ENS Rennes 2024-2025

AGRÉGATION 2025
PLANS ET DÉVELOPPEMENTS

Arthur Maritch-Roy

Introduction

Bienvenue ! Ce document regroupe la majorité du travail que j'ai produit pendant ma préparation à l'agrégation (2024-2025). Il y a d'abord les plans que j'ai rédigés durant l'année (non pas pour les apprendre ensuite par cœur, mais pour rassembler les choses que je pouvais mettre dans la leçon et être sûr de l'ordre dans lequel je voulais les présenter), puis les développements (qui sont plus que des développements souvent et qui contiennent des définitions et résultats préliminaires ainsi que des exemples et des applications). De plus, dans les plans et développements, les choses présentées sont souvent ambitieuses et sont un bilan des choses que je sais pour l'oral. Comme dirait Arnaud, il faut présenter des choses qui sont un peu en-deçà de nos connaissances. Ainsi je jour J, je mettais un peu moins de choses et gardais le reste pour des éventuelles questions du jury, en essayant de l'y guider autant que possible. Je mets aussi mes couplages gardés pour le jour J, mêmes si mes développements peuvent souvent aller dans quelques leçons de plus.

Un conseil que j'aurais aimé avoir est de commencer l'année par se rassurer en revoyant des choses simples d'abord, avec comme ligne de conduite le programme de l'écrit. Ensuite vers la fin de l'année (novembre/décembre), commencer à réfléchir à des thèmes qui vous plaisent qui seront source de développements et de parties de plan. Il est possible d'avoir le meilleur classement avec des choses simples. Les choses compliquées sont bien à mon sens que si elles vous plaisent et qu'elles se recasent très bien. Je vous conseille de choisir vos combats, et de ne pas creuser tous les compléments que vous aurez.

Enfin, si vous avez une question ou qu'il y a une erreur importante dans le document, n'hésitez pas à me contacter via mon adresse de l'ENS :

arthur point maritch tiret roy arobase ens tiret rennes point fr

Bonne lecture et bon courage !

Table des matières

I Plans de leçons d'algèbre **15**

1 101. Groupe opérant sur un ensemble. Exemples et applications. **17**
1 Action d'un groupe sur un ensemble 18
2 Actions classiques d'un groupe sur lui-même 19
3 Actions de groupe et algèbre linéaire et bilinéaire 21
4 Actions de groupe et géométrie 22

2 102. Groupe des nombres complexes de module 1. Racines de l'unité. Applications. **25**
1 Groupes des nombres complexes de module un et racines de l'unité 26
2 Dualité dans les groupes 27
3 Cyclotomie 29

3 103. Conjugaison dans un groupe. Exemples de sous-groupes distingués et de groupes quotients. Applications. **31**
1 Action par conjugaison 32
2 Sous-groupes distingués 33
3 Groupe quotient 34

4 104. Groupes finis. Exemples et applications **37**
1 Généralités sur les groupes finis 38
2 Groupes abéliens finis 39
3 Des actions de groupes pour mieux les étudier 41
4 Groupes de matrices 43

5 105. Groupe des permutations d'un ensemble fini. Applications. **45**
1 Groupe symétrique sur un ensemble fini 45
2 Signature et groupe alterné 47
3 Loi uniforme sur les permutations 47

6 106. Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $\mathrm{GL}(E)$. Applications. **49**
1 Groupe linéaire d'un espace vectoriel 50
2 Actions et sous-groupes 51
3 Groupe orthogonal 53

7 108. Exemples de parties génératrices d'un groupe. Applications. 57
1 Groupe engendré par une partie . . . 57
2 Groupes symétriques et alternés . . . 59
3 Groupe linéaire, groupe orthogonal, quaternions . . . 60

8 120. Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications. 63
1 L'anneau des classes de congruence modulo n . . . 63
2 Arithmétique . . . 65
3 Caractères et transformée de Fourier discrète . . . 66
4 Utilisation de la théorie des corps . . . 68

9 121. Nombres premiers. Applications. 71
1 Les nombres premiers et leur répartition . . . 71
2 Théorie des groupes . . . 73
3 Corps finis et réduction modulo un nombre premier . . . 75

10 122. Anneaux principaux. Exemples et applications. 77
1 Anneaux principaux et euclidiens . . . 77
2 Arithmétique des anneaux principaux . . . 78
3 Réduction de matrices à valeurs dans un anneau principal . . . 80

11 123. Corps finis. Applications. 83
1 Corps finis : cardinal, existence, unicité . . . 83
2 Carrés dans les corps finis . . . 84
3 Polynômes et corps finis . . . 85
4 Matrices sur un corps fini . . . 86

12 125. Extensions de corps. Exemples et applications. 89
1 Extensions de corps, éléments algébriques et transcendants . . . 90
2 Corps de rupture, corps de décomposition . . . 91
3 Corps finis . . . 91
4 Nombres constructibles à la règle et au compas . . . 92

13 127. Exemples de nombres remarquables. Exemples d'anneaux de nombres remarquables. Exemples et applications. 95
1 Nombres premiers . . . 96
2 Entiers de Gauss . . . 97
3 Irrationalité, algébricité, transcendance, approximation diophantienne . . . 98
4 Quaternions . . . 100

14 141. Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications. 103
1 Polynômes irréductibles à coefficients dans un anneau . . . 104
2 Théorie des corps . . . 105
3 Les nombres algébriques sur le corps des rationnels . . . 107

15 142. PGCD et PPCM, algorithmes de calcul. Applications. 109
1 PGCD et PPCM dans un anneau abstrait 109
2 Calcul algorithmique 112
3 Réduction de matrices à valeurs dans un anneau principal 112

16 144. Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications. 115
1 Racines d'un polynôme 115
2 Extensions de corps, nombres algébriques 117
3 Localisation des racines d'un polynôme à coefficients complexes 119

17 148. Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications. 121
1 Dimension d'un espace vectoriel 122
2 Rang 123
3 Topologie des espaces de dimension finie 124
4 Théorie des corps 126

18 149. Déterminant. Exemples et applications. 129
1 Formes multilinéaires et déterminant 129
2 Calcul de déterminant 131
3 Propriétés géométriques et analytiques 133

19 150. Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications. 135
1 L'algèbre des polynômes d'un endomorphisme 135
2 Réduction des endomorphismes 137
3 Exponentielle de matrices 142

20 151. Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications. 145
1 Sous-espaces stables par un endomorphisme 146
2 Diagonalisation, triagonalisation 147
3 Théorèmes de réduction 149

21 152. Endomorphismes diagonalisables en dimension finie. Applications. 151
1 Endomorphismes diagonalisables 152
2 Propriétés selon le corps de base 153
3 Décomposition de Dunford 154
4 Endomorphismes normaux 155

22 153. Valeurs propres, vecteurs propres. Calculs exacts ou approchés d'éléments propres. Applications. 159
1 Valeurs propres, vecteurs propres 159
2 Recherche de valeurs propres 162

3 Matrices symétriques, hermitiennes 163

23 155. Exponentielle de matrices. Applications. 167
1 Exponentielle de matrices 167
2 Exponentielle d'endomorphismes remarquables 168
3 Systèmes différentiels linéaires 169

24 156. Endomorphismes trigonalisables. Endomorphismes nilpotents. 171
1 Endomorphismes nilpotents 171
2 Endomorphismes trigonalisables 172

25 157. Matrices symétriques réelles, matrices hermitiennes. 175
1 Endomorphismes auto-adjoints 175
2 Cas des endomorphismes auto-adjoints positifs, définis positifs 178

26 158. Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie). 181
1 Endomorphismes normaux 182
2 Endomorphismes symétriques 183
3 Endomorphismes orthogonaux 184
4 Décomposition polaire 186

27 159. Formes linéaires et dualité en dimension finie. Exemples et applications. 187
1 Formes linéaires et espace dual 188
2 Orthogonalité 189
3 Application transposée 191
4 Dualité dans les espaces euclidiens et géométrie 193
5 Dualité dans les espaces de matrices 194

28 161. Espaces vectoriels et espaces affines euclidiens : distances, isométries. 197
1 Généralités : applications affines, distances, isométries 198
2 Étude du groupe orthogonal 201
3 Classification des isométries 203
4 Isométries préservant des parties du plan et de l'espace 203

29 162. Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques. 205
1 Généralités sur les systèmes linéaires 205
2 Résolution pratique des systèmes linéaires 208
3 Cas des anneaux principaux 209

30 170. Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité. Applications. 211
1 Formes quadratiques et orthogonalité 211
2 Classification des formes quadratiques 213

31 171. Formes quadratiques réelles. Coniques. Exemples et applications. 217

32 181. Convexité dans $\mathbb{R}^n$. Applications en algèbre et en géométrie. 219

1 Barycentres 219

2 Parties convexes 221

33 190. Méthodes combinatoires, problèmes de dénombrement. 225

1 Dénombrements élémentaires 226

2 Théorie des groupes et algèbre linéaire 227

3 Formule d'inversion de Möbius 228

4 Séries génératrices 229

5 Loi uniforme sur les permutations 230

34 191. Exemples d'utilisation de techniques d'algèbre en géométrie. 233

1 Théorie des groupes et géométrie 234

2 L'algèbre linéaire et le déterminant 235

3 Corps de nombres et opérations géométriques 237

4 Convexité et barycentres 239

5 Constructibilité à la règle et au compas 241

II Plans de leçons d'analyse et probabilités 243

35 201. Espaces de fonctions. Exemples et applications. 245

1 Applications linéaires continues 246

2 Fonctions continues sur compact 247

3 Espaces de Lebesgue 249

4 Espaces de Sobolev en dimension un 251

36 203. Utilisation de la notion de compacité. 253

1 Compacité sur un espace métrique ou vectoriel normé 254

2 Fonctions continues sur un compact 254

3 Compacité en dimension finie 256

4 Compacité dans les espaces de fonctions 257

37 204. Connexité. Exemples et applications. 259

1 Espaces connexes 259

2 Le théorème des valeurs intermédiaires 261

3 Passage du local au global 262

38 205. Espaces complets. Exemples et applications. 265

1 Espaces complets 265

2 Utilisation de la complétude 267

3 Espaces de Hilbert 268

39 206. Exemples d'utilisation de la notion de dimension finie en analyse. 271
1 Utilisation de la compacité . . . 271
2 Analyse hilbertienne . . . 273
3 Calcul différentiel . . . 274

40 208. Espaces vectoriels normés, applications linéaires continues. Exemples. 277
1 Espaces vectoriels normés . . . 277
2 Application linéaires continues . . . 279
3 Espaces de Hilbert . . . 281
4 Parties compactes d'un espace vectoriel normé . . . 283

41 209. Approximation d'une fonction par des fonctions régulières. Exemples d'applications. 285
1 Approximation de fonctions continues sur un compact . . . 285
2 Approximation dans les espaces de Lebesgue . . . 286

42 213. Espaces de Hilbert. Exemples d'applications. 291
1 Espaces préhilbertiens et euclidiens . . . 291
2 Espaces de Hilbert . . . 294
3 Analyse de Fourier . . . 295
4 Espaces de Sobolev en dimension un . . . 297

43 214. Théorème d'inversion locale, théorème des fonctions implicites. illustrations en analyse et en géométrie. 299
1 Le théorème d'inversion locale . . . 300
2 Théorème des fonctions implicites . . . 302
3 Sous-variétés . . . 302

44 215. Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications. 305
1 Différentiabilité . . . 306
2 Résultats fondamentaux . . . 309
3 Optimisation . . . 311

45 218. Formules de Taylor. Exemples et applications. 313
1 Fonctions de la variable réelle . . . 313
2 Suites récurrentes . . . 317
3 En dimension supérieure . . . 318

46 219. Extremums : existence, caractérisation, recherche. Exemples et applications. 321
1 Extremums, cadre général . . . 322
2 Calcul différentiel . . . 323
3 Espaces de Hilbert et méthodes hilbertiennes . . . 324

47 220. Illustrer par des exemples la théorie des équations différentielles ordinaires. 327
1 Théorie de Cauchy-Lipschitz . . . 327
2 Résolution explicite . . . 329
3 Étude qualitative . . . 331

48 221. Équations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications. 333
1 Théorie générale . . . 333
2 Étude qualitative . . . 335
3 Méthodes hilbertiennes et problèmes aux limites . . . 336

49 223. Suites réelles et complexes. Convergence, valeurs d'adhérence. Exemples et applications. 339
1 Convergence de suites numériques . . . 340
2 Séries numériques . . . 341
3 Suites récurrentes . . . 344

50 224. Exemples de développements asymptotiques de suites et de fonctions. 347
1 Développements asymptotiques, développements limités . . . 348
2 Séries numériques . . . 349
3 Étude de suites récurrentes . . . 351
4 Asymptotique d'intégrales à paramètres . . . 352

51 226. Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations. 355
1 Cas d'une fonction de la variable réelle . . . 355
2 Cas des fonctions sur un espace de Banach . . . 358

52 228. Continuité, dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications. 361
1 Continuité d'une fonction de la variable réelle . . . 362
2 Dérivabilité . . . 364
3 Monotonie et convexité . . . 367
4 Suites et séries de fonctions . . . 369

53 229. Fonctions monotones, fonctions convexes. Exemples et applications. 373
1 Fonctions monotones . . . 373
2 Fonctions convexes de la variable réelle . . . 375
3 Cas général . . . 377

54 230. Séries de nombres réelles et complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples. 379
1 Séries numériques . . . 379
2 Séries de fonctions définies à partir d'une série . . . 381

3 Probabilités : séries génératrices, séries de variables aléatoires 384

55 234. Fonctions et espaces de fonctions Lebesgue-intégrables. 387
1 Définition de l'intégrale de Lebesgue 387
2 Espaces de Lebesgue 389
3 Ajout de dérivée, les espaces de Sobolev 391

56 235. Problèmes d'interversion de symboles en analyse. 393
1 Suites et séries de fonctions 393
2 Intégration 396
3 Théorie de Fourier 398

57 236. Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables. 403
1 Méthodes générales de calcul d'intégrales 404
2 Intégrales à paramètres 406
3 Intégration complexe 407
4 Calcul approché d'intégrales 408

58 239. Fonctions définies par un intégrale dépendant d'un paramètre. Exemples et applications. 411
1 Théorèmes de régularité 412
2 Convolution 413
3 Transformée de Fourier 414
4 Asymptotique d'intégrale à paramètres 416

59 241. Suites et séries de fonctions. Exemples et contre-exemples. 419
1 Convergence de suites et séries de fonction 419
2 Séries entières 421
3 Séries de Fourier 423

60 243. Séries entières, propriétés de la somme. Exemples et applications. 427
1 Définition d'une série entière 427
2 Développement en série entière 428
3 Série génératrice 430

61 245. Fonctions holomorphes et méromorphes sur un ouvert de $\mathbb{C}$. Exemples et applications. 433
1 Fonctions holomorphes 433
2 Conséquences du théorème de Cauchy 435
3 Fonctions méromorphes 437

62 246. Séries de Fourier. Exemples et applications. 439
1 Coefficients et séries de Fourier 439
2 Convergence de la série de Fourier 440

3 Utilisation des séries de Fourier 442

63 250. Transformation de Fourier. Applications. 445
1 Transformée de Fourier d'une fonction intégrable 445
2 Classe de Schwartz et prolongement de la transformée de Fourier 447
3 Transformée de Fourier de mesures 448

64 253. Utilisation de la notion de convexité en analyse. 451
1 Ensembles convexes 451
2 Fonctions convexes 453
3 Optimisation, points fixes 454

65 261. Loi d'une variable aléatoire :caractérisations, exemples, applications. 457
1 Loi d'une variable aléatoire 457
2 Caractérisation de la loi 460
3 Loi uniforme sur les permutations 463

66 262. Convergences d'une suite de variables aléatoires. Théorèmes limite. Exemples et applications. 465
1 Convergence en probabilité, convergence presque sûre 466
2 Convergence en norme 467
3 Séries de variables indépendantes 468
4 Convergence en loi 468
5 Statistiques 469

67 264. Variables aléatoires discrètes. Exemples et applications. 473
1 Variables aléatoires discrètes 474
2 Moments et fonction génératrice 475
3 Des exemples des processus aléatoires discrets 476
4 Loi uniforme sur les permutations 477

68 266. Utilisation de la notion d'indépendance en probabilités. 479
1 Notion d'indépendance 480
2 Loi de variables indépendantes 480
3 Vecteurs gaussiens 482
4 Loi uniforme sur les permutations 484

III Développements d'algèbre et géométrie 487

69 Théorème de Perron-Frobenius 489

70 Déterminants de Gram et de Cauchy 493

71 Cycles d'une permutation aléatoire 497

72 Formes quadratiques, décomposition en valeurs singulières **499**

73 Endomorphismes sur un corps fini **505**

74 Loi de réciprocité quadratique **509**

75 Quaternions et matrices orthogonales **513**

76 Théorème de Gauss-Wantzel **515**

77 Sous-groupes compacts de matrices réelles **519**

78 Forme normale de Smith **521**

79 Caractères et marches aléatoires dans les groupes abéliens **527**

80 Matrices inversibles à coefficients entiers **531**

81 Théorème de Dirichlet faible **533**

82 Corps finis et polynômes irréductibles **535**

83 Automorphismes du groupe symétrique **537**

84 Décomposition de Dunford **539**

85 Une équation de Mordell **543**

IV Développement d'analyse et probabilités **545**

86 Fonction gamma et formule des compléments **547**

87 Théorème de Charkovski **551**

88 Théorème de Krein-Milman **555**

89 Théorème de Hadamard-Lévy **559**

90 Lemme de Morse **563**

91 Méthode de Laplace généralisée **565**

92 Indécomposabilité de la loi de Poisson **567**

93 Théorème de Fejér **571**

94 Théorème de Banach-Steinhaus et séries de Fourier **575**

95 Théorème de Liapounov **577**

96 Théorème limite central et extensions aux densités **579**

97 Théorème taubérien fort **585**

98 Inversion de Fourier pour la fonction caractéristique **589**

99 Séries de variables aléatoires **591**

100 Formule sommatoire de Poisson **595**

101 Processus de Galton-Watson **599**

102 Nombre de descentes d'une permutation aléatoire **603**

103 Espaces de Sobolev et problèmes aux limites **607**

104 Optimisation dans un espace de Hilbert **611**

V Couplages **615**

Première partie

Plans de leçons d'algèbre

Chapitre 1

101. Groupe opérant sur un ensemble. Exemples et applications.

Sommaire

1 Action d'un groupe sur un ensemble 18
1.1 Présentation 18
1.2 Orbites et stabilisateurs 18
1.3 Quelques problèmes de dénombrement 19
2 Actions classiques d'un groupe sur lui-même 19
2.1 Action par translation 19
2.2 Action par conjugaison 20
2.3 Théorèmes de Sylow 20
3 Actions de groupe et algèbre linéaire et bilinéaire 21
3.1 Actions à droite et à gauche, équivalence 21
3.2 Similitude, congruence 21
4 Actions de groupe et géométrie 22
4.1 Géométrie affine 22
4.2 Isométries préservant une partie du plan ou de l'espace 23
4.3 Quaternions et isométries de l'espace 23

Présentation du plan

Soient G un groupe et X un ensemble non vide.

1 Action d'un groupe sur un ensemble

1.1 Présentation

Définition. Une *action de groupe* est la donnée d'une opération $G \times X \to X$, notée avec un point et qui vérifie :

— $\forall x \in X,\ 1 \cdot x = x$

— $\forall g, h \in G, \forall x \in X, g \cdot (h \cdot x) = (gh) \cdot x$.

De manière équivalente, cela revient à se donner un morphisme de G dans $\mathfrak{S}_X$ $(g \mapsto (x \mapsto g \cdot x))$.

Définition. Une action est *fidèle* si le morphisme qui lui correspond est injectif, c'est-à-dire si $(g \cdot x = x \forall x)$ implique $(g = 1)$. Une action est *libre* si pour $g \neq 1$, $g \cdot x \neq x$ pour tout $x \in X$. Une action est *transitive* si pour tout couple (x, y) d'éléments de X, il existe $g \in G$ tel que $g \cdot x = y$.

Exemple. Le groupe symétrique $\mathfrak{S}_n$ agit transitivement et fidèlement (mais pas librement) sur $\{1, \ldots, n\}$ par $\sigma \cdot k = \sigma(k)$. Le groupe linéaire $\mathrm{GL}(E)$ agit transitivement et fidèlement (mais pas librement) sur $E \backslash \{0\}$ si E est un espace vectoriel par $P \cdot x = Px$. Le groupe orthogonal $\mathrm{O}_n(\mathbb{R})$ agit fidèlement mais pas transitivement sur $\mathbb{R}^n \backslash \{0\}$. Pour rendre l'action transitive, il faudrait restreindre l'ensemble sur lequel on agit à une sphère.

Remarque. Si G agit sur X alors $G/\ker(\varphi)$ (où φ est l'action en question) agit fidèlement sur X.

Exemple. Le groupe projectif général linéaire $\mathrm{PGL}(E)$ agit fidèlement sur $\mathbb{P}\,E$.

Dans la suite, on suppose donnée une action de G sur X.

1.2 Orbites et stabilisateurs

Définition. La relation $x \sim y \Leftrightarrow \exists g \in G, g \cdot x = y$ est une relation d'équivalence sur X. Les classes d'équivalences pour cette relation sont appelées *orbites*, et forment donc une partition de X.

Définition. Soit $x \in X$ le *stabilisateur* de x est le sous-groupe de G formés des g tels que $g \cdot x = x$, noté $\mathrm{Stab}(x)$.

Remarque. Une action est transitive si et seulement si on n'a qu'une seule orbite.

Exemple. Les orbites pour l'action de $\mathrm{O}_n(\mathbb{R})$ sur $\mathbb{R}^n$ sont les sphères centrées en 0. La partition de $\{1, \ldots, n\}$ en orbites sous l'action des itérées d'une permutation $\sigma \in \mathfrak{S}_n$ donne sa décomposition en produit de cycles à supports disjoints. Le stabilisateur d'un $k \in \{1, \ldots, n\}$ est isomorphe à $\mathfrak{S}_{n-1}$.

Proposition. Le groupe $G/\mathrm{Stab}(x)$ est en bijection avec l'orbite de x, pour tout $x \in X$.

Théorème. (Équation aux classes)
Si X et G sont finis, et si $x_1, \ldots, x_n$ sont des représentants des orbites de X sous l'action de G, alors

$$|X| = \sum_{i=1}^{n} \frac{|G|}{|\mathrm{Stab}(x_i)|}.$$

Théorème. (formule de Burnside)
Toujours dans le cas où X et G sont finis, alors le nombre d'orbites sous l'action de G est égal à

$$\frac{1}{|G|} \sum_{g \in G} |\{x \in X, g \cdot x = x\}|.$$

1.3 Quelques problèmes de dénombrement

Proposition. Le nombre de matrices diagonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|}.$$

Proposition. Le nombre de matrices nilpotentes dans $\mathcal{M}_n(\mathbb{F}_q)$ est $q^{d(d-1)}$.

Proposition. Le nombre de matrices trigonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|} \prod_{i=1}^{q} q^{n_i(n_i-1)}.$$

Proposition. On dispose d'assiettes de c couleurs différentes, à placer sur une table pour n convives. Le nombre de dispositions possibles des assiettes (à rotation près donc) est de $\frac{1}{n}\sum_{d|n}\varphi(d)\,c^{n/d}$.

2 Actions classiques d'un groupe sur lui-même

2.1 Action par translation

Définition. Sur le groupe G, on peut définir une action libre, transitive et fidèle par $g \cdot h := gh$.

Théorème. (de Lagrange)
Soit H un sous-groupe de G. Alors $|G| = |H|[G : H]$. En particulier, le cardinal d'un sous-groupe divise toujours le cardinal d'un groupe.

Théorème. (de Cayley)
Tout groupe fini G est isomorphe à un sous-groupe de $\mathfrak{S}_{|G|}$.

2.2 Action par conjugaison

Définition. L'*action par conjugaison* de G sur lui-même est définie par $g \cdot h := ghg^{-1}$. Pour cette action, le stabilisateur de g est appelé *centralisateur* et correspond aux éléments qui commutent avec g.

Remarque. Le principe de la conjugaison est de fabriquer, à partir d'un g, un nouvel élément avec des caractéristiques similaires à celles de g. De plus, la modification de ces caractéristiques est souvent facile à relier à l'élément par lequel on conjugue. Si $g \in G$ et $x \in X$, on a

$$\mathrm{Stab}(g \cdot x) = g\mathrm{Stab}(x)g^{-1}.$$

Exemple. Dans un espace affine, si t_v désigne la transvection de vecteur v alors pour tout automorphisme affine u et tout vecteur x on a

$$t_{u(x)} = u \circ t_x \circ u^{-1}.$$

Exemple. Soit σ in $\mathfrak{S}_n$ dont la décomposition en produit de cycles à supports disjoints est

$$\sigma = (i_1 \ \dots \ i_k) \dots (i_r \ \dots i_n).$$

Alors pour $\tau \in \mathfrak{S}_n$, on a

$$\tau\sigma\tau^{-1} = (\tau(i_1) \ \dots \ \tau(i_k)) \dots (\tau(i_r) \ \dots \tau(i_n)).$$

Théorème. Tout automorphisme de $\mathfrak{S}_n$, pour $n \neq 6$ est intérieur, *i.e.* est de la forme $\varphi_\tau : \sigma \mapsto \tau\sigma\tau^{-1}$.

Théorème. (de Wedderburn)
Toute algèbre à division finie est un corps (*i.e.* est commutative).

2.3 Théorèmes de Sylow

Proposition. Le centre d'un p-groupe est non trivial.

Définition. Soit G un groupe de cardinal $p^\alpha m$ avec p qui ne divise pas m. Un *p-sous-groupe de Sylow* de G est un sous-groupe de G de cardinal p^α.

Théorème. (Théorèmes de Sylow)
Soit G un groupe fini et p un diviseur premier de $|G|$.

- G contient au moins un p-sous-groupe de Sylow. Plus précisément, pour H un p-sous-groupe de G, il existe un p-Sylow de G qui contient H.
- Les p-Sylow de G sont tous conjugués. Il forment donc une orbite pour l'action par conjugaison, et leur nombre divise $|G|$.
- Le nombre k de p-Sylow de G est congru à 1 modulo p.

Application. Un groupe d'ordre 63 n'est pas simple.

3 Actions de groupe et algèbre linéaire et bilinéaire

3.1 Actions à droite et à gauche, équivalence

Définition. Les groupes $\mathrm{GL}_n(K)$ et $\mathrm{GL}_m(K)^{\mathrm{op}}$ agissent sur $\mathcal{M}_{n,m}(K)$ par multiplication à gauche et à droite respectivement.

Remarque. Le théorème du pivot de Gauss (ou l'algorithme donnant la forme de Hermite si on considère des matrices sur un anneau principal) indique que pour chacune de ces actions, chaque orbite contient une matrice échelonnée (en lignes et en colonnes respectivement).

Définition. On peut aussi faire agir les deux groupes à la fois, et obtenir la relation d'*équivalence* sur $\mathcal{M}_{n,m}(K)$:

$$(P, Q) \cdot M := PMQ^{-1}.$$

Proposition. Il existe exactement $\min(m,n)$ orbites pour cette action, et on peut donner un représentant simple pour chaque orbite (la matrice de taille (n,m) avec un bloc I_r en haut à gauche et des zéros partout ailleurs). Le r qui apparaît alors est le rang de toute matrice dans cette orbite. On obtient donc un invariant sous l'action d'équivalence.

3.2 Similitude, congruence

Définition. Le groupe $\mathrm{GL}_n(K)$ agit sur $\mathcal{M}_n(K)$ par conjugaison $P \cdot M := PMP^{-1}$ et deux matrices sont semblables si et seulement si elles sont dans la même orbite.

Définition. L'analogue de la similitude est la *congruence*, c'est-à-dire le fait que deux matrices représentent la même forme quadratique mais dans deux bases différentes. L'action en question est $M \cdot B \mapsto M^\top BM$ de $\mathrm{GL}_n(K)^{\mathrm{op}}$ dans $\mathcal{S}_n(K)$ (ou éventuellement $\mathcal{S}_n^+(\mathbb{R})$, $\mathcal{S}_n^{++}(\mathbb{R})$ si $K=\mathbb{R}$).

Théorème. (classification des formes quadratiques)
Soit K un corps de caractéristique différente de 2. On s'intéresse aux formes quadratiques sur un K-espace vectoriel de dimension n.

- Si K est algébriquement clos, deux formes quadratiques sont équivalentes si et seulement si elles ont même rang.
- Si $K=\mathbb{R}$, toute forme quadratique est équivalente à $x_1^2+\ldots+x_p^2-x_{p+1}^2-\ldots-x_{p+q}$ avec $p+q\leq n$. Une orbite pour l'action par congruence est déterminée par un couple (p,q) appelé *signature*.
- Si K est fini, une forme quadratique non dégénérée de discriminant $\overline{a}\in K^*/(K^*)^2$ est équivalente à $x_1^2+\ldots+ax_n^2$, on a donc deux classes d'équivalences de formes non dégénérées.

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors il possède un produit scalaire invariant si et seulement s'il est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

4 Actions de groupe et géométrie

4.1 Géométrie affine

Définition. Un *espace affine* est la donnée d'un espace de points E, d'un espace vectoriel V qui agit fidèlement, librement et transitivement sur E (en tant que groupe abélien).

Remarque. Lorsque l'on travaille sur un espace vectoriel, on donne au vecteur nul un rôle particulier. En géométrie affine, on n'a pas cela, et cela est dû à la définition même d'un espace affine. En faisant agir notre espace vectoriel sur un espace de points, on donne alors à chaque point "le même rôle" et chaque point peut être une origine.

Proposition. Soit E un espace affine euclidien, et $\mathcal{X}$ un ensemble de parties de E de la forme $\mathcal{X}=\{s(X_0), s \text{ similitude de } E, X_0\subset E\}$. Soit φ une isométrie affine qui préserve X_0 ($\varphi(X_0)=X_0$), alors pour $X:=s(X_0)$, $s\circ\varphi\circ s^{-1}$ est une isométrie affine qui préserve X. On peut donc parler du groupe des isométries affines qui laissent invariant $\mathcal{X}$ plutôt que de parler des isométries affines

laissant stable X_0. Inversement, traiter le cas de X_0 permet de traiter le cas de tout élément de $\mathcal{X}$.

Définition. Si G est un sous-groupe des isométries d'un espace affine, on peut le faire agir sur les parties de cet espace affine par $u \cdot X := u(X)$.

Définition. Pour une partie X de l'espace affine $\mathbb{R}^n$, on note $\mathrm{Is}(X)$ (resp. $\mathrm{Is}^+(X)$) l'ensemble des isométries affines (resp. déplacements) f de $\mathbb{R}^n$ telles que $f(X) = X$. Ces deux ensembles forment deux sous-groupes de $\mathrm{Isom}(\mathbb{R}^n)$.

Proposition. Si X est une partie convexe, alors les éléments de $\mathrm{Is}(X)$ permutent les points extrémaux de X.

4.2 Isométries préservant une partie du plan ou de l'espace

Définition. Pour $n \geq 3$, on note D_n le groupe $\mathrm{Is}(P_n)$, où P_n est le polygone régulier à n côtés centré en 0 et dont 1 est l'un des côtés. On l'appelle *groupe diédral* d'ordre $2n$.

Théorème. D_n est un groupe de cardinal $2n$, engendré par r la rotation d'angle $\frac{2\pi}{n}$ et la symétrie orthogonale s d'axe $\mathbb{R}(1,0)$, avec la relation $(rs)^2 = \mathrm{id}$. De plus, $D_n^+ := \mathrm{Is}^+(P_n) \simeq \mathbb{Z}/n\mathbb{Z}$.

On note T le tétraèdre régulier et C le cube de $\mathbb{R}^3$.

Théorème. Le groupe $\mathrm{Is}(T)$ est isomorphe à $\mathfrak{S}_4$ et $\mathrm{Is}^+(T)$ à $\mathfrak{A}_4$.

Théorème. Le groupe $\mathrm{Is}^+(C)$ est isomorphe à $\mathfrak{S}_4$ et le groupe $\mathrm{Is}(C)$ est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

4.3 Quaternions et isométries de l'espace

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a,b,c,d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_n$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Chapitre 2

102. Groupe des nombres complexes de module 1. Racines de l'unité. Applications.

Sommaire

1 Groupes des nombres complexes de module un et racines de l'unité 26

1.1 Exponentielle complexe 26

1.2 Relèvement de l'angle 26

1.3 Racines de l'unité 27

2 Dualité dans les groupes 27

2.1 Caractères et groupe dual 27

2.2 Structure des groupes abéliens finis 28

2.3 Transformée de Fourier sur un groupe abélien fini 28

3 Cyclotomie 29

3.1 Polynômes cyclotomiques 29

3.2 Irréductibilité des polynômes cyclotomiques 30

Présentation du plan

1 Groupes des nombres complexes de module un et racines de l'unité

1.1 Exponentielle complexe

Définition. On définit l'*exponentielle complexe* par la série entière suivante, de rayon de convergence infini :

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}.$$

Remarque. On voit que exp est égale à sa dérivée.

Proposition. L'exponentielle complexe réalise un morphisme de groupes surjectif de $(\mathbb{C}, +)$ dans $(\mathbb{C}^*, \times)$. Si on note $\mathrm{e} := \exp(1)$, cela justifie la notation $\exp(z) = \mathrm{e}^z$.

Proposition. Le noyau de exp est inclus dans $\mathrm{i}\mathbb{R}$. Ainsi, il est de la forme $\mathrm{i}a\mathbb{Z}$ pour un $a > 0$. On définit $\pi := \frac{a}{2}$.

Définition. Pour $z \in \mathbb{C}$, on définit $\cos(z)$ et $\sin(z)$ comme les parties réelle et imaginaire de $\exp(z)$. Ce sont des fonctions 2π périodiques de $\mathbb{R}$ dans $[-1, 1]$, et on a toutes les formules habituelles.

Proposition. En particulier, on a les formules de Moivre et d'Euler.

1.2 Relèvement de l'angle

Définition. On note $\mathbb{U}$ l'ensemble des nombres complexes de module un.

Proposition. On a un isomorphisme de groupes entre $\mathbb{R}/(2\pi\mathbb{Z})$ et $\mathbb{U}$. On définit l'*argument* de $z \in \mathbb{C}^*$, noté $\arg(z)$, l'unique $\theta \in [0, 2\pi[$ tel que $z = |z|\mathrm{e}^{\mathrm{i}\theta}$. On obtient ainsi la décomposition polaire de z.

Remarque. On peut aussi voir l'isomorphisme avec $\mathrm{SO}_2(\mathbb{R})$ ($\theta \mapsto R_\theta$). De plus, on peut voir que c'est un homéomorphisme. En calculant $\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x$, on retrouve la formule de l'aire d'un cercle. On peut faire pareil pour les périmètres.

Théorème. (de relèvement $\mathcal{C}^1$)
Soit γ un chemin $\mathcal{C}^1$ tracé de I dans $\mathbb{C}^*$. Alors il existe une fonction $\rho \in \mathcal{C}^1(I, \mathbb{R}_+^*)$ et une fonction $\theta \in \mathcal{C}^1(I, \mathbb{R})$ telles que $\gamma(t) = \rho(t)\mathrm{e}^{\mathrm{i}\theta(t)}$.

1.3 Racines de l'unité

Proposition. Soit $n \geq 1$, il existe un unique sous-groupe de $(\mathbb{C}^*, \times)$ de cardinal n, qui est cyclique et égal à

$$\mathbb{U}_n = \{\mathrm{e}^{\frac{2\mathrm{i}k\pi}{n}}, 0 \leq k \leq n-1\}.$$

Ce sont les racines de $X^n - 1$ sur $\mathbb{C}$.

Définition. On appelle *racine de l'unité* tout élément de l'un des $\mathbb{U}_n$.

Remarque. L'ensemble des racines de l'unité est un groupe dont chaque élément possède un ordre fini, sans être pour autant d'exposant fini.

Proposition. Soit M une matrice d'ordre fini d de $\mathrm{GL}_n(\mathbb{C})$. Alors les valeurs propres de M sont des racines d-ièmes de l'unité.

Application. L'ensemble des ordres possibles pour une matrice d'ordre fini sur $\mathrm{GL}_n(\mathbb{Z})$ est fini.

Définition. Soit $n \geq 1$. Une racine *primitive* n-ième de l'unité est un générateur de $\mathbb{U}_n$, *i.e.* un élément de la forme $\mathrm{e}^{\frac{2\mathrm{i}k\pi}{n}}$ avec $k \wedge n = 1$.

Proposition. En partitionnant $\mathbb{U}_n$ selon ses générateurs, on obtient, avec la notation

$$\mathcal{P}_d = \{\text{racines primitives } d\text{-ièmes de l'unité}\}$$

, la relation :

$$\mathbb{U}_n = \bigcup_{d|n} \mathcal{P}_d.$$

En particulier, $n = \prod_{d|n} \varphi(d)$ si φ est l'indicatrice d'Euler.

Application. Soit K un corps fini. Alors $(K^*, \times)$ est un groupe cyclique.

2 Dualité dans les groupes

2.1 Caractères et groupe dual

Définition. Soit G un groupe fini. On appelle *caractère* de G tout morphisme de groupes de G dans $(\mathbb{C}^*, \times)$.

Définition. Pour G un groupe fini, on note $\widehat{G}$ le *groupe dual* de G, défini comme l'ensemble des caractères de G.

Remarque. Si n est l'ordre de G, tout caractère de G prend ses valeurs dans $\mathbb{U}_n$.

Exemple. Le groupe dual de $\mathfrak{S}_n$ est $\{1, \varepsilon\}$ pour tout n.

Exemple. Un groupe cyclique est isomorphe à son dual.

Proposition. Le groupe dual d'un groupe fini G est un groupe abélien fini.

2.2 Structure des groupes abéliens finis

Proposition. Soit G un groupe abélien fini et H un sous-groupe de G. Tout caractère de H peut être prolongé à un caractère de G.

Théorème. (de structure des groupes abéliens finis)
Soit G un groupe abélien fini non trivial. Il existe un unique entier $r \geq 1$ et des uniques $n_r | n_{r-1} | \dots | n_1$ supérieurs à 2 tels que

$$G \simeq \mathbb{U}_{n_1} \times \cdots \times \mathbb{U}_{n_r}.$$

Application. Tout groupe abélien fini est isomorphe à son dual et canoniquement isomorphe à son bidual par $g \mapsto (\chi \mapsto \chi(g))$ par le théorème de structure des groupes abéliens finis.

2.3 Transformée de Fourier sur un groupe abélien fini

Définition. Soit H un groupe abélien fini, on peut définir un produit scalaire sur $\mathbb{C}^H$ par la formule :

$$\langle f_1, f_2 \rangle_H = \frac{1}{|H|} \sum_{h \in H} \overline{f_1(h)} f_2(h).$$

Définition. Soit $f \in \mathbb{C}^G$. On définit $\widehat{f} \in \mathbb{C}^{\widehat{G}}$ par la formule :

$$\widehat{f}(\chi) := \sum_{g \in G} \overline{\chi(g)} f(g) = |G| \langle \chi, f \rangle_G.$$

Proposition. Les caractères forment une base orthonormée de $\mathbb{C}^G$.

Proposition. (Égalité de Bessel-Parseval)
Pour toute fonction $f \in \mathbb{C}^G$, puisque la famille $\widehat{G}$ est une base orthonormée, on a

$$\sum_{g \in G} |f(g)|^2 = |G| \sum_{g \in G} |\langle \chi, f \rangle|^2 = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|^2.$$

Proposition. (Inversion de Fourier)
Soit $f \in \mathbb{C}^G$. Alors pour tout $g \in G$:

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \chi(g) \widehat{f}(\chi)$$

Définition. On peut définir un produit de convolution (associatif, commutatif et distributif par rapport à $+$) sur $\mathbb{C}^H$ pour H un groupe abélien fini par la formule suivante :

$$(f_1 * f_2)(h) := \sum_{k \in G} f_1(k) f_2(hk^{-1}).$$

Proposition. Soient f_1 et f_2 deux fonctions de $\mathbb{C}^G$. Alors $\widehat{f_1 * f_2} = \widehat{f_1} \widehat{f_2}$.

Théorème. On considère une particule sur $\mathbb{Z}/N\mathbb{Z} = (\overline{0}, \ldots, \overline{N-1})$, initialement en $\overline{0}$. À chaque étape, la particule se déplace d'un cran vers la droite ou vers la gauche ou stationne de manière équiprobable. On peut modéliser cela par une chaîne de Markov dont la loi initiale est $\delta_{\overline{0}}$ et dont la loi de transition est $\mu = \frac{1}{3}(\delta_{-\overline{1}} + \delta_{\overline{0}} + \delta_{\overline{1}})$. La loi de X_n la position de la particule au rang n a pour loi

$$\mu_n = \mu_0 * \mu^{*n} = \mu^{*n}.$$

La chaîne de Markov en question converge vers la loi uniforme notée μ_∞ et on a une approximation de la vitesse de convergence en variation totale :

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N} \mathrm{e}^{-\frac{\pi^2 n}{N^2}}.$$

3 Cyclotomie

3.1 Polynômes cyclotomiques

Définition. Soit $n \geq 1$. Le n-ième *polynôme cyclotomique* est le polynôme

$$\Phi_n(X) = \prod_{\zeta \in \mathcal{P}_n} (X - \zeta) \in \mathbb{C}[X].$$

Proposition. Le polynôme Φ_n est unitaire, de degré $\varphi(n)$.

Exemple. Pour p premier, $\Phi_p(X) = \frac{X^p-1}{X-1}$.

Proposition. Soit $n \geq 1$, alors

$$X^n - 1 = \prod_{d|n} \Phi_d(X).$$

Par inversion de Möbius, on a alors

$$\Phi_n(X) = \prod_{d|n} (X^d - 1)^{\mu(n/d)}$$

Corollaire. Les polynômes cyclotomiques sont à coefficients entiers.

Théorème. (progression arithmétique de Dirichlet, version faible)
Soit $n \geq 1$. Il existe une infinité de nombres premiers congrus à 1 modulo n.

3.2 Irréductibilité des polynômes cyclotomiques

Théorème. Les polynômes cyclotomiques sont irréductibles sur $\mathbb{Q}[X]$.

Application. Le polynôme $X^4 + 1$ est irréductible sur $\mathbb{Q}[X]$ mais n'est irréductible sur aucun corps fini.

Chapitre 3

103. Conjugaison dans un groupe. Exemples de sous-groupes distingués et de groupes quotients. Applications.

Sommaire

1 Action par conjugaison 32
1.1 Définition 32
1.2 Principe de conjugaison 32
1.3 Cas du groupe symétrique 32
2 Sous-groupes distingués 33
2.1 Définition 33
2.2 Centre et groupe dérivé 33
2.3 Théorèmes de Sylow 34
2.4 Groupes simples 34
3 Groupe quotient 34
3.1 Construction et isomorphismes 34
3.2 Cas des quaternions 35

Présentation du plan

Dans tout le plan, G est un groupe.

1 Action par conjugaison

1.1 Définition

Définition. Le groupe G agit sur lui-même par *automorphismes intérieurs* ou par *conjugaison* :

$$g \cdot h := ghg^{-1}.$$

Pour cette action, le stabilisateur de $h \in G$ est appelée *centralisateur* noté $\mathrm{Z}_G(h)$ de h et son orbite est sa *classe de conjugaison*, notée $\mathrm{C}_G(h)$.

Remarque. On agit par automorphismes intérieurs sur G, mais tous les automorphismes de G ne sont pas intérieurs, comme par exemple $A \mapsto (A^{-1})^\top$ sur $\mathrm{GL}_n(\mathbb{R})$.

Remarque. Dans un groupe abélien, la conjugaison ne fait rien.

Proposition. On suppose G fini. Soient $x_1, \ldots, x_n$ un représentant de chaque orbite disjointe au centre, on a

$$|G| = |\mathrm{Z}(G)| + \sum_{i=1}^{n} \frac{|G|}{|\mathrm{Z}(x_i)|}.$$

Application. Le centre d'un p-groupe est non trivial. En particulier, tout groupe d'ordre p^2 est abélien (p désigne un nombre premier).

1.2 Principe de conjugaison

Remarque. Le conjugué ghg^{-1} de h a même nature géométrique que h, mais ses éléments caractéristiques sont l'image par g de ceux de h.

Exemple. Dans un espace affine, si t_v désigne la transvection de vecteur v alors pour tout automorphisme affine u et tout vecteur x on a

$$t_{u(x)} = u \circ t_x \circ u^{-1}.$$

Exemple. Conjuguer dans un espace de matrice donne une matrice qui représente le même endomorphisme. La base dans laquelle il faut se placer pour obtenir cette nouvelle forme est alors la matrice par laquelle on conjugue.

1.3 Cas du groupe symétrique

Proposition. Soit $\sigma \in \mathfrak{S}_n$ et $(i_1 \ \ldots \ i_r)$ un cycle de $\mathfrak{S}_n$. Alors :

$$\sigma \circ (i_1 \ \ldots \ i_r) \circ \sigma^{-1} = (\sigma(i_1) \ \ldots \ \sigma(i_r))$$

Définition. Soit $\sigma \in \mathfrak{S}_n$ on définit son *type* comme l'ensemble des longueurs de ses cycles.

Proposition. Deux permutations sont conjuguées si et seulement si elles ont même type.

Théorème. Tout automorphisme de $\mathfrak{S}_n$, pour $n \neq 6$ est intérieur, *i.e.* est de la forme $\varphi_\tau : \sigma \mapsto \tau\sigma\tau^{-1}$.

2 Sous-groupes distingués

2.1 Définition

Définition. Un sous-groupe H de G est dit *distingué* (ou *normal*) dans G et on note $H \triangleleft G$.

Exemple. Un sous-groupe d'indice 2 est toujours distingué. Si G' est un autre groupe, alors $G \times \{1_{G'}\} \triangleleft G \times G'$.

Proposition. (lemme d'Ore)
Un sous-groupe d'indice p où p est le plus petit nombre premier qui divise $|G|$ est toujours distingué dans G.

Application. Un groupe d'ordre p^2 est isomorphe à $(\mathbb{Z}/p\mathbb{Z})^2$ ou a $\mathbb{Z}/p^2\mathbb{Z}$

Remarque. Soit H un sous-groupe de G. Il est distingué si et seulement si c'est une union de classes de conjugaison.

Proposition. Le noyau d'un morphisme est toujours distingué. Ainsi, $\mathfrak{A}_n \triangleleft \mathfrak{S}_n$ et $\mathrm{SL}_n(K) \triangleleft \mathrm{GL}_n(K)$.

2.2 Centre et groupe dérivé

Proposition. Le centre d'un groupe est toujours distingué dans ce groupe.

Définition. Soient $g, h \in G$, on définit leur *commutateur* par $[g, h] := ghg^{-1}h^{-1}$.

Remarque. Si $H \triangleleft G$ et $h \in H$, $g \in G$ alors $[g, h] \in H$.

Définition. On définit le *groupe dérivé* de G comme le sous-groupe de G engendré par les commutateurs. C'est un sous-groupe distingué dans G, noté $\mathrm{D}(G)$.

Exemple. Le groupe dérivé de $\mathfrak{A}_4$ est le groupe des doubles transpositions, isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$.

Proposition. Le groupe dérivé de $\mathrm{GL}_n(K)$ est $\mathrm{SL}_n(K)$ dès que n est supérieur à 3.

2.3 Théorèmes de Sylow

Définition. Soit G un groupe de cardinal $p^\alpha m$ avec p qui ne divise pas m. Un *p-sous-groupe de Sylow* de G est un sous-groupe de G de cardinal p^α.

Théorème. (Théorèmes de Sylow)
Soit G un groupe fini et p un diviseur premier de $|G|$.

- G contient au moins un p-sous-groupe de Sylow. Plus précisément, pour H un p-sous-groupe de G, il existe un p-Sylow de G qui contient H.
- Les p-Sylow de G sont tous conjugués. Il forment donc une orbite pour l'action par conjugaison, et leur nombre divise $|G|$.
- Le nombre k de p-Sylow de G est congru à 1 modulo p.

Application. Un groupe d'ordre 63 n'est pas simple.

Application. Le groupe $\mathfrak{S}_6$ possède un automorphisme non intérieur.

2.4 Groupes simples

Définition. Un groupe *simple* est un groupe dont tous les sous-groupes distingués sont $\{1\}$ et lui-même.

Exemple. Les seuls groupes simples abéliens sont les $\mathbb{Z}/p\mathbb{Z}$ avec p premier.

Théorème. Le groupe $\mathfrak{A}_n$ est simple pour $n \geq 5$.

Théorème. Le groupe $\mathrm{SO}_3(\mathbb{R})$ est simple.

3 Groupe quotient

3.1 Construction et isomorphismes

Proposition. Soit H un sous-groupe distingué dans G. Il existe une unique structure de groupe sur G/H qui fait de la projection canonique un morphisme de groupes.

Remarque. Dans ce cas, le noyau de la projection est H, ce qui montre que si on a une telle structure de groupe sur un quotient, le sous-groupe par lequel on quotiente est nécessairement normal dans G.

Remarque. On retrouve avec ce formalisme la notation $\mathbb{Z}/n\mathbb{Z}$.

Théorème. Soit $f : G \to G'$ un morphisme de groupes. Soit $H \triangleleft G$ inclus dans $\ker(f)$. Alors il existe un unique morphisme $\overline{f} : G/H \to G'$ qui passe au quotient, c'est-à-dire tel que $\overline{f} \circ \pi = f$, où π est la projection canonique sur H. Si f est surjectif, alors $\overline{f}$ l'est aussi. Si $\ker(f) = H$, il est injectif. En particulier, $G/\ker(f) \simeq \mathrm{Im}(f)$.

Théorème. Soit H un sous-groupe normal dans G. La projection canonique induit une bijection entre les sous-groupes distingués de G contenant H et les sous-groupes distingués de G/H.

Application. Un p-groupe vérifie la réciproque de Lagrange.

3.2 Cas des quaternions

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a, b, c, d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_n$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Chapitre 4

104. Groupes finis. Exemples et applications

Sommaire

1 Généralités sur les groupes finis 38
1.1 Le théorème de Lagrange 38
1.2 Groupes cycliques 38
1.3 Groupes symétriques 39
2 Groupes abéliens finis 39
2.1 Dualité dans les groupes finis 39
2.2 Classification des groupes abéliens finis 40
2.3 Transformée de Fourier rapide 40
3 Des actions de groupes pour mieux les étudier 41
3.1 Généralités et dénombrement 41
3.2 Théorèmes de Sylow 42
3.3 Vision de certains groupes de manière géométrique 42
4 Groupes de matrices 43
4.1 Sur un corps quelconque 43
4.2 Sur le corps des complexes 43
4.3 Sur l'anneau des entiers 43

Présentation du plan

Dans tout le plan, G est un groupe fini, *i.e.* de cardinal fini.

1 Généralités sur les groupes finis

1.1 Le théorème de Lagrange

Définition. On appelle *ordre* d'un groupe fini son cardinal. L'ordre d'un élément $x \in G$ est le cardinal du sous-groupe engendré par x.

Définition. Sur G, on peut définir une relation d'équivalence associée à un sous-groupe H par $x \sim y \Leftrightarrow xy^{-1} \in H$. L'ensemble des classes d'équivalence est appelé *quotient* de G sur H, noté G/H. Son cardinal est appelée *indice* de H dans G, noté $[G : H]$.

Théorème. (de Lagrange)
Soit H un sous-groupe de G. Alors $|H|[G : H] = |G|$. En particulier, le cardinal d'un sous-groupe divise le cardinal du sous-groupe.

Définition. Un sous-groupe H de G est dit *distingué* dans G s'il est stable par tout automorphisme intérieur, c'est-à-dire si pour tout $h \in H$ et $g \in G$, $ghg^{-1} \in G$.

Proposition. On peut munir G/H d'une structure de groupe qui rend la projection canonique continue si et seulement si H est un sous-groupe distingué de G.

Exemple. En particulier, dans un groupe abélien, tous les sous-groupes sont distingués donc on peut toujours quotienter pour obtenir de nouveaux groupes. On peut par exemple construire $\mathbb{Z}/n\mathbb{Z}$ comme ensemble des classes d'équivalence modulo n.

Définition. Un groupe est simple si ses seuls sous-groupes distingués sont $\{1\}$ et le groupe lui-même.

1.2 Groupes cycliques

Définition. Un groupe *cyclique* est un groupe monogène et fini.

Proposition. Un groupe cyclique G est isomorphe à $(\mathbb{Z}/|G|\mathbb{Z})$ ou de manière équivalente à $\mathbb{U}_{|G|}$.

Proposition. Dans un groupe cyclique d'ordre d, il y a $\varphi(d)$ générateurs d'ordre d, et tout sous-groupe d'un groupe cyclique est cyclique.

Proposition. Un groupe cyclique vérifie la réciproque du théorème de Lagrange. Plus précisément, pour $d||G|$, G possède un unique sous-groupe d'ordre d.

Proposition. Un groupe de cardinal premier est cyclique.

Théorème. Tout sous-groupe multiplicatif fini d'un corps est cyclique.

1.3 Groupes symétriques

Définition. Soit X un ensemble. Le *groupe symétrique* sur X est le groupe des bijections de X muni de la composition, noté $\mathfrak{S}_X$. Dans le cas où X est fini, c'est un groupe fini de cardinal $|X|!$. On peut simplifier en le voyant comme le groupe des permutations de $\{1, \ldots, n\}$, noté $\mathfrak{S}_n$ où $n = |X|$.

Définition. Une permutation c de $\mathfrak{S}_n$ telle qu'il existe $\{j_1, \ldots, j_k\} = J \subset \{1, \ldots, n\}$ tel que $c(j_i) = j_{i+1}$ pour $i \in \{1, \ldots, k-1\}$, $c(j_k) = j_1$ et $c(l) = l$ si $l \notin J$ est un *cycle*, et J est son support. Une *transposition* est un cycle dont le support est de cardinal 2.

Proposition. Toute permutation est produit unique (à permutation du produit près) de cycles à supports disjoints. En particulier, $\mathfrak{S}_n$ est engendré par les transpositions.

Proposition. L'ordre d'une permutation est le ppcm des longueurs des cycles présents dans sa décomposition en produit de cycles à supports disjoints.

Définition. Il existe un unique morphisme de groupes de $\mathfrak{S}_n$ vers $\mathbb{C}^*$, noté ε et appelé *signature*. Il correspond à $(-1)^N$, où N est le nombre de transpositions présentes dans la décomposition d'une permutation en un produit de transpositions.

Définition. On appelle *groupe alterné* le noyau de la signature, et on le note $\mathfrak{A}_n$. C'est le seul groupe d'indice 2 de $\mathfrak{S}_n$.

Proposition. Les cycles de longueur 3 sont conjugués dans $\mathfrak{A}_n$ et engendrent $\mathfrak{A}_n$.

Théorème. Le groupe alterné $\mathfrak{A}_n$ est simple pour $n \geq 5$.

2 Groupes abéliens finis

2.1 Dualité dans les groupes finis

Définition. Soit G un groupe fini. On appelle *caractère* de G tout morphisme de groupes de G dans $(\mathbb{C}^*, \times)$.

Définition. Pour G un groupe fini, on note $\widehat{G}$ le *groupe dual* de G, défini comme l'ensemble des caractères de G.

Remarque. Si n est l'ordre de G, tout caractère de G prend ses valeurs dans $\mathbb{U}_n$.

Exemple. Le groupe dual de $\mathfrak{S}_n$ est $\{1, \varepsilon\}$ pour tout n.

Exemple. Un groupe cyclique est isomorphe à son dual.

Proposition. Le groupe dual d'un groupe fini G est un groupe abélien fini.

2.2 Classification des groupes abéliens finis

Proposition. Soit G un groupe abélien fini et H un sous-groupe de G. Tout caractère de H peut être prolongé à un caractère de G.

Théorème. (de structure des groupes abéliens finis)
Soit G un groupe abélien fini non trivial. Il existe un unique entier $r \geq 1$ et des uniques $n_r | n_{r-1} | \dots | n_1$ supérieurs à 2 tels que

$$G \simeq \mathbb{U}_{n_1} \times \cdots \times \mathbb{U}_{n_r}.$$

Application. Tout groupe abélien fini est isomorphe à son dual et canoniquement isomorphe à son bidual par $g \mapsto (\chi \mapsto \chi(g))$ par le théorème de structure des groupes abéliens finis.

2.3 Transformée de Fourier rapide

Définition. Soit H un groupe abélien fini, on peut définir un produit scalaire sur $\mathbb{C}^H$ par la formule :

$$\langle f_1, f_2 \rangle_H = \frac{1}{|H|} \sum_{h \in H} \overline{f_1(h)} f_2(h).$$

Définition. Soit $f \in \mathbb{C}^G$. On définit $\widehat{f} \in \mathbb{C}^{\widehat{G}}$ par la formule :

$$\widehat{f}(\chi) := \sum_{g \in G} \overline{\chi(g)} f(g) = |G| \langle \chi, f \rangle_G.$$

Proposition. Les caractères forment une base orthonormée de $\mathbb{C}^G$.

Proposition. (Égalité de Bessel-Parseval)
Pour toute fonction $f \in \mathbb{C}^G$, puisque la famille $\widehat{G}$ est une base orthonormée, on a

$$\sum_{g \in G} |f(g)|^2 = |G| \sum_{g \in G} |\langle \chi, f \rangle|^2 = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|^2.$$

Proposition. (Inversion de Fourier)
Soit $f \in \mathbb{C}^G$. Alors pour tout $g \in G$:

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \chi(g) \widehat{f}(\chi)$$

Définition. On peut définir un produit de convolution (associatif, commutatif et distributif par rapport à $+$) sur $\mathbb{C}^H$ pour H un groupe abélien fini par la formule suivante :

$$(f_1 * f_2)(h) := \sum_{k \in G} f_1(k) f_2(hk^{-1}).$$

Proposition. Soient f_1 et f_2 deux fonctions de $\mathbb{C}^G$. Alors $\widehat{f_1 * f_2} = \widehat{f_1} \widehat{f_2}$.

Théorème. On considère une particule sur $\mathbb{Z}/N\mathbb{Z} = (\overline{0}, \ldots, \overline{N-1})$, initialement en $\overline{0}$. À chaque étape, la particule se déplace d'un cran vers la droite ou vers la gauche ou stationne de manière équiprobable. On peut modéliser cela par une chaîne de Markov dont la loi initiale est $\delta_{\overline{0}}$ et dont la loi de transition est $\mu = \frac{1}{3}(\delta_{\overline{-1}} + \delta_{\overline{0}} + \delta_{\overline{1}})$. La loi de X_n la position de la particule au rang n a pour loi

$$\mu_n = \mu_0 * \mu^{*n} = \mu^{*n}.$$

La chaîne de Markov en question converge vers la loi uniforme notée μ_∞ et on a une approximation de la vitesse de convergence en variation totale :

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N} \mathrm{e}^{-\frac{\pi^2 n}{N^2}}.$$

3 Des actions de groupes pour mieux les étudier

3.1 Généralités et dénombrement

Définition. Une *action de groupe* est la donnée d'une opération $G \times X \to X$, notée avec un point et qui vérifie :

— $\forall x \in X,\ 1 \cdot x = x$

— $\forall g, h \in G, \forall x \in X, g \cdot (h \cdot x) = (gh) \cdot x$.

De manière équivalente, cela revient à se donner un morphisme de G dans $\mathfrak{S}_X$ $(g \mapsto (x \mapsto g \cdot x))$.

La relation $x \sim y \Leftrightarrow \exists g \in G, g \cdot x = y$ est une relation d'équivalence sur X. Les classes d'équivalences pour cette relation sont appelées *orbites*, et forment donc une partition de X.
Soit $x \in X$ le *stabilisateur* de x est le sous-groupe de G formés des g tels que $g \cdot x = x$, noté $\mathrm{Stab}(x)$.

Exemple. Le groupe G peut agir sur lui-même par *translation* avec $g \cdot h := gh$.

Application. Tout groupe fini de cardinal n est isomorphe à un sous-groupe de $\mathfrak{S}_n$.

Théorème. (Équation aux classes)
Si X et G sont finis, et si $x_1, \ldots, x_n$ sont des représentants des orbites de X sous l'action de G, alors

$$|X| = \sum_{i=1}^{n} \frac{|G|}{|\mathrm{Stab}(x_i)|}.$$

Théorème. (formule de Burnside)
Toujours dans le cas où X et G sont finis, alors le nombre d'orbites sous l'action de G est égal à

$$\frac{1}{|G|} \sum_{g \in G} |\{x \in X, g \cdot x = x\}|.$$

Application. On dispose d'assiettes de c couleurs différentes, à placer sur une table pour n convives. Le nombre de dispositions possibles des assiettes (à rotation près donc) est de $\frac{1}{n} \sum_{d|n} \varphi(d)\, c^{n/d}$.

Définition. Le groupe G agit sur lui-même par *automorphismes intérieurs* ou par *conjugaison* :

$$g \cdot h := ghg^{-1}.$$

Pour cette action, le stabilisateur de $h \in G$ est appelée *centralisateur* noté $\mathrm{Z}_G(h)$ de h et son orbite est sa *classe de conjugaison*, notée $\mathrm{C}_G(h)$.

Proposition. On suppose G fini. Soient $x_1, \ldots, x_n$ un représentant de chaque orbite disjointe au centre, on a

$$|G| = |\mathrm{Z}(G)| + \sum_{i=1}^{n} \frac{|G|}{|\mathrm{Z}(x_i)|}.$$

Application. Le centre d'un p-groupe est non trivial. En particulier, tout groupe d'ordre p^2 est abélien (p désigne un nombre premier).

3.2 Théorèmes de Sylow

Définition. Soit G un groupe de cardinal $p^\alpha m$ avec p qui ne divise pas m. Un *p-sous-groupe de Sylow* de G est un sous-groupe de G de cardinal p^α.

Théorème. (Théorèmes de Sylow)
Soit G un groupe fini et p un diviseur premier de $|G|$.

- G contient au moins un p-sous-groupe de Sylow. Plus précisément, pour H un p-sous-groupe de G, il existe un p-Sylow de G qui contient H.
- Les p-Sylow de G sont tous conjugués. Il forment donc une orbite pour l'action par conjugaison, et leur nombre divise $|G|$.
- Le nombre k de p-Sylow de G est congru à 1 modulo p.

Application. Un groupe d'ordre 63 n'est pas simple.

3.3 Vision de certains groupes de manière géométrique

Définition. Si G est un sous-groupe des isométries d'un espace affine, on peut le faire agir sur les parties de cet espace affine par $u \cdot X := u(X)$.

Définition. Pour une partie X de l'espace affine $\mathbb{R}^n$, on note $\mathrm{Is}(X)$ (resp. $\mathrm{Is}^+(X)$) l'ensemble des isométries affines (resp. déplacements) f de $\mathbb{R}^n$ telles que $f(X) = X$. Ces deux ensembles forment deux sous-groupes de $\mathrm{Isom}(\mathbb{R}^n)$.

Proposition. Si X est une partie convexe, alors les éléments de $\mathrm{Is}(X)$ permutent les points extrémaux de X.

Définition. Pour $n \geq 3$, on note D_n le groupe $\mathrm{Is}(P_n)$, où P_n est le polygone régulier à n côtés centré en 0 et dont 1 est l'un des côtés. On l'appelle *groupe diédral* d'ordre $2n$.

Théorème. D_n est un groupe de cardinal $2n$, engendré par r la rotation d'angle $\frac{2\pi}{n}$ et la symétrie orthogonale s d'axe $\mathbb{R}(1,0)$, avec la relation $(rs)^2 = \mathrm{id}$. De plus, $D_n^+ := \mathrm{Is}^+(P_n) \simeq \mathbb{Z}/n\mathbb{Z}$.

On note T le tétraèdre régulier et C le cube de $\mathbb{R}^3$.

Théorème. Le groupe $\mathrm{Is}(T)$ est isomorphe à $\mathfrak{S}_4$ et $\mathrm{Is}^+(T)$ à $\mathfrak{A}_4$.

Théorème. Le groupe $\mathrm{Is}^+(C)$ est isomorphe à $\mathfrak{S}_4$ et le groupe $\mathrm{Is}(C)$ est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

4 Groupes de matrices

4.1 Sur un corps quelconque

Soit K de caractéristique différente de 2.

Proposition. Soit G un sous-groupe d'exposant 2 de $\mathrm{GL}_n(K)$. Alors $|G| \leq 2^n$.

Corollaire. Les groupes $\mathrm{GL}_n(K)$ et $\mathrm{GL}_m(K)$ sont isomorphes si et seulement si $m = n$.

4.2 Sur le corps des complexes

Théorème. (de Burnside)
Tout sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ d'exposant fini est fini.

4.3 Sur l'anneau des entiers

Proposition. Soit $M \in \mathrm{GL}_n(\mathbb{Z})$ d'ordre fini d. Alors M est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité.

Théorème. Soit M une matrice de $\mathrm{GL}_n(\mathbb{Z})$ d'ordre fini. L'ensemble des ordres possibles pour M est fini.

Théorème. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Alors son cardinal est borné par 3^{n^2}.

Chapitre 5

105. Groupe des permutations d'un ensemble fini. Applications.

Sommaire

1 Groupe symétrique sur un ensemble fini 45
1.1 Définition 45
1.2 Décomposition en produit de cycles et générateurs 46
1.3 Conjugaison 46
1.4 Matrices de permutations 47
2 Signature et groupe alterné 47
2.1 Définitions 47
2.2 Structure 47
3 Loi uniforme sur les permutations 47
3.1 Nombre d'inversions 47
3.2 Nombre de cycles 48
3.3 Nombre de descentes 48

Présentation du plan

1 Groupe symétrique sur un ensemble fini

1.1 Définition

Définition. Soit X un ensemble. Le *groupe symétrique* sur X est le groupe des bijections de X muni de la composition, noté $\mathfrak{S}_X$. Dans le cas où X est fini, c'est un groupe fini de cardinal $|X|!$. On peut simplifier en le voyant comme le groupe des permutations de $\{1, \dots, n\}$, noté $\mathfrak{S}_n$ où $n = |X|$.

Théorème. (de Cayley)
Soit G un groupe de cardinal n. Alors G est isomorphe à un sous-groupe de $\mathfrak{S}_n$.

Définition. Une permutation c de $\mathfrak{S}_n$ telle qu'il existe $\{j_1, \dots, j_k\} = J \subset \{1, \dots, n\}$ tel que $c(j_i) = j_{i+1}$ pour $i \in \{1, \dots, k-1\}$, $c(j_k) = j_1$ et $c(l) = l$ si $l \notin J$ est un *cycle*, et J est son support. Une *transposition* est un cycle dont le support est de cardinal 2.

Proposition. Pour $n \geq 3$, le centre de $\mathfrak{S}_n$ est trivial.

Définition. Les points fixes d'une permutation $\sigma \in \mathfrak{S}_n$ sont les $k \in \{1, \dots, n\}$ tels que $\sigma(k) = k$. Le complémentaire de l'ensemble des points fixes s'appelle le *support*, définition qui colle avec celle donnée pour les cycles.

Remarque. Deux permutations dont les supports sont disjoints commutent.

1.2 Décomposition en produit de cycles et générateurs

Remarque. Le groupe $\mathfrak{S}_n$ agit naturellement sur $\{1, \dots, n\}$, et on peut donc partitionner $\{1, \dots, n\}$ sous l'action d'une permutations et ainsi obtenir le résultat suivant.

Théorème. Toute permutation est produit unique (à permutation du produit près) de cycles à supports disjoints. En particulier, $\mathfrak{S}_n$ est engendré par les transpositions. La taille de ces cycles (sans ordre donc) est appelé *type* de la permutation.

Application. On considère T et C le tétraèdre et le cube de l'espace affine $\mathbb{R}^3$.

— Le groupe $\mathrm{Is}(T)$ est isomorphe à $\mathfrak{S}_4$ et $\mathrm{Is}^+(T)$ à $\mathfrak{A}_4$.

— Le groupe $\mathrm{Is}^+(C)$ est isomorphe à $\mathfrak{S}_4$ et le groupe $\mathrm{Is}(C)$ est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

Remarque. Parmi les transpositions, on peut trouver des générateurs simples ; les $(i\, i+1)$ ou les $(1\, i)$ par exemple.

Proposition. L'ordre d'une permutation est le ppcm des longueurs des cycles présents dans sa décomposition en produit de cycles à supports disjoints.

1.3 Conjugaison

Proposition. Soit $\sigma \in \mathfrak{S}_n$ et $(i_1 \dots i_r)$ un cycle de $\mathfrak{S}_n$. Alors :

$$\sigma \circ (i_1 \dots i_r) \circ \sigma^{-1} = (\sigma(i_1) \dots \sigma(i_r))$$

Proposition. Deux permutations sont conjuguées si et seulement si elles ont même type.

Théorème. Tout automorphisme de $\mathfrak{S}_n$, pour $n \neq 6$ est intérieur, *i.e.* est de la forme $\varphi_\tau : \sigma \mapsto \tau\sigma\tau^{-1}$.

1.4 Matrices de permutations

Proposition. Soit K un corps. On peut injecter $\mathfrak{S}_n$ dans $\mathrm{GL}_n(K)$ par

$$\sigma \mapsto P_\sigma := (\delta_{i,\sigma(j)})_{1 \leq i,j \leq n}.$$

Application. Cela, couplé au théorème de Cayley, permet de montrer les théorème des Sylow.

2 Signature et groupe alterné

2.1 Définitions

Définition. Il existe un unique morphisme de groupes de $\mathfrak{S}_n$ vers $\mathbb{C}^*$, noté ε et appelé *signature*. Il correspond à $(-1)^N$, où N est le nombre de transpositions présentes dans la décomposition d'une permutation en un produit de transpositions.

Remarque. La signature d'un cycle de longueur k est $(-1)^{k+1}$.

Définition. On appelle *groupe alterné* le noyau de la signature, et on le note $\mathfrak{A}_n$. C'est le seul groupe d'indice 2 de $\mathfrak{S}_n$.

2.2 Structure

Proposition. Les cycles de longueur 3 sont conjugués dans $\mathfrak{A}_n$ et engendrent $\mathfrak{A}_n$.

Théorème. Le groupe alterné $\mathfrak{A}_n$ est simple pour $n \geq 5$.

Proposition. Le groupe dérivé de $\mathfrak{S}_n$ est $\mathfrak{A}_n$ pour $n \geq 3$.

Remarque. Le groupe $\mathfrak{A}_4$ n'est pas simple, son groupe dérivé est isomorphe au groupe de Klein.

3 Loi uniforme sur les permutations

3.1 Nombre d'inversions

Définition. La loi uniforme sur $\mathfrak{S}_n$ est définie par $\mathbb{P}(\{\sigma\}) := \frac{1}{n!}$.

Définition. Le *nombre d'inversions* d'une permutation σ est le nombre de couples (i, j) tels que $i < j$ et $\sigma(j) > \sigma(i)$. On note I_n qui à σ associe le nombre d'inversions de $\sigma \in \mathfrak{S}_n$.

Proposition. Si σ suit une loi uniforme sur $\mathfrak{S}_{n-1}$, et k suit une loi uniforme sur $\{1, \ldots, n\}$, alors la permutation obtenue en insérant n à la k-ième place dans σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. La loi de I_n est celle d'une somme de n variables indépendantes de loi $\mathcal{U}(\{0, k-1\})$ où $k \in \{1, \ldots, n\}$. En particulier, $\frac{6}{n^{3/2}}(I_n - \mathbb{E}(I_n))$ converge en loi vers une loi normale $\mathcal{N}(0, 1)$.

3.2 Nombre de cycles

Proposition. On va construire une permutation aléatoire selon ses cycles. On met d'abord un 1. Puis si on a i nombres utilisés, on tire une $\mathcal{B}(\frac{1}{n-i+1})$. Si elle vaut 1, on clôt le cycle. Sinon on ne touche à rien. On choisit un nombre au hasard que l'on rajoute à la suite. On obtient alors une loi uniforme sur $\mathfrak{S}_n$.

Application. On note $K_n(\sigma)$ le nombre de cycles d'une permutation $\sigma \in \mathfrak{S}_n$. Sous la loi uniforme, K_n est une somme de $\mathcal{B}(\frac{1}{k})$ indépendantes, pour $k \in \{1, \ldots, n\}$. En particulier, $\frac{K_n - \ln(n)}{\sqrt{\ln(n)}}$ converge en loi vers une loi normale $\mathcal{N}(0, 1)$.

3.3 Nombre de descentes

Définition. Soit $\sigma \in \mathfrak{S}_n$ on dit que σ présente une *descente au rang* i si $\sigma(i+1) < \sigma(i)$. On note $A_{n,k}$ le nombre de permutations de $\mathfrak{S}_n$ qui possèdent exactement k descentes.

Proposition. Soit $(U_1, \ldots, U_n) \sim \mathcal{U}(0,1)^{\otimes n}$. Pour presque tout $\omega \in \Omega$, il existe $\sigma(\omega)$ tel que $U_{\sigma(1)} < \ldots < U_{\sigma(n)}$. Alors σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. Avec les mêmes notations, la loi du nombre de descentes est la loi $\lfloor U_1 + \ldots + U_n \rfloor$. De plus, pour $\alpha \in [0,1]$, alors, lorsque n tend vers $+\infty$:

$$A(n, \alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}} \, \mathrm{e}^{-6(\alpha - 1/2)^2}.$$

Chapitre 6

106. Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de $\mathrm{GL}(E)$. Applications.

Sommaire

1 Groupe linéaire d'un espace vectoriel 50
1.1 Endomorphismes et matrices 50
1.2 Déterminant 50
1.3 Pivot de Gauss 50
2 Actions et sous-groupes 51
2.1 Translation, équivalence, similitude, congruence 51
2.2 Matrices de permutations 52
2.3 Sous-groupes finis 53
3 Groupe orthogonal 53
3.1 Matrices orthogonales 53
3.2 Réduction 54
3.3 Sous-groupes compacts du groupe linéaire 55

Présentation du plan

Dans tout le plan, E désigne un K-espace vectoriel de dimension finie n, où K est un corps.

1 Groupe linéaire d'un espace vectoriel

1.1 Endomorphismes et matrices

Définition. Un *isomorphisme* (ou *automorphisme*) de E est un endomorphisme de E bijectif. On note $\mathrm{GL}(E)$ l'ensemble des isomorphismes de E.

Proposition. Si $u \in \mathrm{GL}(E)$, alors $u^{-1} \in \mathcal{L}(E)$. En particulier, $\mathrm{GL}(E)$ est un groupe, appelé *groupe général linéaire* (abrégé en groupe linéaire) de E.

Proposition. Soit $u \in \mathcal{L}(E)$, alors $u \in \mathrm{GL}(E)$ si et seulement si u est injectif, si et seulement si u est bijectif.

Application. Soit $A \in \mathcal{M}_n(\mathbb{C})$ à diagonale strictement dominante. Alors A est inversible.

Proposition. Si on fixe une base $\mathcal{B}$ de E, alors $u \in \mathrm{GL}(E)$ si et seulement si $M := \mathrm{Mat}_{\mathcal{B}}(u)$ est inversible, et dans ce cas $\mathrm{Mat}_{\mathcal{B}}(u^{-1}) = M^{-1}$.

Corollaire. Le cardinal de $\mathrm{GL}_n(\mathbb{F}_q)$ est $(q^n-1)(q^n-q)\cdots(q^n-q^{n-1})$.

Application. Il y a $q^{d(d-1)}$ matrice nilpotentes dans $\mathcal{M}_d(\mathbb{F}_q)$.

Définition. On note $\mathrm{GL}_n(K)$ l'ensemble des matrices inversibles de $\mathcal{M}_n(K)$, et on a donc un isomorphisme non canonique entre $\mathrm{GL}(E)$ et $\mathrm{GL}_n(K)$.

1.2 Déterminant

Proposition. Soit $u \in \mathcal{L}(E)$. Alors $u \in \mathrm{GL}(E)$ si et seulement si $\det(u) \neq 0$ (si et seulement si la matrice de u dans toute base a un déterminant non nul), si et seulement si 0 n'est pas valeur propre de u.

Application. Soit A un sous-anneau (intègre) de K. Alors $M \in \mathcal{M}_n(A)$ est inversible dans $\mathcal{M}_n(A)$ si et seulement si $\det(A) \in A^\times$.

Proposition. Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Alors $\mathrm{GL}_n(\mathbb{K})$ est un ouvert dense de $\mathcal{M}_n(\mathbb{K})$.

Définition. On définit le groupe *spécial linéaire*, noté $\mathrm{SL}(E)$ comme le groupe des endomorphismes de déterminant 1. De même, on définit $\mathrm{SL}_n(K)$ comme les matrices de déterminant 1.

1.3 Pivot de Gauss

Définition. Soit H un hyperplan de E d'équation $f \in E^*$ et $u \in \mathrm{GL}(E)$ qui n'est pas l'identité. On dit que u est une *transvection* s'il existe $a \notin H$ non nul tel que pour tout $x \in E$, $u(x) = x + f(x)a$. On dit que u est une *dilatation* (de rapport α) s'il existe $a \notin H$ non nul tel que $u_{|H} = \mathrm{id}_H$ et $u_{|\mathrm{Vect}(a)} = \alpha\mathrm{id}$ avec $\alpha \neq 0$.

Remarque. Un endomorphisme est une transvection si et seulement s'il existe une base telle que sa matrice dans cette base est $I_n + \lambda E_{i,j}, \lambda \neq 0$ avec $i \neq j$, et est une dilatation si et seulement s'il existe une base telle que sa matrice dans cette base est $I_n + (\alpha-1)E_{i,i}, \alpha \neq 0$. On parle alors de matrices de *transvection élémentaire* et *dilatation élémentaire*.

Théorème. (Pivot de Gauss)
Toute matrice de $\mathrm{GL}_n(K)$ est produit de transvections élémentaires et d'une dilatation élémentaire, dont le rapport est le déterminant de la matrice. En particulier, $\mathrm{SL}_n(K)$ est engendré par les transvections élémentaires.

Corollaire. Le groupe $\mathrm{SL}_n(\mathbb{K})$ est connexe par arcs. De plus $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs, mais pas $\mathrm{GL}_n(\mathbb{R})$, qui a deux composantes connexes par arcs ; à savoir les matrices de déterminant strictement positif et strictement négatifs.

Proposition. Deux dilatations sont conjuguées si et seulement si elles ont le même rapport. Deux transvections quelconques sont conjuguées dans $\mathrm{GL}(E)$ et pour $n \geq 3$ elles le sont aussi dans $\mathrm{SL}(E)$.

Proposition. Le centre de $\mathrm{GL}(E)$ est l'ensemble des homothéties. Le centre de $\mathrm{SL}(E)$ est l'ensemble des homothéties dont le rapport est une racine n-ième de l'unité dans K.

Proposition. Le groupe dérivé de $\mathrm{GL}_n(K)$ est $\mathrm{SL}_n(K)$ sauf pour $n = |K| = 2$. Le groupe dérivé de $\mathrm{SL}_n(K)$ est lui-même sauf pour $n = 2$ et $|K| \in \{2, 3\}$.

2 Actions et sous-groupes

2.1 Translation, équivalence, similitude, congruence

Définition. Les groupes $\mathrm{GL}_n(K)$ et $\mathrm{GL}_m(K)^{\mathrm{op}}$ agissent sur $\mathcal{M}_{n,m}(K)$ par multiplication à gauche et à droite respectivement.

Remarque. Le théorème du pivot de Gauss (ou l'algorithme donnant la forme de Hermite si on considères des matrices sur un anneau principal) indique que pour chacune de ces actions, chaque orbite contient une matrice échelonnée (en lignes et en colonnes respectivement).

Proposition. Plus généralement, si $E = F_1 \oplus \ldots \oplus F_r$ est une décomposition de E avec $d_i = \dim(F_i)$, on peut faire agir $\mathrm{GL}_{d_1}(K) \times \ldots \times \mathrm{GL}_{d_r}(K)$ sur l'ensemble des telles décompositions.

Application. On peut en déduire un dénombrement des matrices diagonalisables sur un corps fini.

Définition. On peut aussi faire agir les deux groupes à la fois, et obtenir la relation d'*équivalence* sur $\mathcal{M}_{n,m}(K)$:

$$(P, Q) \cdot M := PMQ^{-1}.$$

Proposition. Il existe exactement $\min(m, n)$ orbites pour cette action, et on peut donner un représentant simple pour chaque orbite (la matrice de taille (n, m) avec un bloc I_r en haut à gauche et des zéros partout ailleurs). Le r qui apparaît alors est le rang de toute matrice dans cette orbite. On obtient donc un invariant sous l'action d'équivalence.

Définition. Le groupe $\mathrm{GL}_n(K)$ agit sur $\mathrm{M}_n(K)$ par conjugaison $P \cdot M := PMP^{-1}$ et deux matrices sont semblables si et seulement si elles sont dans la même orbite.

Remarque. La classification des matrices sous cette action est complètement résolue avec la théorie des facteurs invariants de Frobenius.

Définition. L'analogue de la similitude est la *congruence*, c'est-à-dire le fait que deux matrices représentent la même forme quadratique mais dans deux bases différentes. L'action en question est $M \cdot B \mapsto M^\top BM$ de $\mathrm{GL}_n(K)^{\mathrm{op}}$ dans $\mathcal{S}_n(K)$ (ou éventuellement $\mathcal{S}_n^+(\mathbb{R})$, $\mathcal{S}_n^{++}(\mathbb{R})$ si $K = \mathbb{R}$).

Théorème. (classification des formes quadratiques)
Soit K un corps de caractéristique différente de 2. On s'intéresse aux formes quadratiques sur un K-espace vectoriel de dimension n.

— Si K est algébriquement clos, deux formes quadratiques sont équivalentes si et seulement si elles ont même rang.

— Si $K = \mathbb{R}$, toute forme quadratique est équivalente à $x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$ avec $p+q \leq n$. Une orbite pour l'action par congruence est déterminée par un couple (p, q) appelé *signature*.

— Si K est fini, une forme quadratique non dégénérée de discriminant $\overline{a} \in K^*/(K^*)^2$ est équivalente à $x_1^2+\ldots+ax_n^2$, on a donc deux classes d'équivalences de formes non dégénérées.

2.2 Matrices de permutations

Proposition. Soit K un corps. On peut injecter $\mathfrak{S}_n$ dans $\mathrm{GL}_n(K)$ par

$$\sigma \mapsto P_\sigma := (\delta_{i,\sigma(j)})_{1\leq i,j\leq n}.$$

Remarque. Le déterminant d'une matrice de permutation est la signature de la permutation en question (pourvu que la caractéristique du corps soit différente de 2).

Corollaire. Tout groupe fini est isomorphe à un sous-groupe de $\mathrm{GL}_n(\mathbb{F}_p)$ où p est un nombre premier.

Application. On peut utiliser ceci pour montrer les théorème des Sylow.

2.3 Sous-groupes finis

Proposition. Soit K de caractéristique différente de 2, alors le cardinal d'un sous-groupe d'exposant 2 de $\mathrm{GL}_n(K)$ est borné par 2^n et la borne est atteinte.

Corollaire. Si $\mathrm{car}(K) \neq 2$ alors $\mathrm{GL}_n(K) \simeq \mathrm{GL}_m(K)$ si et seulement si $m = n$.

Théorème. (de Burnside)
Tout sous-groupe de $\mathrm{GL}_n(\mathbb{C})$ d'exposant fini est fini.

Proposition. Soit $M \in \mathrm{GL}_n(\mathbb{Z})$ d'ordre fini d. Alors M est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité.

Théorème. Soit M une matrice de $\mathrm{GL}_n(\mathbb{Z})$ d'ordre fini. L'ensemble des ordres possibles pour M est fini.

Théorème. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Alors son cardinal est borné par 3^{n^2}.

3 Groupe orthogonal

Ici, on suppose que E est un espace euclidien.

3.1 Matrices orthogonales

Définition. Soit $u \in \mathcal{L}(E)$. On dit que u est *orthogonal* si $uu^* = \mathrm{id}_E$. On note $\mathrm{O}(E)$ l'ensemble des endomorphismes orthogonaux de E.

Proposition. Soit u un endomorphisme orthogonal de E, alors :

- u est injectif donc inversible
- u^* est orthogonal
- $\det(u) = \pm 1$
- Ceci équivaut au fait que dans toute base, la matrice M de u vérifie $MM^\top = I_n$
- Ceci équivaut au fait que dans toute base, les lignes et les colonnes de la matrice de u forment une base orthonormée de E
- Ceci équivaut au fait que la matrice de u dans toute base envoie toute base orthonormée sur une base orthonormée.
- Ceci équivaut à $\|u(x)\| = \|x\|$ pour tout x.

On note alors $\mathrm{O}_n(\mathbb{R})$ l'ensemble des matrices d'endomorphismes orthogonaux, appelées *matrices orthogonales*.

Corollaire. L'ensemble $\mathrm{O}(E)$ est un sous-groupe compact de $\mathrm{GL}(E)$.

Définition. On appelle groupe *spécial orthogonal* le groupe des endomorphismes orthogonaux de déterminant positif, qui forme un sous-groupe fermé de $\mathrm{O}(E)$. De même on définit $\mathrm{SO}_n(\mathbb{R})$.

3.2 Réduction

Définition. Une symétrie orthogonale $u \in \mathrm{O}(E)$ est une isométrie involutive. On a alors la somme directe orthogonale :

$$E = \ker(u - \mathrm{id}) \oplus^\perp \ker(u + \mathrm{id}) =: E_1 \oplus^\perp E_{-1}.$$

Si $\dim(E_{-1}) = 1$, on parle de réflexion, si $\dim(E_{-1}) = 2$, de renversement.

Théorème. Le groupe orthogonal est engendré par les réflexions. Plus précisément, toute isométrie vectorielle s'écrit comme produit d'au plus n réflexions.

Le groupe spécial orthogonal est engendré par les renversements, et tout élément de $\mathrm{SO}(E)$ s'écrit comme produit d'au plus n renversements.

Corollaire. Toute isométrie affine de $\mathcal{E}$ est le produit d'au plus $(n+1)$ réflexions orthogonales (une réflexion orthogonale affine est une symétrie orthogonale par rapport à un hyperplan affine).

Proposition. Le centre de $\mathrm{O}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ et le centre de $\mathrm{SO}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ si n est paire et est trivial sinon.

Proposition. Deux symétries orthogonales s et s' sont conjuguées dans $\mathrm{O}(E)$ si et seulement si elles ont même nature, *i.e.* si et seulement si $\dim(E_1(s)) = \dim(E_1(s'))$.

Corollaire. Le groupe dérivé de $\mathrm{O}(E)$ est $\mathrm{SO}(E)$, et le groupe dérivé de $\mathrm{SO}(E)$ est $\mathrm{SO}(E)$.

Théorème. (Réduction des isométries)
Soit $u \in \mathrm{O}(E)$. Il existe une base orthonormée de E dans laquelle la matrice de u est diagonale par blocs, avec un bloc égal à I_r, un autre bloc égal à $-I_s$ et q blocs 2×2 qui sont des matrices de rotation :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

avec $n = r + s + 2q$. En particulier, $u \in \mathrm{SO}(E)$ si et seulement si s est pair.

Corollaire. L'ensemble $\mathrm{SO}(E)$ est une partie connexe par arcs de $\mathcal{L}(E)$, et $\mathrm{O}(E)$ a deux composantes connexes par arcs, qui sont $\mathrm{SO}(E)$ et $\mathrm{O}(E) \backslash \mathrm{SO}(E)$.

3.3 Sous-groupes compacts du groupe linéaire

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors il possède un produit scalaire invariant si et seulement s'il est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$

Théorème. (point fixe de Kakutani)
Soit G un sous-groupe compact de $\mathrm{GL}(E)$, et K une partie convexe et compacte de E stable par G. Alors il existe u un point fixe sous G dans K.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Chapitre 7

108. Exemples de parties génératrices d'un groupe. Applications.

Sommaire

1 Groupe engendré par une partie 57
1.1 Définition et premiers exemples 57
1.2 Groupe cyclique 58
1.3 Classification des groupes abéliens finis 58
2 Groupes symétriques et alternés 59
2.1 Générateurs du groupe symétrique 59
2.2 Conjugaison dans le groupe symétrique 59
2.3 Groupe alterné et signature 60
3 Groupe linéaire, groupe orthogonal, quaternions 60
3.1 Pivot de Gauss 60
3.2 Générateurs et réduction du groupe orthogonal 61
3.3 Quaternions et rotations en trois dimensions 62

Présentation du plan

Dans tout le plan, G désignera un groupe.

1 Groupe engendré par une partie

1.1 Définition et premiers exemples

Définition. Soit A une partie de G. On appelle *groupe engendré* par G le plus petit sous-groupe de G contenant A, et on le note $\langle A \rangle$. On dit alors que A est une partie génératrice de $\langle A \rangle$.

Proposition. Soit A une partie de G.

— $\langle A \rangle = \bigcap_{\substack{H \text{ sg } G \\ H \supset A}} H$

— $\langle A \rangle = \{a_1^{\varepsilon_1} \cdots a_N^{\varepsilon_N}, N \in \mathbb{N}, a_i \in A, \varepsilon_i = \pm 1\}$.

Exemple. Dans $\mathbb{R}$ ou $\mathbb{C}$, le sous-groupe engendré par $\{1\}$ pour $+$ est $\mathbb{Z}$.

Exemple. Dans $\mathbb{C}$, le sous-groupe engendré par $\{1, \mathrm{i}\}$ est $\mathbb{Z}[\mathrm{i}]$.

Définition. Un groupe engendré par un ensemble fini d'éléments est dit de *type fini*. Un groupe engendré par un élément est dit *monogène*.

Exemple. Un sous-groupe de $(\mathbb{R}, +)$ est soit monogène soit dense.

Définition. Soient $g, h \in G$, on définit leur *commutateur* par $[g, h] := ghg^{-1}h^{-1}$.

Définition. On définit le *groupe dérivé* de G comme le sous-groupe de G engendré par les commutateurs. C'est un sous-groupe distingué dans G, noté $\mathrm{D}(G)$.

1.2 Groupe cyclique

Définition. Un groupe *cyclique* est un groupe monogène et fini.

Proposition. Un groupe cyclique G est isomorphe à $(\mathbb{Z}/|G|\mathbb{Z})$ ou de manière équivalente à $\mathbb{U}_{|G|}$. Ainsi, un groupe monogène est ou bien $\mathbb{Z}$ ou bien un $\mathbb{Z}/n\mathbb{Z}$.

Proposition. Dans un groupe cyclique d'ordre d, il y a $\varphi(d)$ générateurs d'ordre d, et tout sous-groupe d'un groupe cyclique est cyclique.

Proposition. Un groupe cyclique vérifie la réciproque du théorème de Lagrange. Plus précisément, pour $d||G|$, G possède un unique sous-groupe d'ordre d.

Proposition. Un groupe de cardinal premier est cyclique.

Théorème. Tout sous-groupe multiplicatif fini d'un corps est cyclique.

Proposition. Le groupe multiplicatif $(\mathbb{Z}/n\mathbb{Z})^\times$ est cyclique si et seulement si $n = 2, 4, p^r, 2p^r$ avec $p \geq 3$ premier et $r \geq 1$.

1.3 Classification des groupes abéliens finis

Théorème. (forme normale de Smith) Soit $M \in \mathcal{M}_{n,p}(A)$ où A est un anneau principal. Il existe $(P, Q) \in \mathrm{GL}_n(A) \times \mathrm{GL}_p(A)$ telles que

$$PMQ = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Avec $d_1 | \ldots | d_r$. De plus, les (d_i) sont uniques à association près. Cette forme diagonale est appelée *forme (normale) de Smith* de la matrice M, et les (d_i) sont appelés *facteurs invariants*.

Corollaire. Soit G un groupe abélien de type fini, il existe des uniques $s, n_1 | \cdots | n_r$ tels que

$$G \simeq \mathbb{Z}^s \oplus (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z}).$$

Remarque. Avec le théorème chinois, on peut trouver le nombre de groupe abéliens d'ordre donné.

2 Groupes symétriques et alternés

2.1 Générateurs du groupe symétrique

Théorème. Toute permutation est produit unique (à permutation du produit près) de cycles à supports disjoints. En particulier, $\mathfrak{S}_n$ est engendré par les transpositions. La taille de ces cycles (sans ordre donc) est appelé *type* de la permutation.

Proposition. (Systèmes de générateurs du groupe symétrique)

- Les cycles
- Les transpositions
- Les transpositions $(1\, i)$
- Les transpositions $(i\, i+1)$
- La transposition $(1\, 2)$ et le cycle $(1 \cdots n)$

sont des générateurs de $\mathfrak{S}_n$.

Application. On considère T et C le tétraèdre et le cube de l'espace affine $\mathbb{R}^3$.

- Le groupe $\mathrm{Is}(T)$ est isomorphe à $\mathfrak{S}_4$ et $\mathrm{Is}^+(T)$ à $\mathfrak{A}_4$.
- Le groupe $\mathrm{Is}^+(C)$ est isomorphe à $\mathfrak{S}_4$ et le groupe $\mathrm{Is}(C)$ est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

2.2 Conjugaison dans le groupe symétrique

Proposition. Soit $\sigma \in \mathfrak{S}_n$ et $(i_1 \ \ldots \ i_r)$ un cycle de $\mathfrak{S}_n$. Alors :

$$\sigma \circ (i_1 \ \ldots \ i_r) \circ \sigma^{-1} = (\sigma(i_1) \ \ldots \ \sigma(i_r))$$

Proposition. Deux permutations sont conjuguées si et seulement si elles ont même type.

Théorème. Tout automorphisme de $\mathfrak{S}_n$, pour $n \neq 6$ est intérieur, *i.e.* est de la forme $\varphi_\tau : \sigma \mapsto \tau\sigma\tau^{-1}$.

2.3 Groupe alterné et signature

Définition. Il existe un unique morphisme de groupes de $\mathfrak{S}_n$ vers $\mathbb{C}^*$, noté ε et appelé *signature*. Il correspond à $(-1)^N$, où N est le nombre de transpositions présentes dans la décomposition d'une permutation en un produit de transpositions.

Remarque. La signature d'un cycle de longueur k est $(-1)^{k+1}$.

Définition. On appelle *groupe alterné* le noyau de la signature, et on le note $\mathfrak{A}_n$. C'est le seul groupe d'indice 2 de $\mathfrak{S}_n$.

Proposition. Les cycles de longueur 3 sont conjugués dans $\mathfrak{A}_n$ et engendrent $\mathfrak{A}_n$.

Théorème. Le groupe alterné $\mathfrak{A}_n$ est simple pour $n \geq 5$.

Proposition. Le groupe dérivé de $\mathfrak{S}_n$ est $\mathfrak{A}_n$ pour $n \geq 3$.

Remarque. Le groupe $\mathfrak{A}_4$ n'est pas simple, son groupe dérivé est isomorphe au groupe de Klein.

3 Groupe linéaire, groupe orthogonal, quaternions

3.1 Pivot de Gauss

Définition. Soit H un hyperplan de E d'équation $f \in E^*$ et $u \in \mathrm{GL}(E)$ qui n'est pas l'identité. On dit que u est une *transvection* s'il existe $a \notin H$ non nul tel que pour tout $x \in E$, $u(x) = x + f(x)a$. On dit que u est une *dilatation* (de rapport α) s'il existe $a \notin H$ non nul tel que $u_{|H} = \mathrm{id}_H$ et $u_{|\mathrm{Vect}(a)} = \alpha\mathrm{id}$ avec $\alpha \neq 0$.

Remarque. Un endomorphisme est une transvection si et seulement s'il existe une base telle que sa matrice dans cette base est $I_n + \lambda E_{i,j}, \lambda \neq 0$ avec $i \neq j$, et est une dilatation si et seulement s'il existe une base telle que sa matrice dans cette base est $I_n + (\alpha - 1)E_{i,i}, \alpha \neq 0$. On parle alors de matrices de *transvection élémentaire* et *dilatation élémentaire*.

Théorème. (Pivot de Gauss)
Toute matrice de $\mathrm{GL}_n(K)$ est produit de transvections élémentaires et d'une dilatation élémentaire, dont le rapport est le déterminant de la matrice. En particulier, $\mathrm{SL}_n(K)$ est engendré par les transvections élémentaires.

Corollaire. Le groupe $\mathrm{SL}_n(\mathbb{K})$ est connexe par arcs. De plus $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs, mais pas $\mathrm{GL}_n(\mathbb{R})$, qui a deux composantes connexes par arcs ; à savoir les matrices de déterminant strictement positif et strictement négatifs.

Proposition. Deux dilatations sont conjuguées si et seulement si elles ont le même rapport. Deux transvections quelconques sont conjuguées dans $\mathrm{GL}(E)$ et pour $n \geq 3$ elles le sont aussi dans $\mathrm{SL}(E)$.

Proposition. Le centre de $\mathrm{GL}(E)$ est l'ensemble des homothéties. Le centre de $\mathrm{SL}(E)$ est l'ensemble des homothéties dont le rapport est une racine n-ième de l'unité dans K.

Proposition. Le groupe dérivé de $\mathrm{GL}_n(K)$ est $\mathrm{SL}_n(K)$ sauf pour $n = |K| = 2$. Le groupe dérivé de $\mathrm{SL}_n(K)$ est lui-même sauf pour $n = 2$ et $|K| \in \{2, 3\}$.

3.2 Générateurs et réduction du groupe orthogonal

Définition. Une symétrie orthogonale $u \in \mathrm{O}(E)$ est une isométrie involutive. On a alors la somme directe orthogonale :

$$E = \ker(u - \mathrm{id}) \oplus^{\perp} \ker(u + \mathrm{id}) =: E_1 \oplus^{\perp} E_{-1}.$$

Si $\dim(E_{-1}) = 1$, on parle de réflexion, si $\dim(E_{-1}) = 2$, de renversement.

Théorème. Le groupe orthogonal est engendré par les réflexions. Plus précisément, toute isométrie vectorielle s'écrit comme produit d'au plus n réflexions.

Le groupe spécial orthogonal est engendré par les renversements, et tout élément de $\mathrm{SO}(E)$ s'écrit comme produit d'au plus n renversements.

Corollaire. Toute isométrie affine de $\mathcal{E}$ est le produit d'au plus $(n+1)$ réflexions orthogonales (une réflexion orthogonale affine est une symétrie orthogonale par rapport à un hyperplan affine).

Proposition. Le centre de $\mathrm{O}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ et le centre de $\mathrm{SO}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ si n est paire et est trivial sinon.

Proposition. Deux symétries orthogonales s et s' sont conjuguées dans $\mathrm{O}(E)$ si et seulement si elles ont même nature, *i.e.* si et seulement si $\dim(E_1(s)) = \dim(E_1(s'))$.

Corollaire. Le groupe dérivé de $\mathrm{O}(E)$ est $\mathrm{SO}(E)$, et le groupe dérivé de $\mathrm{SO}(E)$ est $\mathrm{SO}(E)$.

Théorème. (Réduction des isométries)
Soit $u \in \mathrm{O}(E)$. Il existe une base orthonormée de E dans laquelle la matrice de u est diagonale par blocs, avec un bloc égal à I_r, un autre bloc égal à $-I_s$ et q blocs 2×2 qui sont des matrices de rotation :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

avec $n = r + s + 2q$. En particulier, $u \in \mathrm{SO}(E)$ si et seulement si s est pair.

Corollaire. L'ensemble $\mathrm{SO}(E)$ est une partie connexe par arcs de $\mathcal{L}(E)$, et $\mathrm{O}(E)$ a deux composantes connexes par arcs, qui sont $\mathrm{SO}(E)$ et $\mathrm{O}(E) \backslash \mathrm{SO}(E)$.

3.3 Quaternions et rotations en trois dimensions

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a, b, c, d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_n$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Chapitre 8

120. Anneaux $\mathbb{Z}/n\mathbb{Z}$. Applications.

Sommaire

1 L'anneau des classes de congruence modulo n 63
1.1 Construction et premières propriétés 63
1.2 Théorème chinois 64
1.3 Éléments remarquables 65
2 Arithmétique 65
2.1 Théorèmes de Fermat, de Wilson 65
2.2 Chiffrement RSA 65
2.3 Systèmes de congruences, équations diophantiennes 66
3 Caractères et transformée de Fourier discrète 66
3.1 Dualité dans les groupes finis 66
3.2 Transformée de Fourier discrète 67
4 Utilisation de la théorie des corps 68
4.1 Carrés modulo un nombre premier 68
4.2 Symbole de Jacobi et test de Solovay-Strassen 68
4.3 Polynômes 69

Présentation du plan

1 L'anneau des classes de congruence modulo n

1.1 Construction et premières propriétés

Définition. On définit $\mathbb{Z}/n\mathbb{Z}$ pour $n \geq 2$ comme le quotient de l'anneau $\mathbb{Z}$ par son idéal $n\mathbb{Z}$. Il correspond aux classes d'équivalence pour la relation d'équivalence "être congru modulo n". C'est un anneau de caractéristique n. On note $\overline{k} = k \mod n$.

Proposition. Soit $n \geq 2$.

— L'ordre de $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$ est $\frac{n}{k \wedge n}$.

— L'élément $\overline{k}$ engendre $\mathbb{Z}/n\mathbb{Z}$ pour la loi $+$ si et seulement si $k \wedge n = 1$ si et seulement si $\overline{k}$ est inversible pour $\times$.

— L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est premier, on le note alors $\mathbb{F}_p$.

Remarque. Dans $\mathbb{Z}/n\mathbb{Z}$, tout élément non nul et non inversible est un diviseur de 0.

Définition. Soit $n \geq 1$, on définit $\varphi(n) := \{k \in \mathbb{N}^*, k \wedge n = 1\}$ l'*indicatrice d'Euler* de n.

Corollaire. (formule d'Euler-Möbius)
Dans $\mathbb{Z}/n\mathbb{Z}$, pour la loi $+$, on a donc exactement $\varphi(d)$ éléments d'ordre d, ce qui donne :

$$n = \sum_{d|n} \varphi(d).$$

Application. Le groupe des inversibles d'un corps fini est cyclique.

Proposition. Soient $n, m \geq 2$.

— $\mathrm{Hom}(\mathbb{Z}, \mathbb{Z}/n\mathbb{Z}) \simeq \mathbb{Z}/n\mathbb{Z}$

— $\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, Z) = \{0\}$

— $\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{Z}/m\mathbb{Z}) \simeq \mathbb{Z}/(n \wedge m)\mathbb{Z}$.

1.2 Théorème chinois

Théorème. (théorème chinois)
Soient m et n deux entiers positifs premiers entre eux. Alors

$$\mathbb{Z}/(nm)\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}\,,\ 1 \mod mn \mapsto (1 \mod n, 1 \mod m)$$

définit un unique isomorphisme d'anneaux.

Application. Soient m et n premiers entre eux, alors $\varphi(mn) = \varphi(n)\varphi(m)$. Cela permet d'avoir une expression de $\varphi(n)$ en fonction de sa décomposition en produit de facteur premiers.

Corollaire. Dans le cas où m et n ne sont pas premiers entre eux, les anneaux $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ et $\mathbb{Z}/(m \vee n)\mathbb{Z} \times \mathbb{Z}/(m \wedge n)\mathbb{Z}$ sont isomorphes mais on n'a pas d'isomorphisme explicite.

Remarque. Plus généralement, on a toujours la suite exacte :

$$1 \to \mathbb{Z}/(m \vee m)\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/(n \wedge m) \to 1.$$

Remarque. À partir d'une relation de Bezout, on peut simplement trouver un antécédent de l'isomorphisme chinois.

1.3 Éléments remarquables

Proposition. L'élément $\overline{k}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si tout diviseur premier de n est aussi un diviseur premier de k. En particulier, si $p_1, \ldots, p_r$ sont des nombres premiers distincts, $\mathbb{Z}/(p_1 \cdots p_r)\mathbb{Z}$ ne possède pas de nilpotent non nul.

Proposition. Soit $n \geq 2$, et $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ sa décomposition en produit de facteurs premiers. Alors $\mathbb{Z}/n\mathbb{Z}$ possède 2^r idempotents.

Proposition. Avec les mêmes notations, $\overline{a}$ est un carré dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si la classe de a modulo $p_i^{\alpha_i}$ est un carré dans $\mathbb{Z}/(p_i^{\alpha_i})\mathbb{Z}$ pour tout i.

2 Arithmétique

2.1 Théorèmes de Fermat, de Wilson

Théorème. (Fermat)
Soit p un nombre premier et $a \in \mathbb{Z}$.

— Si p ne divise pas a, alors $a^{p-1} \equiv 1 \mod p$.

— En général, $a^p \equiv a \mod p$.

Théorème. (de Wilson)
Soit $n \in \mathbb{N}$, n est premier si et seulement si $(n-1)! \equiv -1 \mod n$.

Application. Si $r = \frac{p-1}{2}$, p premier impair, alors $r!$ est une racine carrée de -1 dans $\mathbb{F}_p$.

Corollaire. Il y a une infinité de nombres premiers congrus à 1 modulo 4.

Théorème. (d'Euler)
Soit $a \in \mathbb{Z}$ et $n \in \mathbb{N}^*$. Alors si $a \wedge n = 1$, $a^{\varphi(n)} \equiv 1 \mod n$.

Application. Le chiffre des unités de 7^{222} est 9

2.2 Chiffrement RSA

Théorème. Soient p et q deux (grands) nombres premiers distincts et soit $n := pq$. Soit $c \in \mathbb{N}^*$ tel que $c \wedge \varphi(n) = 1$, c'est-à-dire $c \wedge (p-1)(q-1) = 1$. Soit d l'inverse de c modulo $\varphi(n)$, c'est-à-dire $cd \equiv 1 \mod \varphi(n)$. Alors l'application de $\mathbb{Z}/n\mathbb{Z}$ dans lui-même définie par $f_c := (t \mapsto t^c)$ est bijective, de réciproque f_d.

Remarque. C'est le principe de cryptographie à clef publique, tout le monde connaît la clef c mais déchiffrer revient à connaître d donc $\varphi(n)$ donc p et q, ce qui est très compliqué car il est difficile de trouver les facteurs premiers d'un nombre (non polynomial en le nombre de chiffres).

2.3 Systèmes de congruences, équations diophantiennes

Remarque. Le théorème chinois est un bon moyen de résoudre des systèmes de congruences, à condition que les moduli soient premiers entre eux deux à deux, puisque l'on a un moyen effectif de calculer un antécédent par l'isomorphisme chinois.

Exemple. Le système $x \equiv 1 \mod 2, x \equiv 2 \mod 3, x \equiv 3 \mod 5$ a pour solutions $23 \mod 30$.

Remarque. Dans le cas où le théorème chinois ne s'applique pas, il peut quand même y avoir des solutions, et on a des méthodes pour les trouver.

Remarque. Plus généralement, réduire modulo des nombres bien choisis permet de résoudre certaines équations. Par exemple, $x \equiv 2 \mod 6, x \equiv 3 \mod 10, x \equiv 1 \mod 15$ n'a pas de solutions mais $x \equiv 1 \mod 6, x \equiv 3 \mod 10, x \equiv 1 \mod 15$ en a.

Exemple. $x^2 - 13y^2 = 7$ n'a pas de solutions entières.

3 Caractères et transformée de Fourier discrète

3.1 Dualité dans les groupes finis

Définition. Soit G un groupe fini. On appelle *caractère* de G tout morphisme de groupes de G dans $(\mathbb{C}^*, \times)$.

Définition. Pour G un groupe fini, on note $\widehat{G}$ le *groupe dual* de G, défini comme l'ensemble des caractères de G.

Remarque. Si n est l'ordre de G, tout caractère de G prend ses valeurs dans $\mathbb{U}_n$.

Exemple. Le groupe dual de $\mathfrak{S}_n$ est $\{1, \varepsilon\}$ pour tout n.

Exemple. Un groupe cyclique est isomorphe à son dual.

Proposition. Le groupe dual d'un groupe fini G est un groupe abélien fini.

Proposition. Soit G un groupe abélien fini et H un sous-groupe de G. Tout caractère de H peut être prolongé à un caractère de G.

Théorème. (de structure des groupes abéliens finis)
Soit G un groupe abélien fini non trivial. Il existe un unique entier $r \geq 1$ et des uniques $n_r | n_{r-1} | \dots | n_1$ supérieurs à 2 tels que

$$G \simeq \mathbb{U}_{n_1} \times \cdots \times \mathbb{U}_{n_r}.$$

Remarque. Le théorème chinois permet alors de montrer que pour $d||G|$, G admet au moins un sous-groupe d'ordre d.

Application. Tout groupe abélien fini est isomorphe à son dual et canoniquement isomorphe à son bidual par $g \mapsto (\chi \mapsto \chi(g))$.

3.2 Transformée de Fourier discrète

Définition. Soit H un groupe abélien fini, on peut définir un produit scalaire sur $\mathbb{C}^H$ par la formule :

$$\langle f_1, f_2 \rangle_H = \frac{1}{|H|} \sum_{h \in H} \overline{f_1(h)} f_2(h).$$

Définition. Soit $f \in \mathbb{C}^G$. On définit $\widehat{f} \in \mathbb{C}^{\widehat{G}}$ par la formule :

$$\widehat{f}(\chi) := \sum_{g \in G} \overline{\chi(g)} f(g) = |G| \langle \chi, f \rangle_G.$$

Proposition. Les caractères forment une base orthonormée de $\mathbb{C}^G$.

Proposition. (Égalité de Bessel-Parseval)
Pour toute fonction $f \in \mathbb{C}^G$, puisque la famille $\widehat{G}$ est une base orthonormée, on a

$$\sum_{g \in G} |f(g)|^2 = |G| \sum_{g \in G} |\langle \chi, f \rangle|^2 = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} |\widehat{f}(\chi)|^2.$$

Proposition. (Inversion de Fourier)
Soit $f \in \mathbb{C}^G$. Alors pour tout $g \in G$:

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \chi(g) \widehat{f}(\chi)$$

Définition. On peut définir un produit de convolution (associatif, commutatif et distributif par rapport à +) sur $\mathbb{C}^H$ pour H un groupe abélien fini par la formule suivante :

$$(f_1 * f_2)(h) := \sum_{k \in G} f_1(k) f_2(hk^{-1}).$$

Proposition. Soient f_1 et f_2 deux fonctions de $\mathbb{C}^G$. Alors $\widehat{f_1 * f_2} = \widehat{f_1}\widehat{f_2}$.

Théorème. On considère une particule sur $\mathbb{Z}/N\mathbb{Z} = (\overline{0}, \ldots, \overline{N-1})$, initialement en $\overline{0}$. À chaque étape, la particule se déplace d'un cran vers la droite ou vers la gauche ou stationne de manière équiprobable. On peut modéliser cela par une chaîne de Markov dont la loi initiale est $\delta_{\overline{0}}$ et dont la loi de transition est $\mu = \frac{1}{3}(\delta_{-\overline{1}} + \delta_{\overline{0}} + \delta_{\overline{1}})$. La loi de X_n la position de la particule au rang n a pour loi

$$\mu_n = \mu_0 * \mu^{*n} = \mu^{*n}.$$

La chaîne de Markov en question converge vers la loi uniforme notée μ_∞ et on a une approximation de la vitesse de convergence en variation totale :

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N} \mathrm{e}^{-\frac{\pi^2 n}{N^2}}.$$

4 Utilisation de la théorie des corps

4.1 Carrés modulo un nombre premier

Définition. Soient p un nombre premier et a un entier, on définit le *symbole de Legendre* $\left(\frac{a}{p}\right)$ comme valant 0 si p divise a, 1 si a est un carré non nul modulo p et -1 sinon.

Proposition. Le symbole de Legendre est multiplicatif, pour p premier et a et b deux entiers, alors $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Théorème. (Loi de réciprocité quadratique)
Soient p et q deux nombres premiers, alors

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}.$$

Exemple. Le nombre 219 est un carré modulo 383.

4.2 Symbole de Jacobi et test de Solovay-Strassen

Définition. Pour $n = p_1 \cdots p_r$ un produit de nombres premiers pas forcément distincts, on définit le *symbole de Jacobi*, noté comme le symbole de Legendre, par

$$\left(\frac{a}{p_1 \cdots p_r}\right) = \left(\frac{a}{p_1}\right) \cdots \left(\frac{a}{p_r}\right).$$

Proposition. Si $a \wedge n \neq 1$, alors $\left(\frac{a}{n}\right) = 0$, et sinon il vaut ± 1. De plus, pour a, b, m, n des entiers,

$$\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right) \text{ et } \left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$$

Proposition. (loi de réciprocité quadratique, version Jacobi)
Pour m et n positifs impairs :

$$\left(\frac{m}{n}\right) = \left(\frac{n}{m}\right)(-1)^{\frac{(m-1)(n-1)}{4}}, \left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}, \left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}.$$

Proposition. Soit n un entier positif impair différent de 1. Alors n est premier si et seulement si

$$\forall a, a \wedge n = 1, \left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n.$$

Proposition. Si n est composé, il y a au plus $\frac{\varphi(n)}{2}$ entiers compris entre 1 et $(n-1)$ tels que

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$$

Proposition. (test de Solovay-Strassen)
Le test consiste à tester l'égalité précédente. Si n est composé, il y a moins d'une chance sur 2^k pour que n passe k fois le test.

4.3 Polynômes

Théorème. (critère d'Eisenstein)
Soit $P = a_n X^n + \ldots + a_1 X + a_0 \in \mathbb{Z}[X]$. S'il existe un nombre premier p tel que $\forall i \leq n-1$, p divise a_i, p ne divise pas a_n et p^2 ne divise pas a_0, alors P est irréductible dans $\mathbb{Z}[X]$.

Exemple. Le polynôme $3X^4 + 15X^2 + 10$ est irréductible sur $\mathbb{Z}[X]$.

Définition. Soit $n \geq 1$. Le n-ième *polynôme cyclotomique* est le polynôme

$$\Phi_n(X) = \prod_{\zeta \in \mathcal{P}_n} (X - \zeta) \in \mathbb{C}[X].$$

Proposition. Le polynôme Φ_n est unitaire, de degré $\varphi(n)$.

Exemple. Pour p premier, $\Phi_p(X) = \frac{X^p - 1}{X - 1}$.

Proposition. Soit $n \geq 1$, alors

$$X^n - 1 = \prod_{d|n} \Phi_d(X).$$

Par inversion de Möbius, on a alors

$$\Phi_n(X) = \prod_{d|n} (X^d - 1)^{\mu(n/d)}$$

Corollaire. Les polynômes cyclotomiques sont à coefficients entiers.

Théorème. (progression arithmétique de Dirichlet, version faible)
Soit $n \geq 1$. Il existe une infinité de nombres premiers congrus à 1 modulo n.

Théorème. Les polynômes cyclotomiques sont irréductibles sur $\mathbb{Q}[X]$.

Application. Le polynôme $X^4 + 1$ est irréductible sur $\mathbb{Q}[X]$ mais n'est irréductible sur aucun corps fini.

Chapitre 9

121. Nombres premiers. Applications.

Sommaire

1 Les nombres premiers et leur répartition 71
1.1 Généralités 71
1.2 Répartition des nombres premiers 72
1.3 Tests de primalité 73
2 Théorie des groupes 73
2.1 Théorème de Lagrange et conséquences 73
2.2 Groupes dont le cardinal est la puissance d'un nombre premier 74
2.3 Théorèmes de Sylow 74
3 Corps finis et réduction modulo un nombre premier 75
3.1 Cardinal d'un corps fini 75
3.2 Carrés dans les corps finis 75
3.3 Symbole de Jacobi et test de Solovay-Strassen 75

Présentation du plan

1 Les nombres premiers et leur répartition

1.1 Généralités

Définition. Un nombre entier positif est *premier* s'il est supérieur à 2 et si ses seuls diviseurs sont 1 et lui-même. On note $\mathcal{P}$ l'ensemble des nombres premiers.

Proposition. Il existe une infinité de nombres premiers.

Théorème. Tout nombre entier se décompose de manière unique (à l'ordre des facteurs près) en un produit de nombre premiers.

Définition. On peut donc définir la *valuation p-adique* de $n \in \mathbb{N}^*$, notée $v_p(n)$ comme le plus grand α tel que p^α divise n.

Proposition. Soit $n \in \mathbb{N}^*$, alors

$$\frac{\varphi(n)}{n} = \prod_{v_p(n)\neq 0} \left(1 - \frac{1}{p}\right).$$

Remarque. On peut avoir ce genre de résultats avec plus de fonctions, par exemple :

$$\forall s > 1, \zeta(s) = \prod_{p\in\mathcal{P}} \frac{1}{1-p^{-s}}.$$

Théorème. (de Fermat, d'Euler)
Soit $n \in \mathbb{N}^*$ et $a \in \mathbb{Z}$. Alors

$$a^{\varphi(n)} \equiv 1 \mod n.$$

En particulier si p est premier, alors $a^{p-1} \equiv 1 \mod p$.

1.2 Répartition des nombres premiers

Proposition. La série des inverses des nombres premiers diverge.

Théorème. (de Dirichlet, version faible)
Soit $n \in \mathbb{N}^*$. Il existe une infinité de nombres premiers congrus à 1 modulo n.

Remarque. Le *théorème de la progression arithmétique* de Dirichlet, généralise ce résultat au cas $p \equiv a \mod b$ avec a et b premiers entre eux. Il se base sur l'étude de séries L de Dirichlet en analyse complexe.

Théorème. (de Tchebychev)
Soit $n \geq 1$. Il existe un nombre premier compris entre n et $2n$.

Définition. Soit $n \in \mathbb{N}^*$. On définit $\pi(n)$ le nombre d'entiers positifs premiers inférieurs ou égaux à n.

Proposition. Pour $n \in \mathbb{N}^*$, $\pi(n) \geq \frac{n}{\ln(n)} \ln(2)$.

Remarque. Il existe un théorème plus fort, le *théorème des nombres premiers* de Hadamard et La Vallée-Poussin, qui stipule que $\pi(n)$ est équivalent à $\frac{n}{\ln(n)}$. Il se prouve notamment avec des arguments d'analyse complexe.

Définition. On définit la *fonction de Möbius* $\mu : \mathbb{N}^* \mapsto \mathbb{C}$ telle que $\mu(1) = 1$, $\mu(n) = 0$ si un nombre premier au carré divise n, et $(-1)^r$ si n est un produit de r nombres premiers distincts.

Théorème. (inversion de Möbius)
Soient $f, g : \mathbb{N}^* \to \mathbb{C}$ telles que pour $n \in \mathbb{N}^*$, $f(n) = \sum_{d|n} g(d)$. Alors pour tout $n \in \mathbb{N}^*$,

$$g(n) = \sum_{d|n} g(d)\mu(n/d).$$

Remarque. On peut utiliser ce résultat pour dénombrer les polynômes irréductibles sur un corps fini par exemple.

Application. Soit $(X_n)_n, (Y_n)_n \sim \bigotimes_n \mathcal{U}([\![1,n]\!])$ deux suites indépendantes, alors

$$\mathbb{P}(X_n \wedge Y_n = 1) \xrightarrow[n\to\infty]{} \frac{6}{\pi^2}.$$

1.3 Tests de primalité

Théorème. (de Wilson)
Soit $n \in \mathbb{N}^*$, c'est un nombre premier si et seulement si $(n-1)! \equiv -1 \mod n$.

Remarque. Cela fournit un algorithme de test de primalité en $\mathrm{O}(n)$.

Proposition. L'algorithme du crible d'Eratosthène, qui est naïf, est quant à lui en $\mathrm{O}(\sqrt{n})$, ce qui est mieux que le test de Wilson.

Définition. Soit $n \in \mathbb{N}^*$, on dit que c'est un *nombre de Carmichael* si n n'est pas premier et si pour tout b premier avec n, $b^{n-1} \equiv 1 \mod n$.

Proposition. Soit $n \in \mathbb{N}^*$, s'il existe b premier avec n tel que $b^{n-1} \not\equiv 1 \mod n$. Alors plus de la moitié des éléments $a \in \{1, \dots, n-1\}$ premiers avec n vérifient $a^{n-1} \not\equiv 1 \mod n$.

Corollaire. (test de primalité de Fermat)
Soit $n \in \mathbb{N}^*$. On réalise k fois l'opération suivante : on choisit $b \in \{1, \dots, n-1\}$ uniformément, et si $b \wedge n \neq 1$ ou $b^{n-1} \not\equiv 1 \mod n$, on renvoie "$n$ non premier". Ensuite, si on n'a rien renvoyé au bout de k répétitions, on renvoie "n est premier ou de Carmichael".

Ce procédé a une probabilité d'erreur inférieure ou égale à $\frac{1}{2^k}$.

Remarque. Il existe des procédés qui permettent de se débarrasser de l'alternative Carmichael, il s'agit du test de Solovay-Strassen.

2 Théorie des groupes

2.1 Théorème de Lagrange et conséquences

Remarque. Il existe un unique groupe de cardinal p où p est un nombre premier, c'est le groupe cyclique à p éléments.

Proposition. L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est premier.

Théorème. (critère d'Eisenstein)
Soit $P = a_n X^n + \ldots + a_1 X + a_0 \in \mathbb{Z}[X]$. S'il existe un nombre premier p tel que $\forall i \leq n-1$, p divise a_i, p ne divise pas a_n et p^2 ne divise pas a_0, alors P est irréductible dans $\mathbb{Z}[X]$.

Exemple. Le polynôme $3X^4 + 15X^2 + 10$ est irréductible sur $\mathbb{Z}[X]$.

Proposition. Le groupe $(\mathbb{Z}/n\mathbb{Z})^\times$ est *cyclique* si et seulement si $n = 2, 4, p^\alpha, 2p^\alpha$ avec p premier impair et $\alpha \geq 1$.

Proposition. (lemme d'Ore)
Un sous-groupe d'indice p où p est le plus petit nombre premier qui divise $|G|$ est toujours distingué dans G.

Application. Un groupe d'ordre p^2 est isomorphe à $(\mathbb{Z}/p\mathbb{Z})^2$ ou a $\mathbb{Z}/p^2\mathbb{Z}$

2.2 Groupes dont le cardinal est la puissance d'un nombre premier

Définition. Un p-groupe, où $p \in \mathcal{P}$ est un groupe d'ordre p^α avec $\alpha \geq 1$.

Proposition. Le centre d'un p-groupe est non-trivial.

Proposition. Un p-groupe vérifie la réciproque du théorème de Lagrange.

2.3 Théorèmes de Sylow

Définition. Soit G un groupe de cardinal $p^\alpha m$ avec p qui ne divise pas m. Un *p-sous-groupe de Sylow* de G est un sous-groupe de G de cardinal p^α.

Théorème. (Théorèmes de Sylow)
Soit G un groupe fini et p un diviseur premier de $|G|$.

- G contient au moins un p-sous-groupe de Sylow. Plus précisément, pour H un p-sous-groupe de G, il existe un p-Sylow de G qui contient H.
- Les p-Sylow de G sont tous conjugués. Il forment donc une orbite pour l'action par conjugaison, et leur nombre divise $|G|$.
- Le nombre k de p-Sylow de G est congru à 1 modulo p.

Application. Un groupe d'ordre 63 n'est pas simple.

Application. Le groupe $\mathfrak{S}_6$ possède un automorphisme non intérieur.

3 Corps finis et réduction modulo un nombre premier

3.1 Cardinal d'un corps fini

Théorème. Le cardinal d'un corps fini est forcément une puissance d'un nombre premier. Pour $n \in \mathbb{N}^*$ et p premier, il existe un unique corps de cardinal p^n, que l'on note $\mathbb{F}_{p^n}$.

Théorème. Les automorphismes d'un corps fini de caractéristique p sont les itérées du morphisme de Frobenius $x \mapsto x^p$, et est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

3.2 Carrés dans les corps finis

Proposition. Les carrés non nuls de $\mathbb{F}_q$ forment un sous-groupe de $\mathbb{F}_q^*$ d'indice 2.

Application. Sur un corps fini, une forme quadratique non dégénérée de discriminant $\overline{a} \in K^*/(K^*)^2$ est équivalente à $x_1^2 + \ldots + ax_n^2$, on a donc deux classes d'équivalences de formes non dégénérées.

Théorème. (des deux carrés)
Un entier strictement positif est somme de deux carrés si et seulement si la valuation p-adique de tout diviseur premier p congru à 3 modulo 4 est paire.

Définition. Soient p un nombre premier et a un entier, on définit le *symbole de Legendre* $\left(\frac{a}{p}\right)$ comme valant 0 si p divise a, 1 si a est un carré non nul modulo p et -1 sinon.

Proposition. Le symbole de Legendre est multiplicatif, pour p premier et a et b deux entiers, alors $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Théorème. (Loi de réciprocité quadratique)
Soient p et q deux nombres premiers, alors

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}.$$

Exemple. Le nombre 219 est un carré modulo 383.

3.3 Symbole de Jacobi et test de Solovay-Strassen

Définition. Pour $n = p_1 \cdots p_r$ un produit de nombres premiers pas forcément distincts, on définit le *symbole de Jacobi*, noté comme le symbole de Legendre, par

$$\left(\frac{a}{p_1 \cdots p_r}\right) = \left(\frac{a}{p_1}\right) \cdots \left(\frac{a}{p_r}\right).$$

Proposition. Si $a \wedge n \neq 1$, alors $\left(\frac{a}{n}\right) = 0$, et sinon il vaut ± 1. De plus, pour a, b, m, n des entiers,

$$\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right) \text{ et } \left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$$

Proposition. (loi de réciprocité quadratique, version Jacobi)
Pour m et n positifs impairs :

$$\left(\frac{m}{n}\right) = \left(\frac{n}{m}\right)(-1)^{\frac{(m-1)(n-1)}{4}}, \left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}, \left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}.$$

Proposition. Soit n un entier positif impair différent de 1. Alors n est premier si et seulement si

$$\forall a, a \wedge n = 1, \left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n.$$

Proposition. Si n est composé, il y a au plus $\frac{\varphi(n)}{2}$ entiers compris entre 1 et $(n-1)$ tels que

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$$

Proposition. (test de Solovay-Strassen)
Le test consiste à tester l'égalité précédente. Si n est composé, il y a moins d'une chance sur 2^k pour que n passe k fois le test.

Chapitre 10

122. Anneaux principaux. Exemples et applications.

Sommaire

1 Anneaux principaux et euclidiens 77
1.1 Caractère principal d'un anneau 77
1.2 Le cas particulier des anneaux euclidiens 78
2 Arithmétique des anneaux principaux 78
2.1 Caractère factoriel d'un anneau principal 78
2.2 PGCD et théorème de Bezout 79
2.3 Sous-anneaux de nombres remarquables et équations diophantiennes . . 80
3 Réduction de matrices à valeurs dans un anneau principal 80
3.1 Résolution d'un système d'équation diophantiennes linéaires 80
3.2 Forme normale de Smith 81

Présentation du plan

Dans tout le plan, A désignera un anneau (unitaire et commutatif) et K un corps.

1 Anneaux principaux et euclidiens

1.1 Caractère principal d'un anneau

Définition. Un idéal I de A est dit *principal* s'il est engendré par un élément.

Définition. L'anneau A est dit *principal* s'il est intègre et si tout idéal de A est principal.

Exemple. Un corps est un anneau principal.

Remarque. Si n n'est pas premier, $\mathbb{Z}/n\mathbb{Z}$ n'est pas principal car pas intègre.

1.2 Le cas particulier des anneaux euclidiens

Définition. L'anneau A est euclidien s'il est intègre et s'il existe une *jauge* $\nu : A\backslash\{0\} \to \mathbb{N}$ telle que pour tout couple (a, b) d'éléments non nuls de A, il existe $q, r \in A$ avec $a = bq + r$ et ($r = 0$ ou $\nu(r) < \nu(b)$).

Théorème. Un anneau euclidien est principal.

Exemple. Les anneaux $\mathbb{Z}$ et $K[X]$ sont euclidiens donc principaux (on prend respectivement la valeur absolue et le degré).

Proposition. L'anneau $A[X]$ est principal si et seulement si A est un corps.

Exemple. Dans $\mathbb{Z}[X]$, l'idéal $\langle 2, X\rangle$ n'est pas principal.

Définition. Soit $K \subset L$ une extension de corps. Soit $\alpha \in L$ et φ le morphisme d'évaluation en α de $K[X]$ dans L. Si φ n'est pas injective, α est dit *algébrique* sur K (sinon il est dit *transcendant*). En particulier il existe P non nul unitaire tel que $\ker(\varphi) = \langle P\rangle$; on appelle ce polynôme le *polynôme minimal* de α, noté Π_α. Dans le cas d'un K-espace vectoriel, on peut aussi définir $\Pi_{u,x}$ le générateur de $\{P \in K[X], P(u)(x) = 0\}$.

Application. Cela permet de définir les nombres algébriques et transcendants sur K dans $\mathbb{C}$ ainsi que le polynôme minimal d'un endomorphisme de E un K-espace vectoriel de dimension finie.

Exemple. (Autres exemples d'anneaux euclidiens)

- L'anneau des séries formelles $K[\![X]\!]$ muni de la valuation.
- L'anneau $\mathbb{D}$ des nombres décimaux, avec la jauge ν suivante : si $d = n2^p5^q \in \mathbb{D}^*$, $n \in \mathbb{Z}^*, (p, q) \in \mathbb{Z}^*$, on prend $\nu(d) = |n|$.

Proposition. Soit A un anneau euclidien, alors il existe un élément $x \in A\backslash A^\times$ tel que la restriction $A^\times \cup \{0\} \to A/\langle x\rangle x$ de la projection canonique soit surjective.

Application. Soit $a = \frac{1+i\sqrt{19}}{2}$. Alors $\mathbb{Z}[a]$ est principal non euclidien.

2 Arithmétique des anneaux principaux

2.1 Caractère factoriel d'un anneau principal

Proposition. Soit A principal et p non nul et non inversible dans A. Alors p est irréductible si et seulement s'il est premier, si et seulement si $\langle p\rangle$ est maximal.

Remarque. Si A est intègre et n'est pas un corps , alors $\langle 0\rangle$ n'est pas maximal mais premier.

Remarque. On a toujours, dans un anneau intègre, p premier entraîne p irréductible.

Exemple. Cela permet de définir le corps de rupture d'un polynôme P comme $K[X]/\langle P\rangle$.

Définition. L'anneau A est dit *factoriel* s'il est intègre et si

— tout $a \neq 0$ peut s'écrire $a = up_1 \cdots p_r$ avec u inversible et (p_i) irréductibles

— cette décomposition est unique à l'ordre des facteurs près et à un inversible près.

Remarque. Dans un anneau factoriel, on a toujours existence d'un PGCD et d'un PPCM.

Proposition. Soit A intègre qui vérifie la propriété d'existence précédente, alors la propriété d'unicité est équivalente au lemme d'Euclide, au théorème de Gauss et à (p premier $\Leftrightarrow$ p irréductible).

Théorème. Si A est principal alors il est factoriel.

Théorème. Si A est factoriel alors $A[X]$ l'est aussi.

Exemple. Les anneaux $Z[X]$ et $K[X, Y]$ sont factoriels mais non principaux.

Exemple. L'anneau $\mathbb{Z}[\mathrm{i}\sqrt{5}]$ n'est pas factoriel car on n'a pas la propriété d'unicité ; $6 = 3 \times 2 = (1 - \mathrm{i}\sqrt{5})(1 + \mathrm{i}\sqrt{5})$.

2.2 PGCD et théorème de Bezout

Définition. Soit A principal et a et b non nuls. Soit d un PGCD de a et b. Il existe u, v dans A tels que $d = au + bv$.

Remarque. Dans un anneau euclidien, l'algorithme d'Euclide étendu permet d'avoir accès à u, v et à d.

Théorème. (Théorème chinois)
Soient $(a_1, \ldots, a_k)$ des éléments de A deux à deux premiers entre eux. Alors le produit tensoriel φ des projections canoniques sur les $\langle a_i \rangle$ est un isomorphisme d'anneaux. De plus, si $1 = \sum_i u_i a_i$ et si $b_i := \prod_{j \neq i} a_j$ alors

$$\varphi^{-1}(x_1, \ldots, x_k) = \sum_i x_i u_i b_i.$$

Remarque. Soit E un K-espace vectoriel de dimension finie, $u \in \mathcal{L}(E)$. On peut appliquer le théorème chinois pour montrer le lemme des noyaux. Ainsi, E est somme directe des sous-espaces caractéristiques de u si Π_u le polynôme minimal de u est scindé. On a alors le diagramme commutatif suivant, où chaque flèche est un isomorphisme.

$$\begin{array}{ccc} K[X]/\langle \Pi_u \rangle & \xrightarrow{\quad c \quad} & \prod_\lambda K[X]/\langle P_\lambda \rangle \\ \Big\downarrow{\scriptstyle \varphi_u} & & \Big\downarrow{\scriptstyle (\varphi_{u_\lambda})_\lambda} \\ K[u] & \xrightarrow[\quad \psi \quad]{} & \prod_\lambda K[u_\lambda] \end{array}$$

On peut utiliser ce diagramme pour montrer la réduction de Dunford et de Jordan.

Remarque. Avec le théorème chinois, on peut voir en quoi les deux résultats suivants sont importants :

— Pour $u \in \mathcal{L}(E)$, il existe $x \in E$ tel que $\Pi_{u,x} = \Pi_u$

— Soit G un groupe abélien d'exposant fini, alors il existe un élément dont l'ordre est l'exposant du groupe.

Ces résultats permettent respectivement de montrer la réduction de Frobenius et le théorème de structure des groupes abéliens finis. On va voir dans la suite qu'il y a un lien plus fort entre ces deux théorèmes.

2.3 Sous-anneaux de nombres remarquables et équations diophantiennes

Proposition. L'anneau des entiers de Gauss $\mathbb{Z}[\mathrm{i}]$ est euclidien.

Théorème. (des deux carrés)
Un entier strictement positif est somme de deux carrés si et seulement si la valuation p-adique de tout diviseur premier p congru à 3 modulo 4 est paire.

Exemple. $41 = 5^2 + 4^2, 61 = 6^2 + 5^2$

Proposition. L'anneau $\mathbb{Z}[\mathrm{i}\sqrt{2}]$ est euclidien donc principal et factoriel.

Application. L'équation de Mordell $y^2 = x^3 - 2$ a pour unique solutions $(3, \pm 5)$.

3 Réduction de matrices à valeurs dans un anneau principal

3.1 Résolution d'un système d'équation diophantiennes linéaires

Théorème. (forme normale de Hermite)
Soit A un anneau principal et $M \in \mathcal{M}_{n,p}(A)$. Il existe H une unique matrice échelonnée en colonnes telle qu'il existe $P \in \mathrm{GL}_p(A)$ telle que $MP = H$. On dit que H est la *forme de Hermite* de A.

Exemple. On a

$$\begin{pmatrix} 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} -2 & 1 & 5 \\ 0 & 1 & 0 \\ 1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

Les solutions entières de $2x + 3y + 5z = 0$ sont donc les $(t - 5s, t, -t - 2s), t, s \in \mathbb{Z}$.

3.2 Forme normale de Smith

Théorème. (forme normale de Smith) Soit $M \in \mathcal{M}_{n,p}(A)$ où A est un anneau principal. Il existe $(P, Q) \in \mathrm{GL}_n(A) \times \mathrm{GL}_p(A)$ telles que

$$PMQ = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Avec $d_1 | \dots | d_r$. De plus, les (d_i) sont uniques à association près. Cette forme diagonale est appelée *forme (normale) de Smith* de la matrice M, et les (d_i) sont appelés *facteurs invariants*.

Corollaire. Soit G un groupe abélien de type fini, il existe des uniques $s, n_1 | \cdots | n_r$ tels que

$$G \simeq \mathbb{Z}^s \oplus (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z}).$$

Remarque. Avec le théorème chinois, on peut trouver le nombre de groupe abéliens d'ordre donné.

Remarque. En appliquant l'algorithme de Smith à une matrice de la forme $XI_n - M$, $M \in \mathcal{M}_n(K)$, on trouve les facteurs invariants de M. En fait, avec la théorie des modules, on peut montrer le théorème des facteurs invariants de Frobenius avec le théorème de Smith.

Chapitre 11

123. Corps finis. Applications.

Sommaire

1 Corps finis : cardinal, existence, unicité 83
1.1 Caractéristique et cardinal 83
1.2 Existence et unicité 84
1.3 Groupe multiplicatif 84
2 Carrés dans les corps finis 84
2.1 Dénombrement 84
2.2 Symbole de Legendre 85
3 Polynômes et corps finis 85
3.1 Réduction 85
3.2 Polynômes irréductibles sur un corps fini 85
3.3 Polynômes cyclotomiques 86
4 Matrices sur un corps fini 86

Présentation du plan

1 Corps finis : cardinal, existence, unicité

1.1 Caractéristique et cardinal

Proposition. Soit K un corps fini inclus dans L un autre corps fini. Alors L possède une structure de K-espace vectoriel, donc le cardinal de L est une puissance de celui de K.

Définition. Soit K un corps, le *sous-corps premier* de K est le plus petit corps (au sens de l'inclusion) inclus dans K.

Proposition. Soit K un corps fini, le sous-corps premier de K est isomorphe à $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$, où p est un nombre premier, appelé *caractéristique* de K.

Corollaire. Le cardinal d'un corps est donc forcément une puissance d'un nombre premier.

Définition. Soit K un corps fini de caractéristique p, alors l'application $x \mapsto x^p$ est un automorphisme de corps, appelé *automorphisme de Frobenius*.

1.2 Existence et unicité

Théorème. Soit p un nombre premier, et $q := p^n$ avec $n \in \mathbb{N}^*$. Il existe un corps à q éléments, c'est un corps de décomposition de $X^q - X$ sur $\mathbb{F}_p$. En particulier, il est unique à isomorphisme près, on le note $\mathbb{F}_q$.

1.3 Groupe multiplicatif

Théorème. Soit K un corps fini, alors $K^\times$ est un groupe (multiplicatif) cyclique.

Remarque. On peut de même montrer que tout sous-groupe fini des inversibles d'un corps est cyclique. Ce résultat tombe en défaut dans un corps non commutatif. Par exemple, $\mathbb{H}_8$ le groupe des quaternions est inclus dans le corps des quaternions $\mathbb{H}$ mais n'est pas cyclique.

2 Carrés dans les corps finis

2.1 Dénombrement

Définition. On note $\mathbb{F}_q^2$ les carrés de $\mathbb{F}_q$ et $\mathbb{F}_q^{*2}$ les carrés non nuls de $\mathbb{F}_q$.

Proposition. Si q est une puissance de deux alors $\mathbb{F}_q^2 = \mathbb{F}_q$. Sinon, $|\mathbb{F}_q^2| = \frac{q+1}{2}$.

Proposition. Dans le cas où q n'est pas une puissance de 2, $x \in \mathbb{F}_q^{*2}$ si est seulement si $x^{\frac{q-1}{2}} = 1$.

Application. Il existe une infinité de nombres premiers congrus à 1 modulo 4.

Théorème. (des deux carrés)
Un entier strictement positif est somme de deux carrés si et seulement si la valuation p-adique de tout diviseur premier p congru à 3 modulo 4 est paire.

Proposition. Pour $a, b, c \in \mathbb{F}_q$, l'équation $ax^2 + bx^2 = c$ admet des solutions dans $\mathbb{F}_q$.

Application. Sur un corps fini, une forme quadratique non dégénérée de discriminant $\overline{a} \in K^*/(K^*)^2$ est équivalente à $x_1^2 + \ldots + ax_n^2$, on a donc deux classes d'équivalences de formes non dégénérées.

2.2 Symbole de Legendre

Définition. Soient p un nombre premier et a un entier, on définit le *symbole de Legendre* $\left(\frac{a}{p}\right)$ comme valant 0 si p divise a, 1 si a est un carré non nul modulo p et -1 sinon.

Proposition. Le symbole de Legendre est multiplicatif, pour p premier et a et b deux entiers, alors $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Théorème. (Loi de réciprocité quadratique)
Soient p et q deux nombres premiers, alors

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}.$$

Exemple. Le nombre 219 est un carré modulo 383.

3 Polynômes et corps finis

3.1 Réduction

Théorème. (critère d'Eisenstein)
Soit $P = a_n X^n + \ldots + a_1 X + a_0 \in \mathbb{Z}[X]$. S'il existe un nombre premier p tel que $\forall i \leq n-1$, p divise a_i, p ne divise pas a_n et p^2 ne divise pas a_0, alors P est irréductible dans $\mathbb{Z}[X]$.

Exemple. Le polynôme $3X^4 + 15X^2 + 10$ est irréductible sur $\mathbb{Z}[X]$.

Théorème. Soit A un anneau factoriel et K son corps des fractions. Soit I un idéal premier de A et $B = A/I$, soit L le corps des fractions de B. Soit P un polynôme de $K[X]$, dont le coefficient dominant n'est pas dans I. Si $\overline{P}$ est irréductible sur B ou L, il l'est sur K.

Exemple. Soit p un nombre premier, alors $X^p - X - 1$ est irréductible sur $\mathbb{Z}[X]$.

3.2 Polynômes irréductibles sur un corps fini

Proposition. Soit K un corps et P un polynôme sur $K[X]$. Alors P est irréductible sur K si et seulement si P n'a de racines dans aucune extension de K de degré inférieur à $n/2$.

Exemple. Le polynôme $X^4 + 1$ est irréductible sur $\mathbb{Z}$ mais sur aucun corps $\mathbb{F}_p$, p premier.

Proposition. Soit K un corps et P de degré n un irréductible de $K[X]$. Si L est une extension de K de degré m premier avec n, alors P est irréductible sur $L[X]$.

Exemple. Le polynôme $X^3 + X + 1$ est irréductible sur $\mathbb{F}_{25}$.

Théorème. Soit $I(n, q)$ le nombre d'irréductibles unitaires sur le corps à q éléments. Alors

$$I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \sim \frac{q^n}{n}.$$

3.3 Polynômes cyclotomiques

Définition. Pour $K = \mathbb{Q}$ ou $\mathbb{F}_p$ (p premier) et $n \in \mathbb{N}^*$ premier avec la caractéristique de K, on note K_n le corps de décomposition de $X^n - 1$ sur K. Une *racine primitive de l'unité* d'ordre n est un générateur multiplicatif de $\mu_n(K_n)$ l'ensemble des racines n-ièmes de l'unité de K_n. On note $\mu_n^*(K_n)$ l'ensemble des racines n-ièmes primitives de l'unité.

Définition. On définit le n-ième *polynôme cyclotomique* sur K par

$$\Phi_{n,K}(X) = \prod_{\zeta \in \mu_n^*(K_n)} (X - \zeta).$$

Proposition. On a la formule de récurrence suivante, pour $n \in \mathbb{N}^*$:

$$X^n - 1 = \prod_{d|n} \Phi_d.$$

Corollaire. Le polynôme $\Phi_{n,\mathbb{Q}}$ est à coefficients entiers pour tout n. De plus, $\Phi_{n,\mathbb{F}_p}$ est la réduction modulo p de son analogue sur $\mathbb{Q}$.

Théorème. (de Wedderburn)
Toute algèbre à division finie est commutative.

Théorème. (de la progression arithmétique de Dirichlet, version faible)
Soit $n \in \mathbb{N}^*$, il existe une infinité de nombre premiers congrus à 1 modulo n.

Théorème. Les polynômes cyclotomiques sont irréductibles sur $\mathbb{Q}$ et $\mathbb{Z}$.

4 Matrices sur un corps fini

Proposition. Le groupe $|\mathrm{GL}_n(\mathbb{F}_q)|$ a pour cardinal $(q^n-1)(q^n-q)\cdots(q^n-q^{n-1})$.

Proposition. Tout groupe fini s'injecte dans $\mathrm{GL}_n(\mathbb{F}_q)$.

Remarque. On peut utiliser le résultat précédent pour montrer les théorèmes de Sylow.

Proposition. Le nombre de matrices diagonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|}.$$

Proposition. Il y a $q^{q(q-1)}$ matrices nilpotentes dans $\mathcal{M}_d(\mathbb{F}_q)$.

Proposition. Le nombre de matrices trigonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|} \prod_{i=1}^{q} q^{n_i(n_i-1)}.$$

Chapitre 12

125. Extensions de corps. Exemples et applications.

Sommaire

1 Extensions de corps, éléments algébriques et transcendants 90
1.1 Extension de corps 90
1.2 Éléments algébriques et transcendants 90
2 Corps de rupture, corps de décomposition 91
2.1 Corps de rupture 91
2.2 Corps de décomposition 91
3 Corps finis 91
3.1 Construction 91
3.2 Polynômes irréductibles 92
4 Nombres constructibles à la règle et au compas 92
4.1 Premiers exemples et critère de constructibilité 92
4.2 Polygones constructibles 93

Présentation du plan

Dans tout le plan, K, L et M seront des corps.

1 Extensions de corps, éléments algébriques et transcendants

1.1 Extension de corps

Définition. Une *extension de corps* de K est un surcorps L de K. L'extension se note alors L/K.

Exemple. On a la chaîne d'extensions $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Proposition. Soit L/K une extension de corps, alors L possède une structure naturelle de K-espace vectoriel.

Exemple. Si K et L sont deux corps finis, alors $|L| = |K|^n$ pour un entier $n \geq 1$.

Définition. La dimension de L sur K s'appelle *degré* de l'extension, on le note $[L : K]$. S'il est fini, on parle d'*extension finie*.

Exemple. Soit $\alpha = \frac{1+\sqrt{5}}{2}$. Alors $\mathbb{Q}[\alpha]$ est une extension sur $\mathbb{Q}$ de degré 2.

Théorème. Soit $K \subset L \subset M$ une chaîne d'extensions de degré fini, alors $[M : K] = [M : L][L : K]$.

1.2 Éléments algébriques et transcendants

Définition. Soit L/K une extension de corps et $\alpha \in L$. L'ensemble $I = \{P \in K[X], P(\alpha) = 0\}$ est un idéal de $K[X]$, donc principal. S'il est nul on dit que α est *transcendant* sur K et sinon on dit qu'il est *algébrique* sur K. Le générateur unitaire de I est appelé *polynôme minimal* de α. C'est un polynôme irréductible, son degré est appelé *degré* de α (c'est aussi le degré de $K[\alpha]/K$).

Exemple. Les complexes i et $\sqrt[3]{2}$ sont algébriques sur $\mathbb{Q}$.

Remarque. Dans $\mathbb{C}$, les algébriques sur $\mathbb{Q}$ sont dénombrables donc "il y a beaucoup plus de transcendants". Cependant, il est souvent difficile de montrer qu'un nombre donné est transcendant. Le premier à y parvenir est Joseph Liouville en 1844.

Proposition. Soit α un réel algébrique de degré $d > 1$ sur $\mathbb{Q}$. Il existe $A > 0$ telle que pour tout $\frac{p}{q} \in \mathbb{Q}$,

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{A}{q^d}.$$

Corollaire. Le nombre $\sum_{n=0}^{+\infty} 10^{-n!}$ est transcendant sur $\mathbb{Q}$.

Théorème. (Lindemann, ADMIS)
Les nombres e et π sont transcendants.

Proposition. Si L/K est une extension de corps, l'ensemble des algébriques sur K forme un sous-corps de L.

2 Corps de rupture, corps de décomposition

2.1 Corps de rupture

Définition. Soit L/K une extension de corps. On dit que c'est un *corps de rupture* de $P \in K[X]$, P irréductible, sur K s'il existe α tel que $P(\alpha) = 0$ et tel que $L = K(\alpha)$.

Proposition. Soit $P \in K[X]$ irréductible sur K. Alors il existe un corps de rupture de P sur K, unique à isomorphisme près.

2.2 Corps de décomposition

Définition. Soit P un polynôme de $K[X]$. On appelle *corps de décomposition* de P sur K une extension L de K telle que

— Dans $L[X]$, P est produit de facteurs de degré 1.

— Le corps L est minimal pour cette propriété.

Proposition. Pour tout $P \in K[X]$, il existe un corps de décomposition de P sur K, unique a isomorphisme près.

3 Corps finis

3.1 Construction

Proposition. Soit K un corps fini inclus dans L un autre corps fini. Alors L possède une structure de K-espace vectoriel, donc le cardinal de L est une puissance de celui de K.

Définition. Soit K un corps, le *sous-corps premier* de K est le plus petit corps (au sens de l'inclusion) inclus dans K.

Proposition. Soit K un corps fini, le sous-corps premier de K est isomorphe à $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$, où p est un nombre premier, appelé *caractéristique* de K.

Corollaire. Le cardinal d'un corps est donc forcément une puissance d'un nombre premier.

Définition. Soit K un corps fini de caractéristique p, alors l'application $x \mapsto x^p$ est un automorphisme de corps, appelé *automorphisme de Frobenius*.

Théorème. Soit p un nombre premier, et $q := p^n$ avec $n \in \mathbb{N}^*$. Il existe un corps à q éléments, c'est un corps de décomposition de $X^q - X$ sur $\mathbb{F}_p$. En particulier, il est unique à isomorphisme près, on le note $\mathbb{F}_q$.

Théorème. Tout sous-groupe fini de $K^\times$ est cyclique.

3.2 Polynômes irréductibles

Proposition. Soit K un corps et P un polynôme sur $K[X]$. Alors P est irréductible sur K si et seulement si P n'a de racines dans aucune extension de K de degré inférieur à $n/2$.

Exemple. Le polynôme $X^4 + 1$ est irréductible sur $\mathbb{Z}$ mais sur aucun corps $\mathbb{F}_p$, p premier.

Proposition. Soit K un corps et P de degré n un irréductible de $K[X]$. Si L est une extension de K de degré m premier avec n, alors P est irréductible sur $L[X]$.

Exemple. Le polynôme $X^3 + X + 1$ est irréductible sur $\mathbb{F}_{25}$.

Théorème. Soit $I(n, q)$ le nombre d'irréductibles unitaires sur le corps à q éléments. Alors

$$I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \sim \frac{q^n}{n}.$$

4 Nombres constructibles à la règle et au compas

4.1 Premiers exemples et critère de constructibilité

Définition. On dispose d'une règle non graduée, d'un compas et d'un repère orthonormé où l'unité est placée. Un nombre est dit *constructible* si l'on peut, en un nombre fini de droites et de cercles tracés, écrire $z \in \mathbb{C}$ comme point d'intersection de telles droites et cercles.

Remarque. Un complexe $z \in \mathbb{C}$ est constructible si et seulement si ses parties réelle et imaginaire le sont.

Exemple. Tous les nombres entiers (de Gauss) sont constructibles. Le nombre $\sqrt{2}$ est constructible.

Proposition. Soient x et y deux réels constructibles non nuls. Alors $x + y$, xy, $\frac{x}{y}$ et $\sqrt{x}$ sont constructibles.

Théorème. (Wantzel)
Le réel a est constructible si et seulement s'il existe une suite finie de corps $(L_i)_{0 \leq i \leq n}$ telle que

- $L_0 = \mathbb{Q}$
- L_{i+1} est une extension de degré deux de L_i
- $a \in L_n$.

Remarque. En particulier, le degré du polynôme minimal d'un nombre constructible est une puissance de deux, ce qui garantit l'impossibilité de la duplication du cube.

4.2 Polygones constructibles

Théorème. (de Gauss-Wantzel) Un polygone régulier à n côtés est constructible si et seulement si n est le produit d'une puissance de deux et d'un produit de nombres premiers de Fermat distincts.

Exemple. On n'a pas de méthode exacte pour couper une tarte en sept.

Chapitre 13

127. Exemples de nombres remarquables. Exemples d'anneaux de nombres remarquables. Exemples et applications.

Sommaire

1 Nombres premiers 96
1.1 Définition 96
1.2 Répartition des nombres premiers 96
2 Entiers de Gauss 97
2.1 Définition et structure 97
2.2 Applications en arithmétique 98
3 Irrationalité, algébricité, transcendance, approximation diophantienne 98
3.1 Nombres irrationnels et approximation diophantienne 98
3.2 Nombres algébriques et transcendants 98
3.3 Constructions à la règle et au compas 99
4 Quaternions 100
4.1 Définition 100
4.2 Isomorphisme exceptionnel 101

Présentation du plan

1 Nombres premiers

1.1 Définition

Définition. Un nombre entier positif est *premier* s'il est supérieur à 2 et si ses seuls diviseurs sont 1 et lui-même. On note $\mathcal{P}$ l'ensemble des nombres premiers.

Proposition. Il existe une infinité de nombres premiers.

Théorème. Tout nombre entier se décompose de manière unique (à l'ordre des facteurs près) en un produit de nombre premiers.

Définition. On peut donc définir la *valuation p-adique* de $n \in \mathbb{N}^*$, notée $v_p(n)$ comme le plus grand α tel que p^α divise n.

Proposition. Soit $n \in \mathbb{N}^*$, alors

$$\frac{\varphi(n)}{n} = \prod_{v_p(n)\neq 0} \left(1 - \frac{1}{p}\right).$$

Remarque. On peut avoir ce genre de résultats avec plus de fonctions, par exemple :

$$\forall s > 1, \zeta(s) = \prod_{p\in\mathcal{P}} \frac{1}{1-p^{-s}}.$$

Théorème. (de Fermat, d'Euler)
Soit $n \in \mathbb{N}^*$ et $a \in \mathbb{Z}$. Alors

$$a^{\varphi(n)} \equiv 1 \mod n.$$

En particulier si p est premier, alors $a^{p-1} \equiv 1 \mod p$.

1.2 Répartition des nombres premiers

Proposition. La série des inverses des nombres premiers diverge.

Théorème. (de Dirichlet, version faible)
Soit $n \in \mathbb{N}^*$. Il existe une infinité de nombres premiers congrus à 1 modulo n.

Remarque. Le *théorème de la progression arithmétique* de Dirichlet, généralise ce résultat au cas $p \equiv a \mod b$ avec a et b premiers entre eux. Il se base sur l'étude de séries L de Dirichlet en analyse complexe.

Théorème. (de Tchebychev)
Soit $n \geq 1$. Il existe un nombre premier compris entre n et $2n$.

Définition. Soit $n \in \mathbb{N}^*$. On définit $\pi(n)$ le nombre d'entiers positifs premiers inférieurs ou égaux à n.

Proposition. Pour $n \in \mathbb{N}^*$, $\pi(n) \geq \frac{n}{\ln(n)} \ln(2)$.

Remarque. Il existe un théorème plus fort, le *théorème des nombres premiers* de Hadamard et La Vallée-Poussin, qui stipule que $\pi(n)$ est équivalent à $\frac{n}{\ln(n)}$. Il se prouve notamment avec des arguments d'analyse complexe.

Définition. On définit la *fonction de Möbius* $\mu : \mathbb{N}^* \mapsto \mathbb{C}$ telle que $\mu(1) = 1$, $\mu(n) = 0$ si un nombre premier au carré divise n, et $(-1)^r$ si n est un produit de r nombres premiers distincts.

Théorème. (inversion de Möbius)
Soient $f, g : \mathbb{N}^* \to \mathbb{C}$ telles que pour $n \in \mathbb{N}^*$, $f(n) = \sum_{d|n} g(d)$. Alors pour tout $n \in \mathbb{N}^*$,

$$g(n) = \sum_{d|n} g(d)\mu(n/d).$$

Remarque. On peut utiliser ce résultat pour dénombrer les polynômes irréductibles sur un corps fini par exemple.

Application. Soit $(X_n)_n, (Y_n)_n \sim \bigotimes_n \mathcal{U}([\![1, n]\!])$ deux suites indépendantes, alors

$$\mathbb{P}(X_n \wedge Y_n = 1) \xrightarrow[n\to\infty]{} \frac{6}{\pi^2}.$$

2 Entiers de Gauss

2.1 Définition et structure

Définition. On définit l'ensemble des *entiers de Gauss* $\mathbb{Z}[\mathrm{i}] = \{a + \mathrm{i}b, (a, b) \in \mathbb{Z}\}$. C'est un sous-anneau de $\mathbb{C}$.

Proposition. L'application $N : z \mapsto |z|^2$ est une application multiplicative de $\mathbb{Z}[\mathrm{i}]$ dans $\mathbb{N}^*$.

Théorème. L'application N fait de $\mathbb{Z}[\mathrm{i}]$ un anneau euclidien donc principal et factoriel.

Remarque. Pour $\alpha \in \mathbb{C}\backslash\mathbb{Q}$ une racine d'un polynôme de $\mathbb{Q}[X]$, l'anneau $\mathbb{Z}[\alpha] = \{a+\alpha b, (a, b) \in \mathbb{Z}^2\}$ n'est pas forcément euclidien. À titre d'exemples, $\mathbb{Z}[\mathrm{i}]$ et $\mathbb{Z}[\mathrm{i}\sqrt{2}]$ sont euclidiens mais $\mathbb{Z}[\frac{1+\mathrm{i}\sqrt{19}}{2}]$ n'est pas euclidien.

2.2 Applications en arithmétique

Théorème. (des deux carrés)
Un entier strictement positif est somme de deux carrés si et seulement si la valuation p-adique de tout diviseur premier p congru à 3 modulo 4 est paire.

Exemple. $41 = 5^2 + 4^2, 61 = 6^2 + 5^2$

Proposition. (Une équation de Mordell)
L'équation de Mordell $y^2 = x^3 - 1$ a pour unique solution $(0, 1)$ dans $\mathbb{Z}^2$.

Remarque. Avec l'anneau $\mathbb{Z}[\mathrm{i}\sqrt{2}]$, on peut adapter les deux théorèmes précédents ; on peut résoudre $n = a^2 + 2b^2$ pour $n \in \mathbb{N}^*$ et aussi résoudre l'équation de Mordell $y^2 = x^3 - 1$.

3 Irrationalité, algébricité, transcendance, approximation diophantienne

3.1 Nombres irrationnels et approximation diophantienne

Remarque. La dénombrabilité de $\mathbb{Q}$ montre que presque tous les nombres (pour la mesure de Lebesgue) sont irrationnels.

Proposition. Le nombre $\sqrt{2}$ est irrationnel.

Théorème. (approximation de Dirichlet)
Soit $x \in \mathbb{R}$ et soit $N \geq 2$ un entier. Il existe deux entiers relatifs p, q avec $q \in \mathbb{N}^*$ inférieur à N tels que $|x - \frac{p}{q}| \leq \frac{1}{qN} \leq \frac{1}{q^2}$.

Remarque. Le jeu de l'approximation diophantienne consiste à augmenter l'exposant 2, plus précisément à trouver la borne supérieure des exposants μ tels que le nombre de rationnels $\frac{p}{q}$ que l'on peut placer entre x et $x \pm \frac{1}{q^\mu}$ est infini. Ce nombre μ est appelé *mesure d'irrationalité* de x. Plus $\mu(x)$ est grand, mieux x est approché par des rationnels, et, on va le voir, moins il est rationnel.

Proposition. La mesure d'un rationnel est égale à 1 et celle d'un irrationnel est supérieure ou égale à 2. La mesure de $\sqrt{2}$ est exactement égale à 2.

3.2 Nombres algébriques et transcendants

Définition. Soit $\alpha \in \mathbb{C}$. L'ensemble $I = \{P \in \mathbb{Q}[X], P(\alpha) = 0\}$ est un idéal de $\mathbb{Q}[X]$, donc principal. S'il est nul on dit que α est *transcendant* sur $\mathbb{Q}$ et sinon on dit qu'il est *algébrique* sur $\mathbb{Q}$. Le générateur unitaire de I est appelé *polynôme minimal* de α. C'est un polynôme irréductible, son degré est appelé *degré* de α (c'est aussi le degré de $\mathbb{Q}[\alpha]/\mathbb{Q}$ en tant qu'extension de corps).

Exemple. Les complexes i et $\sqrt[3]{2}$ sont algébriques sur $\mathbb{Q}$.

Remarque. Dans $\mathbb{C}$, les algébriques sur $\mathbb{Q}$ sont dénombrables donc "il y a beaucoup plus de transcendants". Cependant, il est souvent difficile de montrer qu'un nombre donné est transcendant. Le premier à y parvenir est Joseph Liouville en 1844.

Définition. Un réel x est dit *de Liouville* lorsque pour tout $n \in \mathbb{N}^*$, et tout $A > 0$, il existe $q \geq 2$ et $p \neq 1$ tels que $0 < |x - \frac{p}{q}| < \frac{A}{q^n}$.

Remarque. Les nombres de Liouville sont ceux qui ont une mesure d'irrationalité infinie.

Théorème. Les nombres de Liouville sont transcendants.

Corollaire. Le nombre $\sum_{n=0}^{+\infty} 10^{-n!}$ est transcendant sur $\mathbb{Q}$.

Théorème. (Lindemann, ADMIS)
Les nombres e et π sont transcendants.

Proposition. Les nombres algébriques forment un sous-corps de $\mathbb{Q}$ (c'en est même la clôture algébrique).

Remarque. On ne sait pas pour autant si $\pi + e$ est transcendant ou pas.

3.3 Constructions à la règle et au compas

Définition. On dispose d'une règle non graduée, d'un compas et d'un repère orthonormé où l'unité est placée. Un nombre est dit *constructible* si l'on peut, en un nombre fini de droites et de cercles tracés, écrire $z \in \mathbb{C}$ comme point d'intersection de telles droites et cercles.

Remarque. Un complexe $z \in \mathbb{C}$ est constructible si et seulement si ses parties réelle et imaginaire le sont.

Exemple. Tous les nombres entiers (de Gauss) sont constructibles. Le nombre $\sqrt{2}$ est constructible.

Proposition. Soient x et y deux réels constructibles non nuls. Alors $x+y$, xy, $\frac{x}{y}$ et $\sqrt{x}$ sont constructibles.

Théorème. (Wantzel)
Le réel a est constructible si et seulement s'il existe une suite finie de corps $(L_i)_{0\leq i\leq n}$ telle que

— $L_0 = \mathbb{Q}$

— L_{i+1} est une extension de degré deux de L_i

— $a \in L_n$.

Remarque. En particulier, le degré du polynôme minimal d'un nombre constructible est une puissance de deux, ce qui garantit l'impossibilité de la duplication du cube.

Théorème. (de Gauss-Wantzel) Un polygone régulier à n côtés est constructible si et seulement si n est le produit d'une puissance de deux et d'un produit de nombres premiers de Fermat distincts.

Exemple. On n'a pas de méthode exacte pour couper une tarte en sept.

4 Quaternions

4.1 Définition

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a,b,c,d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

4.2 Isomorphisme exceptionnel

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_n$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Chapitre 14

141. Polynômes irréductibles à une indéterminée. Corps de rupture. Exemples et applications.

Sommaire

1 Polynômes irréductibles à coefficients dans un anneau 104
1.1 Généralités 104
1.2 Critères d'irréductibilité 104
2 Théorie des corps 105
2.1 Extensions, éléments algébriques 105
2.2 Cyclotomie 106
2.3 Corps de rupture et de décomposition 106
3 Les nombres algébriques sur le corps des rationnels 107
3.1 Structure et nombres transcendants 107
3.2 Constructions à la règle et au compas 107

Présentation du plan

Dans tout le plan, A désigne un anneau commutatif unitaire intègre.

1 Polynômes irréductibles à coefficients dans un anneau

1.1 Généralités

Définition. Un polynôme $P \in A[X]$ non nul est dit *irréductible* lorsqu'il est non inversible et qu'il ne peut pas s'écrire $P = QR$ avec Q et R tous les deux non inversibles.

Exemple. Pour tout $a \in A$, le polynôme $X - a$ est irréductible.

Exemple. Le polynôme $2X$ est irréductible dans $\mathbb{R}[X]$ mais pas dans $\mathbb{Z}[X]$, où il s'écrit $2 \times X$.

Proposition. Soit $P \in A[X]$ de coefficient dominant inversible, $F \in A[X]$. Il existe $Q, R \in A[X]$ tels que $F = PQ+R$ avec $\deg(R) < \deg(P)$. En particulier, $\alpha \in A$ est racine de F si et seulement si $(X - \alpha)$ divise F.

Corollaire. Si K est un corps, alors $K[X]$ est principal (c'est même équivalent).

Proposition. Soit P un irréductible de $A[X]$, alors P n'a pas de racines dans A.

Remarque. La réciproque est fausse : $(X^2 + 1)^2$ n'a pas de racine dans $\mathbb{R}$.

Proposition. Sur un corps, les polynômes de degré 2 ou 3 sans racine sont irréductibles.

Proposition. Les irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 sans racine.

Définition. Soit A un anneau factoriel (on a donc des PGCD). On définit le *contenu* de deux polynômes P et Q comme le PGCD de ses coefficients. S'il est inversible, on dit que le polynôme est *primitif*.

Proposition. (lemme de Gauss)
Le contenu est multiplicatif (modulo un inversible).

Proposition. Les irréductibles de $A[X]$ sont les constantes de A irréductibles en tant qu'éléments de A et les polynômes P non constants primitifs et irréductibles dans $K[X]$ où K est le corps des fractions de A.

Théorème. Si A est factoriel alors $A[X]$ est factoriel.

1.2 Critères d'irréductibilité

Théorème. (critère d'Eisenstein)
Soit A un anneau factoriel et K son corps des fractions. Soit $P = a_n X^n + \cdots + a_0 \in A[X]$ de degré n et p un irréductible de A. On suppose que

- p ne divise pas a_n
- Pour $i \leq n - 1$, p divise a_i
- p^2 ne divise pas a_0.

Alors P est irréductible dans $K[X]$ (donc dans $A[X]$ s'il est primitif).

Exemple. Pour p premier, le polynôme $\frac{X^p-1}{X-1}$ est irréductible sur $\mathbb{Z}$.

Théorème. Soit A un anneau factoriel et K son corps des fractions. Soit I un idéal premier de A et $B = A/I$. C'est un anneau intègre ; on note L son corps des fractions. Soit $P = a_n X^n + \cdots + a_0$, avec $a_n \notin I$. Alors si $\overline{P}$ est irréductible sur B ou L, P est irréductible sur K.

Exemple. Le polynôme $X^2 + Y^2 + 1$ est irréductible dans $\mathbb{R}[X, Y]$. Le polynôme $X^p - X - 1$ est irréductible sur $\mathbb{Z}$ si p est un nombre premier.

2 Théorie des corps

2.1 Extensions, éléments algébriques

Définition. Soit L/K une extension de corps et $\alpha \in L$. L'ensemble $I = \{P \in K[X], P(\alpha) = 0\}$ est un idéal de $K[X]$, donc principal. S'il est nul on dit que α est *transcendant* sur K et sinon on dit qu'il est *algébrique* sur K. Le générateur unitaire de I est appelé *polynôme minimal* de α. C'est un polynôme irréductible, son degré est appelé *degré* de α (c'est aussi le degré de $K[\alpha]/K$).

Exemple. Les complexes i et $\sqrt[3]{2}$ sont algébriques sur $\mathbb{Q}$. Leurs polynômes minimaux respectifs sont $X^2 + 1$ et $X^3 - 2$.

Proposition. Soit K un corps et P un polynôme sur $K[X]$. Alors P est irréductible sur K si et seulement si P n'a de racines dans aucune extension de K de degré inférieur à $n/2$.

Exemple. Le polynôme $X^4 + 1$ est irréductible sur $\mathbb{Z}$ mais sur aucun corps $\mathbb{F}_p$, p premier.

Proposition. Soit K un corps et P de degré n un irréductible de $K[X]$. Si L est une extension de K de degré m premier avec n, alors P est irréductible sur $L[X]$.

Exemple. Le polynôme $X^3 + X + 1$ est irréductible sur $\mathbb{F}_{25}$.

2.2 Cyclotomie

Définition. Pour $K = \mathbb{Q}$ ou $\mathbb{F}_p$ (p premier) et $n \in \mathbb{N}^*$ premier avec la caractéristique de K, on note K_n le corps de décomposition de $X^n - 1$ sur K. Une *racine primitive de l'unité* d'ordre n est un générateur multiplicatif de $\mu_n(K_n)$ l'ensemble des racines n-ièmes de l'unité de K_n. On note $\mu_n^*(K_n)$ l'ensemble des racines n-ièmes primitives de l'unité.

Définition. On définit le n-ième *polynôme cyclotomique* sur K par

$$\Phi_{n,K}(X) = \prod_{\zeta \in \mu_n^*(K_n)} (X - \zeta).$$

Proposition. On a la formule de récurrence suivante, pour $n \in \mathbb{N}^*$:

$$X^n - 1 = \prod_{d|n} \Phi_d.$$

Corollaire. Le polynôme $\Phi_{n,\mathbb{Q}}$ est à coefficients entiers pour tout n. De plus, $\Phi_{n,\mathbb{F}_p}$ est la réduction modulo p de son analogue sur $\mathbb{Q}$.

Théorème. (de Wedderburn)
Toute algèbre à division finie est commutative.

Théorème. (de la progression arithmétique de Dirichlet, version faible)
Soit $n \in \mathbb{N}^*$, il existe une infinité de nombre premiers congrus à 1 modulo n.

Théorème. Les polynômes cyclotomiques sont irréductibles sur $\mathbb{Q}$ et $\mathbb{Z}$.

2.3 Corps de rupture et de décomposition

Définition. Soit L/K une extension de corps. On dit que c'est un *corps de rupture* de $P \in K[X]$, P irréductible, sur K s'il existe α tel que $P(\alpha) = 0$ et tel que $L = K(\alpha)$.

Proposition. Soit $P \in K[X]$ irréductible sur K. Alors il existe un corps de rupture de P sur K, unique à isomorphisme près.

Définition. Soit P un polynôme de $K[X]$. On appelle *corps de décomposition* de P sur K une extension L de K telle que

— Dans $L[X]$, P est produit de facteurs de degré 1.

— Le corps L est minimal pour cette propriété.

Proposition. Pour tout $P \in K[X]$, il existe un corps de décomposition de P sur K, unique a isomorphisme près.

Application. Pour q une puissance de nombre premier, il existe un unique corps à q éléments.

Remarque. Pour construire théoriquement un corps fini, on utilise le corps de décomposition. En machine, si on veut un corps à q éléments pour pouvoir faire des calculs, on préfère définir en tant que corps de rupture. Pour cela, il faudrait des irréductibles de tout degré dans un corps de cardinal p premier. On utilise alors le résultat suivant.

Théorème. Soit $I(n, q)$ le nombre d'irréductibles unitaires sur le corps à q éléments. Alors

$$I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \sim \frac{q^n}{n}.$$

Remarque. Pour avoir notre corps à q éléments, on prend donc un polynôme au hasard et on regarde s'il est irréductible (par exemple avec le teste de Rabin). Grâce au résultat précédent, on a espoir de vite trouver un polynôme irréductible.

3 Les nombres algébriques sur le corps des rationnels

3.1 Structure et nombres transcendants

Remarque. Dans $\mathbb{C}$, les algébriques sur $\mathbb{Q}$ sont dénombrables donc "il y a beaucoup plus de transcendants". Cependant, il est souvent difficile de montrer qu'un nombre donné est transcendant. Le premier à y parvenir est Joseph Liouville en 1844.

Définition. Un réel x est dit *de Liouville* lorsque pour tout $n \in \mathbb{N}^*$, et tout $A > 0$, il existe $q \geq 2$ et $p \neq 1$ tels que $0 < |x - \frac{p}{q}| < \frac{A}{q^n}$.

Remarque. Les nombres de Liouville sont ceux qui ont une mesure d'irrationalité infinie.

Théorème. Les nombres de Liouville sont transcendants.

Corollaire. Le nombre $\sum_{n=0}^{+\infty} 10^{-n!}$ est transcendant sur $\mathbb{Q}$.

Théorème. (Lindemann, ADMIS)
Les nombres e et π sont transcendants.

Théorème. L'ensemble des nombres algébriques sur $\mathbb{Q}$ est la clôture algébrique de $\mathbb{Q}$.

3.2 Constructions à la règle et au compas

Définition. On dispose d'une règle non graduée, d'un compas et d'un repère orthonormé où l'unité est placée. Un nombre est dit *constructible* si l'on peut, en un nombre fini de droites et de cercles tracés, écrire $z \in \mathbb{C}$ comme point d'intersection de telles droites et cercles.

Remarque. Un complexe $z \in \mathbb{C}$ est constructible si et seulement si ses parties réelle et imaginaire le sont.

Exemple. Tous les nombres entiers (de Gauss) sont constructibles. Le nombre $\sqrt{2}$ est constructible.

Proposition. Soient x et y deux réels constructibles non nuls. Alors $x + y$, xy, $\frac{x}{y}$ et $\sqrt{x}$ sont constructibles.

Théorème. (Wantzel)
Le réel a est constructible si et seulement s'il existe une suite finie de corps $(L_i)_{0 \leq i \leq n}$ telle que

- $L_0 = \mathbb{Q}$
- L_{i+1} est une extension de degré deux de L_i
- $a \in L_n$.

Cela se généralise facilement au cas complexe.

Remarque. En particulier, le degré du polynôme minimal d'un nombre constructible est une puissance de deux, ce qui garantit l'impossibilité de la duplication du cube.

Théorème. (de Gauss-Wantzel) Un polygone régulier à n côtés est constructible si et seulement si n est le produit d'une puissance de deux et d'un produit de nombres premiers de Fermat distincts.

Exemple. On n'a pas de méthode exacte pour couper une tarte en sept.

Chapitre 15

142. PGCD et PPCM, algorithmes de calcul. Applications.

Sommaire

1 PGCD et PPCM dans un anneau abstrait 109
1.1 Anneaux à PGCD et à PPCM 109
1.2 Cas des anneaux de polynômes 110
1.3 Relations de Bezout et applications 111
2 Calcul algorithmique 112
2.1 Algorithme d'Euclide 112
2.2 Coût pour des entiers 112
3 Réduction de matrices à valeurs dans un anneau principal 112
3.1 Résolution d'un système d'équation diophantiennes linéaires 112
3.2 Forme normale de Smith 113

Présentation du plan

Dans tout le plan, A sera toujours un anneau commutatif unitaire intègre, et K un corps.

1 PGCD et PPCM dans un anneau abstrait

1.1 Anneaux à PGCD et à PPCM

Définition. La relation de divisibilité induit une relation d'ordre sur $A/(A^\times)$. Pour $(a,b) \in A^2$, on dit que

— $d \in A$ est un *plus grand diviseur commun* ou *PGCD* de a et b si $\overline{d} = \inf(\overline{a}, \overline{b}) =: a \wedge b$.

— $m \in A$ est un *plus petit multiple commun* ou *PPCM* de a et b si $\overline{m} = \inf(\overline{a}, \overline{b}) =: a \vee b$.

On peut définir par associativité le PGCD et le PPCM d'un nombre fini d'éléments.

Proposition. L'élément $d \in A$ est un PGCD de a et b si et seulement si $\langle d \rangle$ est le plus petit idéal principal contenant $\langle a, b \rangle$. De plus, m est un PPCM de a et b si et seulement si $\langle m \rangle = \langle a \rangle \cap \langle b \rangle$.

Remarque. Dans un anneau principal, on a toujours un PGCD et un PPCM.

Remarque. En général il n'y a pas existence du PGCD ni du PPCM. Si c'est le cas, on dit que A est à PGCD (ou à PPCM). Dans $\mathbb{Z}[\mathrm{i}\sqrt{5}]$, 9 et $3(2+\mathrm{i}\sqrt{5})$ n'ont ni PGCD ni PPCM donc $\mathbb{Z}[\mathrm{i}\sqrt{5}]$ n'est pas à PGCD ni à PPCM.

Remarque. En cas d'existence, le PGCD et le PPCM sont uniques à association près.

Proposition. Si a et b ont un PPCM, alors ils ont un PGCD, et le produit des deux est associé à ab.

Remarque. La réciproque est fausse ; dans $\mathbb{Z}[\mathrm{i}\sqrt{5}]$, 2 et $1+\mathrm{i}\sqrt{5}$ n'ont ni PGCD ni PPCM.

Proposition. Un anneau factoriel est à PGCD et à PPCM.

Remarque. L'anneau des fonctions entières $\mathcal{H}(\mathbb{C})$ est à PGCD et à PPCM mais n'est pas factoriel.

Définition. On dit que des éléments de A sont *premiers entre eux* si leur PGCD est inversible.

Proposition. (Lemme de Gauss)
Soit A à pgcd. Si a et b sont premiers entre eux et $a|bc$ alors $a|c$.

1.2 Cas des anneaux de polynômes

Définition. Soit A à PGCD. Le *contenu* de $P \in A[X]$ est le PGCD de ses coefficients, noté $\mathrm{c}(P)$.

Proposition. Le contenu est multiplicatif.

Corollaire. Si A est à PGCD alors $A[X]$ aussi, et pour $P, Q \in A[X]$, $P \wedge_{A[X]} Q = \mathrm{c}(P) \wedge_A \mathrm{c}(Q) \times P \wedge_{K[X]} Q$, où K est le corps des fractions de A.

Remarque. Cela permet de montrer le théorème de transfert de Gauss, qui dit que si A est factoriel alors $A[X]$ l'est aussi.

1.3 Relations de Bezout et applications

Définition. On dit qu'un anneau est *de Bezout* lorsque tout idéal engendré par un nombre fini d'éléments est principal. Il est alors à PGCD et tout couple $(a,b) \in A^2$ vérifie une relation de Bezout : $au + bv = a \wedge b$.

Exemple. L'anneau $K[X,Y]$ est factoriel mais n'est pas de Bezout puisque $\langle X, Y\rangle$ n'est pas principal. L'anneau des fonctions entières est de Bezout bien qu'il ne soit pas factoriel.

Théorème. Soit A un anneau de Bezout, $a_1, \ldots, a_n \in A$ et soit à résoudre $x \equiv x_i \mod a_i$ pour tout $1 \leq i \leq n$. Lorsque les a_i sont premiers entre eux deux à deux, les $b_i := \prod_{j \neq i} a_i$ sont premiers entre eux dans leur ensemble et donc il existe des (u_i) tels que $b_1u_1 + \ldots + b_nu_n = 1$. Les solutions de notre problème forment alors une classe d'équivalence modulo $a_1 \cdots a_n$ représentée par $x_1u_1b_1 + \ldots + x_nu_nb_n$. En fait, l'application $(x_1, \ldots, x_n) \in A/\langle a_1\rangle \times \cdots \times A/\langle a_n\rangle \mapsto x_1u_1b_1 + \ldots + x_nu_nb_n \in A/(a_1 \cdots a_n)A$ est un isomorphisme d'anneaux.

Exemple. Le système de congruence $x \equiv 1 \mod 2, x \equiv 2 \mod 2, x \equiv 3 \mod 5$ a pour solutions $23 + 30\mathbb{Z}$.

Remarque. On peut quand même se débrouiller lorsque nos éléments ne sont pas premiers entre eux deux à deux.

Proposition. Soit A un anneau de Bezout et a et b deux éléments de A. Alors a est inversible modulo b si et seulement si a et b sont premiers entre eux, et dans ce cas, si $au+bv = 1$, l'inverse de a est la classe de u modulo b.

Exemple. Le nombre 157 est inversible modulo 343 d'inverse -142.

Exemple. Dans $\mathbb{F}_8 = \mathbb{F}_2/\langle X^3 + X + 1\rangle$, $\overline{X}$ est inversible, d'inverse $\overline{X}^2 + 1$.

Proposition. (lemme des noyaux)
Soit E un K-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Soit $P = P_1 \cdots P_r \in K[X]$ où les (P_i) sont deux à deux premiers entre eux. Alors

$$\ker(P(u)) = \bigoplus_{i=1}^{r} \ker(P_i(u)),$$

et les projecteurs sont des polynômes en u.

Application. (décomposition de Dunford)
Avec les mêmes notations, si $\chi_u = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$ est scindé, il existe un unique couple (d, n) avec d diagonalisable, n nilpotent tel que $u = n+d$ et $nd = dn$. De plus, n et d sont des polynômes en u et $d = P(u)$, où P est la solution de système de congruences $P \equiv \lambda_i \mod (X - \lambda_i)_i^m$.

2 Calcul algorithmique

2.1 Algorithme d'Euclide

Proposition. Soit A euclidien et soient $a, b \in A$. Si on fait les divisions euclidiennes successives $a = bq + r$, $b = q_1 r + r_1$, etc. alors on s'arrête en un nombre fini d'étapes, et le dernier reste non nul de ce procédé est le PGCD de a et b.

Remarque. Remonter l'algorithme d'Euclide permet de donner une relation de Bezout.

2.2 Coût pour des entiers

Proposition. Soient $0 < b < a$. Si l'algorithme d'Euclide s'arrête en plus de n étapes alors $a \geq dF_{n+2}$ et $b \geq dF_{n+1}$ où $d = a \wedge b$ et $(F_n)_n$ est la suite de Fibonacci. Ainsi le nombre d'étapes nécessaires au calcul du PGCD de a et b est un $\mathrm{O}(\log(b))$.

Remarque. Si l'on suppose que la division euclidienne de a par b coûte $\mathrm{O}(\log(a)\log(b))$ opérations, alors l'algorithme d'Euclide est en $\mathrm{O}(\log(a)\log(b))$. L'algorithme d'Euclide binaire est une variante de l'algorithme d'Euclide, mieux adaptée pour les calculs en machine.

3 Réduction de matrices à valeurs dans un anneau principal

3.1 Résolution d'un système d'équation diophantiennes linéaires

Théorème. (forme normale de Hermite)
Soit A un anneau principal et $M \in \mathcal{M}_{n,p}(A)$. Il existe H une unique matrice échelonnée en colonnes telle qu'il existe $P \in \mathrm{GL}_p(A)$ telle que $MP = H$. On dit que H est la *forme de Hermite* de A.

Exemple. On a

$$\begin{pmatrix} 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} -2 & 1 & 5 \\ 0 & 1 & 0 \\ 1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

Les solutions entières de $2x + 3y + 5z = 0$ sont donc les $(t - 5s, t, -t - 2s), t, s \in \mathbb{Z}$.

3.2 Forme normale de Smith

Théorème. (forme normale de Smith) Soit $M \in \mathcal{M}_{n,p}(A)$ où A est un anneau principal. Il existe $(P, Q) \in \mathrm{GL}_n(A) \times \mathrm{GL}_p(A)$ telles que

$$PMQ = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Avec $d_1 | \dots | d_r$. De plus, les (d_i) sont uniques à association près. Cette forme diagonale est appelée *forme (normale) de Smith* de la matrice M, et les (d_i) sont appelés *facteurs invariants*.

Corollaire. Soit G un groupe abélien de type fini, il existe des uniques $s, n_1 | \cdots | n_r$ tels que

$$G \simeq \mathbb{Z}^s \oplus (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z}).$$

Remarque. Avec le théorème chinois, on peut trouver le nombre de groupe abéliens d'ordre donné.

Remarque. En appliquant l'algorithme de Smith à une matrice de la forme $XI_n - M$, $M \in \mathcal{M}_n(K)$, on trouve les facteurs invariants de M. En fait, avec la théorie des modules, on peut montrer le théorème des facteurs invariants de Frobenius avec le théorème de Smith.

Chapitre 16

144. Racines d'un polynôme. Fonctions symétriques élémentaires. Exemples et applications.

Sommaire

1 Racines d'un polynôme 115
1.1 Propriétés algébriques et analytiques 115
1.2 Lien avec les coefficients 116
2 Extensions de corps, nombres algébriques 117
2.1 Éléments algébriques, transcendants 117
2.2 Corps de rupture, corps de décomposition 117
2.3 Cas particulier des rationnels 118
3 Localisation des racines d'un polynôme à coefficients complexes . . 119
3.1 Résultats de base 119
3.2 Analyse complexe 119

Présentation du plan

Dans tout le plan, K est un corps.

1 Racines d'un polynôme

1.1 Propriétés algébriques et analytiques

Définition. Une *racine* de $P \in K[X]$ est un $\alpha \in K$ tel que $P(\alpha) = 0$.

Proposition. Soit $P \in K[X]$. Alors $\alpha \in K$ est une racine de P si et seulement si $(X - \alpha)$ divise P.

Définition. On définit la *multiplicité* d'une racine α dans P comme le plus grand entier m tel que $(X - \alpha)^m$ divise P.

Corollaire. Un polynôme de $K[X]$ de degré n a au plus n racines sur K (en comptant les multiplicités).

Remarque. Le polynôme $4X \in \mathbb{Z}/8\mathbb{Z}[X]$ montre que ceci n'est plus vrai dans un anneau.

Proposition. Un polynôme qui s'annule en plus de valeurs que son degré est nul.

Application. (Interpolation de Lagrange)
Soient $(a_0, \dots, a_n), (b_0, \dots, b_n)$ des éléments de K de cardinal strictement supérieur à $n+1$. Alors il existe un unique polynôme $P \in K[X]$ tel que pour tout $i \in \{0, \dots, n\}, P(a_i) = b_i$.

Proposition. (formule de Taylor)
On suppose K de caractéristique nulle. Soit $P \in K[X]$, alors pour $a \in K$:

$$P(X) = \sum_{k=0}^{\infty} \frac{P^{(k)}(a)}{k!}(X-a)^k.$$

Proposition. Si K est de caractéristique nulle, alors α est racine de P de multiplicité m si et seulement si $P^{(i)}(\alpha) = 0$ pour $i \leq m-1$ et $P^{(m)}(\alpha) \neq 0$.

Exemple. En caractéristique nulle, cela est mis en défaut (avec X^p dans $\mathbb{F}_p$ par exemple).

Définition. Un polynôme est *scindé* sur K s'il s'écrit comme produit de monômes, *i.e.* si toutes ses racines sont dans K.

Exemple. Le polynôme $X^2 + 1 = (X - \mathrm{i})(X + \mathrm{i})$ est scindé sur $\mathbb{C}$ mais pas sur $\mathbb{R}$.

Théorème. (de d'Alembert-Gauss)
Tout polynôme à coefficients complexes et scindé.

Proposition. Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 sans racine.

1.2 Lien avec les coefficients

Définition. Un polynôme en n variables $P \in K[T_1, \dots, T_n]$ est dit *symétrique* lorsque pour toute permutation σ de $\mathfrak{S}_n$, $P(T_1, \dots, T_n) = P(T_{\sigma(1)}, \dots, T_{\sigma(n)})$.

Exemple. Le polynôme en trois variables $(XYZ + Z^2Y^2 + X^2Y^2 + X^2Z^2)^3$ est symétrique, tandis que $X^2Y + Z$ ne l'est pas.

Définition. Pour $k \in \{0, \dots, n\}$ on définit le k-ième *polynôme symétrique élémentaire* comme suit :

$$\Sigma_k(T_1, \dots, T_n) := \sum_{1 \leq 1 \leq i_1 \leq \dots \leq i_k \leq n} T_{i_1} \cdots T_{i_k}.$$

C'est bien un polynôme symétrique.

Proposition. (relations coefficients-racines)
Soit $P = (X-\alpha_1)\cdots(X-\alpha_n) \in K[X]$. Alors si $P = \sum_{k=0}^{n} a_k X^k$, on a $a_k = (-1)^{n-k}\Sigma_{n-k}(\alpha_1 \dots \alpha_n)$.

Proposition. Soit $M \in \mathrm{GL}_n(\mathbb{Z})$ d'ordre fini d. Alors M est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité.

Théorème. Soit M une matrice de $\mathrm{GL}_n(\mathbb{Z})$ d'ordre fini. L'ensemble des ordres possibles pour M est fini.

Théorème. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Alors son cardinal est borné par 3^{n^2}.

Théorème. Soit $P \in K[T_1, \ldots, T_n]$ un polynôme symétrique. Alors il existe un unique $Q \in K[S_1, \ldots, S_n]$ tel que $P = Q(\Sigma_1, \ldots, \Sigma_n)$.

Application. Si on a des complexes (a_i) tels que $\prod_i (X - a_i)$ est à coefficients entiers, alors pour tout k, $\prod_i (X - a_i^k)$ est à coefficients entiers.

2 Extensions de corps, nombres algébriques

2.1 Éléments algébriques, transcendants

Définition. Soit L/K une extension de corps et $\alpha \in L$. L'ensemble $I = \{P \in K[X], P(\alpha) = 0\}$ est un idéal de $K[X]$, donc principal. S'il est nul on dit que α est *transcendant* sur K et sinon on dit qu'il est *algébrique* sur K. Le générateur unitaire de I est appelé *polynôme minimal* de α. C'est un polynôme irréductible, son degré est appelé *degré* de α (c'est aussi le degré de $K[\alpha]/K$).

Exemple. Les complexes i et $\sqrt[3]{2}$ sont algébriques sur $\mathbb{Q}$.

Proposition. Si L/K est une extension de corps, l'ensemble des algébriques sur K forme un sous-corps de L.

2.2 Corps de rupture, corps de décomposition

Définition. Soit L/K une extension de corps. On dit que c'est un *corps de rupture* de $P \in K[X]$, P irréductible, sur K s'il existe α tel que $P(\alpha) = 0$ et tel que $L = K(\alpha)$.

Proposition. Soit $P \in K[X]$ irréductible sur K. Alors il existe un corps de rupture de P sur K, unique à isomorphisme près.

Définition. Soit P un polynôme de $K[X]$. On appelle *corps de décomposition* de P sur K une extension L de K telle que

— Dans $L[X]$, P est produit de facteurs de degré 1.

— Le corps L est minimal pour cette propriété.

Proposition. Pour tout $P \in K[X]$, il existe un corps de décomposition de P sur K, unique a isomorphisme près.

Remarque. On définit le corps à p^n éléments comme le corps de décomposition de $X^{p^n} - X$ sur $\mathbb{F}_p$, et on obtient donc l'existence et l'unicité d'un corps fini à p^n éléments. Cependant, pour faire des calculs (notamment en machine) on préfèrerait le décrire en tant que corps de rupture. Pour ceci, on peut utiliser le résultat suivant.

Théorème. Soit $I(n,q)$ le nombre d'irréductibles unitaires sur le corps à q éléments. Alors

$$I(n,q) = \frac{1}{n}\sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \sim \frac{q^n}{n}.$$

2.3 Cas particulier des rationnels

Remarque. Dans $\mathbb{C}$, les algébriques sur $\mathbb{Q}$ sont dénombrables donc "il y a beaucoup plus de transcendants". Cependant, il est souvent difficile de montrer qu'un nombre donné est transcendant. Le premier à y parvenir est Joseph Liouville en 1844.

Définition. Un réel x est dit *de Liouville* lorsque pour tout $n \in \mathbb{N}^*$, et tout $A > 0$, il existe $q \geq 2$ et $p \neq 1$ tels que $0 < |x - \frac{p}{q}| < \frac{A}{q^n}$.

Remarque. Les nombres de Liouville sont ceux qui ont une mesure d'irrationalité infinie.

Théorème. Les nombres de Liouville sont transcendants.

Corollaire. Le nombre $\sum_{n=0}^{+\infty} 10^{-n!}$ est transcendant sur $\mathbb{Q}$.

Théorème. (Lindemann, ADMIS)
Les nombres e et π sont transcendants.

Proposition. Les nombres algébriques forment un sous-corps de $\mathbb{Q}$ (c'en est même la clôture algébrique).

Remarque. On ne sait pas pour autant si $\pi + e$ est transcendant ou pas.

3 Localisation des racines d'un polynôme à coefficients complexes

3.1 Résultats de base

Remarque. Les méthodes numériques de Newton ou de la sécante peuvent être utilisées pour calculer des racines de polynômes. Ceux-ci jouissent néanmoins de propriétés supplémentaires, qui laissent penser que l'on peut obtenir des résultats meilleurs.

Remarque. Les valeurs propres d'une matrice sont les racines de son polynôme caractéristique ; on peut donc utiliser des méthodes de localisation des valeurs propres de la matrice compagnon associée à $P \in \mathbb{C}[X]$ pour rechercher les racines de P. Par exemple, si $P = \sum_{k=0}^{n} a_k X^k$, la méthode des disques de Gershgörin donne une zone de localisation des racines en fonction des coefficients (a_k).

Théorème. (de Gauss-Lucas)
Soit $P \in \mathbb{C}[X]$ non constant. Les racines de P' sont dans l'enveloppe convexe de celles de P.

Remarque. Cela généralise le théorème de Rolle aux polynômes complexes.

3.2 Analyse complexe

Définition. Pour γ un lacet lisse par morceaux, on définit $N_\gamma(P)$ le nombre de racines de P à l'intérieur de γ (comptées avec multiplicité).

Théorème. Soit $P \in \mathbb{C}[X]$ non constant, alors $N_\gamma(P) = \int_\gamma \frac{P'}{P}$.

Théorème. (de Jensen)
Soit P un polynôme complexe avec un terme constant non nul. Pour un réel $r > 0$, on note $z_1, \dots, z_p$ les racines de P de module inférieur à r (comptées avec multiplicité). Alors :

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(|P(r\mathrm{e}^{\mathrm{i}\theta})|)\,\mathrm{d}\theta = \log(|P(0)|) + \sum_{j=1}^{p} \log\left(\frac{r}{|z_j|}\right).$$

De manière équivalente :

$$\int_0^r \frac{N_{D(0,t)}(P)}{t}\,\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi} \log\left|\frac{P(r\mathrm{e}^{\mathrm{i}\theta})}{P(0)}\right|\,\mathrm{d}\theta.$$

Définition. Soit $P = \sum_{k=0}^{d} a_k X^k = \prod_{j=1}^{d}(X - z_j) \in \mathbb{C}[X]$. On définit la *hauteur* de P : $\mathrm{H}(P) := \max_{0\le k\le d} |a_k|$ et sa *mesure de Mahler* $\mathrm{M}(P) = |a_d| \prod_{k=1}^{d} \max(|z_k|, 1)$.

Proposition. Soit $P \in \mathbb{C}[X]$. Alors

$$\mathrm{M}(P) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log(|P(\mathrm{e}^{\mathrm{i}\theta})|)\,\mathrm{d}\theta\right).$$

En particulier la mesure de Mahler est multiplicative.

Théorème. Soit $P \in \mathbb{C}[X]$ non nul de degré d. Alors

$$\frac{1}{\binom{d}{\lfloor d/2\rfloor}}\mathrm{H}(P) \le \mathrm{M}(P) \le \sqrt{d+1}\,\mathrm{H}(P).$$

Chapitre 17

148. Dimension d'un espace vectoriel (on se limitera au cas de la dimension finie). Rang. Exemples et applications.

Sommaire

1 Dimension d'un espace vectoriel 122
1.1 Famille libre, génératrice, base 122
1.2 Théorème de la base incomplète 122
1.3 Propriétés de la dimension 122
2 Rang 123
2.1 Rang d'une famille de vecteurs, d'un endomorphisme 123
2.2 Point de vue matriciel 123
3 Topologie des espaces de dimension finie 124
3.1 Généralités 124
3.2 Convexité 124
3.3 Topologie matricielle 125
4 Théorie des corps 126
4.1 Degré d'une extension de corps 126
4.2 Éléments algébriques et corps finis 126
4.3 Cas des complexes algébriques 127

Présentation du plan

Dans tout le plan, K désigne un corps et E un K-espace vectoriel.

1 Dimension d'un espace vectoriel

1.1 Famille libre, génératrice, base

Définition. Une famille $(e_i)_{i\in I}$ d'éléments de E est *génératrice* si tout vecteur de E s'écrit comme combinaison linéaire finie des $(e_i)_i$. Elle est *libre* si $\sum_{j\in J}\lambda_j x_j = 0$ entraîne $\lambda_j = 0$ pour

tout $J \subset I$ fini. C'est une *base* si elle est libre et génératrice. Dans ce cas, tout $x \in E$ s'écrit de manière unique comme combinaison linéaire finie des $(e_i)_i$.

Exemple. La famille $((\delta_{i,j})_j)_i$ est une base de K^n. La famille $(X^n)_n$ est une base de $K[X]$.

Définition. L'espace E est de *dimension finie* s'il admet une famille génératrice finie. Dans le cas contraire, il est de *dimension infinie*.

Exemple. L'espace vectoriel K^n est de dimension finie et $K[X]$ est de dimension infinie.

Dans la suite, on suppose que E est de dimension finie.

1.2 Théorème de la base incomplète

Théorème. Soit L une famille libre de E et G une famille génératrice de E. On peut compléter L en une base de E avec des vecteurs de G. En particulier, on a existence d'une base.

Corollaire. Dans un espace de dimension finie, toutes les bases ont même cardinal, appelé *dimension* de E, noté $\dim_K(E)$ ou $\dim(E)$. On peut donc identifier E et $K^{\dim(E)}$.

Remarque. Pour qu'une famille de cardinal $\dim(E)$ soit une base, il suffit qu'elle soit génératrice ou libre.

Remarque. La dimension dépend du corps de base. Par exemple, $\dim_{\mathbb{R}}(\mathbb{R}) = 1$ mais $\dim_{\mathbb{R}}(\mathbb{C})$.

1.3 Propriétés de la dimension

Proposition. Si E et E' sont deux K-espaces vectoriels de dimension finie, alors $\dim(E \times E') = \dim(E) + \dim(E')$.

Proposition. Soit F un sous-espace vectoriel de E. Il est alors de dimension finie, de dimension inférieure à celle de E. On a égalité des dimensions si et seulement si $F = E$. De plus, on a $\dim(E) = \dim(E/F) + \dim(F)$.

Application. Tout endomorphisme de E possède un polynôme annulateur non nul.

Proposition. (formule de Grassmann)
Soient F et G deux sous-espaces vectoriels de E. Alors $\dim(F+G) = \dim(F)+\dim(G)-\dim(F \cap G)$.

Corollaire. La dimension d'une somme quelconque de sous-espaces est inférieure à la somme des dimensions. La somme est directe si et seulement si on a égalité.

2 Rang

2.1 Rang d'une famille de vecteurs, d'un endomorphisme

Définition. Le *rang* d'une famille de vecteurs est la dimension de l'espace engendré par cette famille.

Proposition. Soit $\mathcal{F}$ une famille de vecteurs de E de dimension n. Alors $\mathrm{rg}(\mathcal{F}) \leq \min(n, k)$. Le cas d'égalité en n dit que la famille est libre, et celui en k qu'elle est liée.

Définition. On définit le *rang* d'une application linéaire $u \in \mathcal{L}(E, F)$ comme la dimension de son image (F quelconque).

Théorème. (du rang)
Le rang est bien défini dans $\mathbb{N}$ et $\dim(E) = \mathrm{rg}(u) + \dim(\ker(u))$ pour tout $u \in \mathcal{L}(E, F)$.

Corollaire. En dimension finie, une application linéaire est bijective si et seulement si elle est injective, si et seulement si elle est surjective.

Remarque. Ceci n'est plus vrai en dimension infinie, comme en témoignent $P \mapsto P'$ et $P \mapsto \int_0^X P$ sur $\mathbb{R}[X]$.

Application. (Interpolation de Lagrange)
Soient $(a_0, \ldots, a_n), (b_0, \ldots, b_n)$ des éléments de K de cardinal strictement supérieur à $n+1$. Alors il existe un unique polynôme $P \in K[X]$ tel que pour tout $i \in \{0, \ldots, n\}, P(a_i) = b_i$.

2.2 Point de vue matriciel

Proposition. Soient E et F de dimension finie, alors $\dim(\mathcal{L}(E, F)) = \dim(E)\dim(F)$.

Théorème. On appelle *rang* d'une matrice la dimension de l'espace engendré par ses colonnes. Le rang d'une application u est alors celui de toute matrice de u dans des bases quelconques. En particulier, dans chaque orbite de $\mathcal{M}_{n,p}(K)$ pour l'action par équivalence, il y a une et une unique matrice diagonale par blocs J_r avec un bloc supérieur gauche égal à I_r et des zéros partout ailleurs, et le rang r caractérise ces orbites ; deux matrices sont équivalentes si et seulement si elles ont même rang.

Théorème. Le rang d'une matrice est la taille maximale d'une sous matrice carrée inversible. En particulier, il ne dépend pas du corps de base.

Proposition. L'algorithme du pivot de Gauss permet de calculer le rang d'une matrice, donc la dimension d'un espace dont on connaît une base.

3 Topologie des espaces de dimension finie

On prend ici $K = \mathbb{R}$ ou $\mathbb{C}$.

3.1 Généralités

Proposition. En dimension finie, toutes les normes sont équivalentes, et les compacts sont exactement les fermés bornés.

Remarque. Le théorème de Riesz précise que si la boule unité d'un espace vectoriel est compacte, l'espace est forcément de dimension finie.

Application. En dimension finie, une application est propre si et seulement si elle est coercitive.

Corollaire. En dimension finie, toutes les applications linéaires sont continues, et on peut définir la norme subordonnée à une norme, qui définit une norme sur l'espace des applications linéaires en question.

Remarque. Ceci n'est plus vrai en dimension infinie.

Proposition. Tout sous-espace vectoriel F de dimension finie d'un espace vectoriel V quelconque est fermé.

Application. L'exponentielle de $A \in \mathcal{M}_n(\mathbb{C})$ est un polynôme en A.

3.2 Convexité

Théorème. (de Carathéodory)
Dans $\mathbb{R}^n$, l'enveloppe convexe d'une partie A est l'ensemble des barycentres à coefficients positifs de familles de $n+1$ points de A.

Application. L'enveloppe convexe d'un compact est compacte.

Proposition. (lemme de Farkas)
Soient $f_1, \dots, f_k$ et g des formes linéaires sur E. Alors

$$\{y \in E, f_i(y) \geq 0 \forall i\} \subset \{g \geq 0\}$$

si et seulement si g est une combinaison linéaire à coefficients positifs des f_i.

Théorème. (de Helly)
Soit $(X_1, \dots, X_{d+2})$ une famille de convexes de $\mathbb{R}^d$. Si pour tout choix de $d+1$ convexes parmi ceux-ci, l'intersection est non vide, l'intersection de tous les (X_i) est non vide.

3.3 Topologie matricielle

Proposition. Soit $S \in \mathcal{S}_p(\mathbb{R})$, alors

$$\max_{\substack{W \in \mathcal{M}_{k,p}(\mathbb{R}) \\ W^\top W = I_k}} \mathrm{Tr}(WSW^\top) = \sum_{i=1}^{k} \lambda_k,$$

avec égalité quand les lignes de W sont les k premiers vecteurs de la base orthonormée qui diagonalise S (ordonnés par ordre décroissant des valeurs propres).

Application. Cela sert en statistiques dans ce que l'on appelle l'*analyse en composantes principales*. La projection sur un sous-espace de dimension k qui maximise la variance de nos données est donnée par les k premiers axes propres de la matrice de covariance empirique.

Théorème. (minmax de Courant-Fischer)
Soit A une matrice hermitienne de taille n, dont les valeurs propres sont $\lambda_1 \geq \cdots \geq \lambda_n$. Si on note G_k l'ensemble des sous-espaces vectoriels de $\mathbb{C}^n$ de dimension k, alors

$$\lambda_k = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle.$$

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \mathrm{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)$.

Proposition. (Eckart-Young)
La matrice $U\Sigma_k V^\top$, diagonale dont la diagonale est constituée des k premières valeurs singulières de A, est la meilleure approximation de A de rang k pour la norme spectrale (subordonnée à la norme deux).

Corollaire. (minmax de Courant-Fischer pour les valeurs singulières)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $p \leq n$, alors si G_k est l'ensemble des sous-espaces de $\mathbb{R}^p$ de dimension k :

$$\sigma_k(A) = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Théorème. (normes de Ky Fan)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$ ($p \leq n$). On appelle *k-ième norme de Ky Fan* la quantité $N_k(A) = \sigma_1(A) + \cdots + \sigma_k(A)$. Alors pour tout $k \in \{1, \ldots, p\}$ la k-ième norme de Ky Fan est une norme.

Remarque. La quantité $N_k(A)/N_p(A)$ permet de mesurer qu'ont les lignes de A à se compresser dans un sous-espace de dimension k. Plus précisément, la norme k de Ky Fan maximise la somme des projections des images par A d'une famille orthonormée de dimension k de $\mathbb{R}^p$ sur une famille orthonormée de dimension k de $\mathbb{R}^n$.

4 Théorie des corps

4.1 Degré d'une extension de corps

Proposition. Soit L/K une extension de corps, alors L possède une structure naturelle de K-espace vectoriel.

Exemple. Si K et L sont deux corps finis, alors $|L| = |K|^n$ pour un entier $n \geq 1$.

Définition. La dimension de L sur K s'appelle *degré* de l'extension, on le note $[L : K]$. S'il est fini, on parle d'*extension finie*.

Exemple. Soit $\alpha = \frac{1+\sqrt{5}}{2}$. Alors $\mathbb{Q}[\alpha]$ est une extension sur $\mathbb{Q}$ de degré 2.

Théorème. Soit $K \subset L \subset M$ une chaîne d'extensions de degré fini, alors $[M : K] = [M : L][L : K]$.

4.2 Éléments algébriques et corps finis

Définition. Soit L/K une extension de corps et $\alpha \in L$. L'ensemble $I = \{P \in K[X], P(\alpha) = 0\}$ est un idéal de $K[X]$, donc principal. S'il est nul on dit que α est *transcendant* sur K et sinon on dit qu'il est *algébrique* sur K. Le générateur unitaire de I est appelé *polynôme minimal* de α. C'est un polynôme irréductible, son degré est appelé *degré* de α (c'est aussi le degré de $K[\alpha]/K$).

Proposition. Si L/K est une extension de corps, l'ensemble des algébriques sur K forme un sous-corps de L.

Définition. Soit L/K une extension de corps. On dit que c'est un *corps de rupture* de $P \in K[X]$, P irréductible, sur K s'il existe α tel que $P(\alpha) = 0$ et tel que $L = K(\alpha)$.

Proposition. Soit $P \in K[X]$ irréductible sur K. Alors il existe un corps de rupture de P sur K, unique à isomorphisme près.

Définition. Soit P un polynôme de $K[X]$. On appelle *corps de décomposition* de P sur K une extension L de K telle que

— Dans $L[X]$, P est produit de facteurs de degré 1.

— Le corps L est minimal pour cette propriété.

Proposition. Pour tout $P \in K[X]$, il existe un corps de décomposition de P sur K, unique a isomorphisme près.

Théorème. Soit $m \in \mathbb{N}^*$. Il existe un corps de cardinal m si et seulement si m est la puissance d'un nombre premier, et dans ce cas le corps en question est unique.

4.3 Cas des complexes algébriques

Remarque. Dans $\mathbb{C}$, les algébriques sur $\mathbb{Q}$ sont dénombrables donc "il y a beaucoup plus de transcendants". Cependant, il est souvent difficile de montrer qu'un nombre donné est transcendant. Le premier à y parvenir est Joseph Liouville en 1844.

Définition. Un réel x est dit *de Liouville* lorsque pour tout $n \in \mathbb{N}^*$, et tout $A > 0$, il existe $q \geq 2$ et $p \neq 1$ tels que $0 < |x - \frac{p}{q}| < \frac{A}{q^n}$.

Théorème. Les nombres de Liouville sont transcendants.

Corollaire. Le nombre $\sum_{n=0}^{+\infty} 10^{-n!}$ est transcendant sur $\mathbb{Q}$.

Définition. On dispose d'une règle non graduée, d'un compas et d'un repère orthonormé où l'unité est placée. Un nombre est dit *constructible* si l'on peut, en un nombre fini de droites et de cercles tracés, écrire $z \in \mathbb{C}$ comme point d'intersection de telles droites et cercles.

Remarque. Un complexe $z \in \mathbb{C}$ est constructible si et seulement si ses parties réelle et imaginaire le sont.

Exemple. Tous les nombres entiers (de Gauss) sont constructibles. Le nombre $\sqrt{2}$ est constructible.

Proposition. Soient x et y deux réels constructibles non nuls. Alors $x + y$, xy, $\frac{x}{y}$ et $\sqrt{x}$ sont constructibles.

Théorème. (Wantzel)
Le réel a est constructible si et seulement s'il existe une suite finie de corps $(L_i)_{0 \leq i \leq n}$ telle que

— $L_0 = \mathbb{Q}$

— L_{i+1} est une extension de degré deux de L_i

— $a \in L_n$.

On généralise cela sans mal pour les complexes constructibles.

Remarque. En particulier, le degré du polynôme minimal d'un nombre constructible est une puissance de deux, ce qui garantit l'impossibilité de la duplication du cube.

Théorème. (de Gauss-Wantzel) Un polygone régulier à n côtés est constructible si et seulement si n est le produit d'une puissance de deux et d'un produit de nombres premiers de Fermat distincts.

Exemple. On n'a pas de méthode exacte pour couper une tarte en sept.

Chapitre 18

149. Déterminant. Exemples et applications.

Sommaire

1 Formes multilinéaires et déterminant 129
1.1 Formes multilinéaires et déterminant d'une famille de vecteurs 129
1.2 Déterminant d'un endomorphisme et d'une matrice 130
2 Calcul de déterminant 131
2.1 Mineurs et cofacteurs 131
2.2 Systèmes linéaires 131
2.3 Déterminants de Gram 132
3 Propriétés géométriques et analytiques 133
3.1 Volume 133
3.2 Continuité et différentielle 133
3.3 Convexité logarithmique 134

Présentation du plan

Dans tout le plan, K est un corps de caractéristique différente de deux, et E un K-espace vectoriel de dimension finie.

1 Formes multilinéaires et déterminant

1.1 Formes multilinéaires et déterminant d'une famille de vecteurs

Définition. On note $\mathcal{L}_p(E,K)$ l'ensemble des formes p-linéaires. C'est un espace vectoriel de dimension $\dim_K(E)^p$.

Définition. Soit f une forme p-linéaire. Elle est dite

— *alternée* si $f(x_1,\dots,x_p)=0$ dès que deux des coordonnées (x_i) sont égales.

— *antisymétrique* si l'échange de deux coordonnées donne deux résultats opposés, ce qui équivaut à $f((x_{\sigma(i)})_i)=\varepsilon(\sigma)f((x_i)_i)$ pour $\sigma\in\mathfrak{S}_p$.

Proposition. Une forme p-linéaire est antisymétrique si et seulement si elle est alternée.

Proposition. Une forme p-linéaire alternée s'annule sur un système lié. Ainsi, si $f \in \mathcal{L}_p(E, K)$, la forme $g = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) f(x_{\sigma(1)}, \ldots, x_{\sigma(p)})$ est alternée (et toujours p-linéaire).

Théorème. L'ensemble des formes n-linéaires alternées sur E est de dimension 1 sur K. Il existe une et une unique forme n-linéaire alternée prenant la valeur 1 sur une base $\mathcal{B}$ de E donnée. On la note $\det_{\mathcal{B}}$ et on l'appelle *déterminant* dans la base $\mathcal{B}$.

Remarque. Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, $\det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B}) = 1$.

Proposition. Une famille est liée si et seulement s'il existe une base telle que le déterminant de la famille dans cette base est nul, si et seulement si c'est le cas dans toute base.

1.2 Déterminant d'un endomorphisme et d'une matrice

Proposition. Soit f un endomorphisme de E et $\mathcal{B} = (e_i)$ une base de E. La quantité $\det_{\mathcal{B}}(f(e_1), \ldots, f(e_n))$ ne dépend pas de $\mathcal{B}$, on l'appelle *déterminant* de f, noté $\det(f)$. En particulier, $\det(\mathrm{id}) = 1$ et pour $f, g \in \mathcal{L}(E), \det(f \circ g) = \det(f) \det(g)$. Ainsi f est bijectif si et seulement si son déterminant est non nul et alors $\det(f^{-1}) = \det(f)^{-1}$.

Définition. On définit de même la déterminant de $A = (a_{i,j}) \in \mathcal{M}_n(K)$ comme le déterminant de ses colonnes dans la base canonique et on a alors :

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Proposition. Soit $A \in \mathcal{M}_n(K)$.

- $\det(A) = \det(A^\top)$
- Si A est triangulaire, $\det(A)$ est le produit de ses éléments diagonaux
- On ne change pas $\det(A)$ en ajoutant à une ligne un combinaison linéaire des autres (idem pour les colonnes).
- Deux matrices semblables ont même déterminant
- Si A est triangulaire par blocs, $\det(A)$ est le produit des déterminants diagonaux.

2 Calcul de déterminant

2.1 Mineurs et cofacteurs

Définition. Un *mineur* de $A \in \mathcal{M}_n(K)$ est un déterminant extrait de A.

Définition. Soit $A \in \mathcal{M}_n(K)$, on définit le *cofacteur* d'indice (i, j) le déterminant de A sans sa ligne i et sa colonne j multiplié par $(-1)^{i+j}$. La matrice des cofacteurs de A est appelée *comatrice*, notée $\widetilde{A}$.

Proposition. Pour $A \in \mathcal{M}_n(K)$ et $i, j \in \{1, \ldots, n\}$

$$\det(A) = \sum_{k=1}^{n} a_{i,k} A_{i,k} = \sum_{k=1}^{n} a_{k,j} A_{k,j}.$$

Exemple. Soit $(\lambda_1, \ldots, \lambda_n) \in K^n$, alors $\det((\lambda_i^{j-1})_{i,j}) = \prod_{i \neq j}(\lambda_i - \lambda_j)$.

Exemple. Pour des complexes $(a_j)_j, (b_j)_j$ tels que $a_i + b_j \neq 0$ pour tout (i,j),

$$\Delta_n := \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & \vdots & \vdots \\ \frac{1}{a_1+b_n} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix} = \frac{\prod_{i<j}(a_j - a_i)\prod_{i<j}(b_j - b_i)}{\prod_{1 \leq i,j \leq n}(a_i + b_j)}.$$

Proposition. (formule de la comatrice)
Soit $A \in \mathcal{M}_n(K)$, alors $A(\widetilde{A})^\top = (\widetilde{A})^\top A = \det(A) I_n$.

Exemple. Cela permet d'obtenir une formule pour le déterminant d'une matrice de taille 2×2.

2.2 Systèmes linéaires

Soit à résoudre le système (S) suivant :

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,q}x_q = b_1 \\ \quad \vdots \\ a_{p,1}x_1 + \ldots + a_{p,q}x_q = b_p. \end{cases}$$

Théorème. (Système de Cramer)
On suppose $p = q = n$, on a alors un système de la forme $AX = B$ avec $A \in \mathcal{M}_n(K)$. La solution s'écrit

$$\forall 1 \leq i \leq n, x_i = \frac{\det(A_1, \ldots, A_{i-1}, B, A_{i+1}, \ldots, A_n)}{\det(A)}.$$

Théorème. (de Rouché-Fontené)
On revient au cas général, $A \in \mathcal{M}_{p,q}(K)$. Soit r le rang de A et A' une sous-matrice carrée de A de taille r inversible. Son déterminant est appelé déterminant principal du système. On a de même les équations et inconnues principales. Si $A' = (a_{i,j})_{i \in I, j \in J}$, on appelle déterminants caractéristiques les déterminants d'ordre $r+1$ obtenus en ajoutant sous A' une ligne de la forme $(a_{k,j})_{j \in J}$ puis une colonne de la forme $((b_i)_{i \in I}, b_k)$ avec $k \notin J$. Ils n'existent que si $p < r$, et sont alors au nombre de $p - r$. Le système (S) admet alors des solutions si et seulement si $p = r$ ou les $p - r$ déterminants caractéristiques sont nuls. Le système est alors équivalent au système des équations principales, les inconnues principales étant déterminées par un système de Cramer à l'aide des inconnues non principales.

Exemple. Si on prend (S) égal à

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 1 \\ x_1 \qquad - x_3 - x_4 = 1 \\ -x_1 + x_2 + x_3 + 2x_4 = m \end{cases}$$

pour m réel, le système n'admet des solutions que si $m = -1$ et dans ce cas $x_1 = 1 + x_3 + x_4$ et $x_2 = -x_4$.

2.3 Déterminants de Gram

Ici E est un espace préhilbertien.

Définition. Soit $(x_1, \ldots, x_n)$ une famille de vecteurs de E. La *matrice de Gram* associé à cette famille est la matrice :

$$M = (\langle x_i, x_j \rangle)_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R}).$$

Son déterminant est appelé *déterminant de Gram* de la famille $(x_1, \ldots, x_n)$.

Proposition. Les matrices de Gram de taille n sont exactement les matrices symétriques positives de taille n.

Proposition. Soit $(x_1, \ldots, x_n)$ une famille de vecteurs de E. Elle est liée si et seulement si son déterminant de Gram est nul.

Théorème. Soit V un sous-espace vectoriel de E admettant une base $(x_1, \ldots, x_n)$. Alors pour tout $u \in E$:

$$d(u, V)^2 = \frac{G(u, x_1, \ldots, x_n)}{G(x_1, \ldots, x_n)}$$

Application. L'expression $\int_0^1 (1 + a_1 + a_2 x^2 + \ldots + a_n x^n)^2 \, \mathrm{d}x$ atteint son minimum $\mu = \frac{1}{(n+1)^2}$ en un unique n-uplet.

Théorème. (Inégalité de Hadamard)
Soit $(x_1, \ldots, x_n)$ une famille de vecteurs non nuls de E. Alors

$$G(x_1, \ldots, x_n) \leq \prod_{i=1}^{n} \|x_i\|^2,$$

avec égalité si et seulement si la famille en question est orthogonale.
Plus généralement,

$$|\det(x_1, \ldots, x_n)| \leq \prod_{i=1}^{n} \|x_i\|.$$

3 Propriétés géométriques et analytiques

On prend ici $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

3.1 Volume

Ici $K = \mathbb{R}$.

Proposition. Le déterminant d'une famille de vecteurs dans la base canonique correspond au volume orienté du parallélotope engendré par cette famille de vecteurs.

Remarque. En particulier, l'inégalité de Hadamard dit que le volume d'un parallélotope est maximal quand celui-ci est rectangle.

Exemple. Soit ABC un triangle dans le plan affine. Alors l'aire de ABC est égale à $\frac{1}{2}|\det(\overrightarrow{AB}, \overrightarrow{BC})|$.

Proposition. Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$ et K un compact de $\mathbb{R}^n$. Si on note λ la mesure de Lebesgue sur $\mathbb{R}^n$, alors $\lambda(AK) = |\det(A)|\lambda(K)$.

Remarque. C'est ce résultat qui donne la première intuition du jacobien comme modification élémentaire de volume.

3.2 Continuité et différentielle

Proposition. Le déterminant est polynomial en les coefficients de la matrice, donc est continu et même $\mathcal{C}^\infty$.

Proposition. Le gradient du déterminant est la comatrice.

Proposition. Le groupe $\mathrm{GL}_n(\mathbb{K})$ est ouvert et $\mathrm{SL}_n(\mathbb{K})$ est fermé.

Proposition. Le groupe $\mathrm{SL}_n(\mathbb{R})$ est une variété de dimension $n^2 - 1$.

3.3 Convexité logarithmique

Proposition. Le déterminant est strictement logarithmiquement convexe sur $\mathcal{S}_n^{++}(\mathbb{R})$.

Théorème. Soit K un compact de $\mathbb{R}^n$ dont l'intérieur contient 0. Il existe un unique ellipsoïde centré en 0 de volume minimal contenant K.

Chapitre 19

150. Polynômes d'endomorphisme en dimension finie. Réduction d'un endomorphisme en dimension finie. Applications.

Présentation du plan

On considère, dans tout le plan, un corps K, et E un K-espace vectoriel de dimension finie d.

1 L'algèbre des polynômes d'un endomorphisme

1.1 Généralités

Proposition. Soit $u \in \mathcal{L}(E)$. Il existe un unique morphisme d'algèbres unitaire φ de $K[X]$ sur $\mathcal{L}(E)$ tel que $\varphi(X) = u$. Il est donné par l'expression $\varphi(\sum_k a_k X^k) = \sum_k a_k u^k =: (\sum_k a_k X^k)(u)$. On note $K[u]$ l'ensemble des polynômes en u, l'image de $K[X]$ par φ.

Remarque. Puisque E est de dimension finie, $K[u]$ est de dimension finie. En particulier, il existe au moins un polynôme $P \in K[X]$ tel que $P(u) = 0$. On appelle un tel polynôme un *polynôme annulateur* de u. Si ce polynôme annulateur a un coefficient constant non nul, alors u est inversible et son inverse est un élément de $K[u]$.

Exemple. Un polynôme annulateur est très utile pour calculer les puissances d'une matrice. En effet, en effectuant une division euclidienne par ce polynôme annulateur, on se ramène à évaluer un polynôme de degré constant (dont le degré ne dépend pas de la puissance)

Proposition. Si u et v sont deux endomorphismes de E qui commutent, alors pour tout $P \in K[X]$, le noyau et l'image de $P(u)$ sont stables par v.

Théorème. (lemme des noyaux)
Soient P, Q deux polynômes de $K[X]$ premiers entre eux. Alors pour $u \in \mathcal{L}(E)$:

$$\ker((PQ)(u)) = \ker(P(u)) \oplus \ker(Q(u)).$$

Application. Si u possède un polynôme annulateur factorisable en un produit d'irréductibles, alors u s'écrit comme une matrice diagonale par blocs, où chaque bloc est de dimension $\dim(\ker(P_i^{m_i}(u)))$.

1.2 Polynôme minimal

Définition. Soit $u \in \mathcal{L}(E)$. L'ensemble des polynômes annulateurs de u est un idéal de $K[X]$. On note Π_u l'unique générateur unitaire de cet idéal, appelé *polynôme minimal* de u.

Proposition. Le morphisme φ défini précédemment induit un isomorphisme de $K[X]/\langle \Pi_u \rangle$ sur $K[u]$. En particulier, $K[u]$ a pour dimension $\deg(\Pi_u)$, et c'est un corps si et seulement si u est irréductible sur K.

Remarque. Un endomorphisme u de polynôme minimal $a_0 + a_1 X + \ldots + X^r$ est inversible dans $\mathcal{L}(E)$ si et seulement si $a_0 \neq 0$, et on a alors

$$u^{-1} = -a_0^{-1}(a_1 \mathrm{id}_E + a_2 u + \ldots + u^{r-1}).$$

On peut calculer l'inverse par la même méthode avec un polynôme annulateur quelconque.

Exemple. On considère dans $\mathcal{M}_3(\mathbb{R})$ la matrice

$$M = \begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}$$

Son polynôme minimal est $(X+1)^2$. Ainsi, M est inversible avec

$$M^{-1} = -M - 2I_3 = \begin{pmatrix} -5 & 0 & -8 \\ -3 & -1 & -6 \\ 2 & 0 & 3 \end{pmatrix}.$$

On voit que la méthode est pratique, il faudrait donc que l'on ait un moyen simple de trouver le polynôme minimal de M (ou au moins un polynôme annulateur).

Proposition. Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E stable par u. Alors $\Pi_{u_{|F}} | \Pi_u$.

1.3 Cas des nilpotents

Définition. Un endomorphisme est dit *nilpotent* d'indice r si $u^r = 0$ sans que u^{r-1} soit nul.

Proposition. Soit $u \in \mathcal{L}(E)$, u est nilpotent d'indice r si et seulement si $\Pi_u = X^r$.

Lemme. (de Fitting)
Pour u nilpotent, il existe un n_0 tel que $E = \ker(u^{n_0}) \oplus \mathrm{Im}(u^{n_0})$.

Application. Le nombre de matrices nilpotentes sur $\mathbb{F}_q$ et de taille d est $q^{d(d-1)}$.

Théorème. (réduction des endomorphismes nilpotents)
Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent. Il existe une unique suite d'entiers $n_1 \leq \ldots \leq n_k$ et

une base de E telles que dans cette base, la matrice de u soit de la forme

$$\begin{pmatrix} J_{n_1} & & (0) \\ & \ddots & \\ (0) & & J_{n_k} \end{pmatrix}$$

où pour $m \in \mathbb{N}^*$, J_m est la matrice de $\mathcal{M}_m(K)$ définie par

$$J_m = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{pmatrix}.$$

2 Réduction des endomorphismes

2.1 Valeurs propres, vecteurs propres, sous-espaces propres

Définition. Soit $u \in \mathcal{L}(E)$. Un *vecteur propre* de u pour la *valeur propre* $\lambda \in K$ est un élément $x \in E$ non nul tel que $u(x) = \lambda x$. On note $E_\lambda(u) := \ker(u - \lambda \mathrm{id}_E)$ le *sous-espace propre* de u pour la valeur propre λ. On note $\mathrm{Sp}(u) = \mathrm{Sp}_K(u)$ l'ensemble des valeurs propres de u, appelé *spectre* de u.

Proposition. Une famille de vecteurs propres associés à des valeurs propres distinctes est libre. En particulier, $u \in \mathcal{L}(E)$ possède un nombre fini de valeurs propres.

Proposition. Si deux endomorphismes commutent, tout sous-espace propre de l'un est stable par l'autre.

Proposition. Les valeurs propres de u figurent parmi les racines d'un polynôme annulateur de u. Plus précisément, les valeurs propres de u sont exactement les racines de Π_u.

Définition. Soit $u \in \mathcal{L}(E)$. On appelle *polynôme caractéristique* de u le polynôme $\chi_u(X) := \det(X\mathrm{id}_E - u) \in K[X]$.

Proposition. Le polynôme caractéristique de u est unitaire de degré n, son coefficient constant est $(-1)^n \det(u)$ et son coefficient en X^{n-1} est $-\mathrm{Tr}(u)$.

Proposition. Le polynôme caractéristique est invariant par similitude.

Proposition. Les valeurs propres de u sont exactement les racines de χ_u. De plus, la multiplicité de λ en tant que racine de χ_u (multiplicité algébrique) est supérieure à $\dim(E_\lambda(u))$ (multiplicité géométrique).

Remarque. Un endomorphisme est nilpotent si et seulement si son polynôme caractéristique est X^n.

Théorème. (de Cayley-Hamilton)
Le polynôme caractéristique de u est un polynôme annulateur de u.

Corollaire. La dimension de $K[u]$ est inférieure ou égale à n. Par ailleurs, $(K[u])^\times = K[u] \cap \mathrm{GL}(E)$.

Remarque. Un endomorphisme est nilpotent si et seulement si son polynôme caractéristique est X^n.

2.2 Diagonalisabilité

Définition. Un endomorphisme est dit *diagonalisable* s'il admet une base de vecteurs propres. Cela équivaut à l'existence d'une base dans laquelle u est diagonale, ou encore à l'identité :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u).$$

Exemple. Un projecteur et une symétrie sont diagonalisables.

Théorème. Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

- u est diagonalisable
- $\sum_{\lambda \in \mathrm{Sp}(u)} \dim(E_\lambda(u)) = n$
- χ_u est scindé sur K et les multiplicités algébrique et géométrique de u coïncident.
- u possède un polynôme annulateur scindé à racines simples
- Π_u est scindé à racines simples.

Exemple. Considérons la matrice

$$A = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0. \end{pmatrix}$$

Son polynôme caractéristique est $X^3 - 5X^2 + 8X - 4$, on vérifie que 1 en est une racine puis on factorise pour avoir $\chi_A = (X-1)(X-2)^2$. On vérifie que le sous-espace propre associé à la valeur propre 1 est $\mathrm{Vect}((1,1,-1))$ en résolvant un système. Pour la valeur propre 2, on résout aussi le système, qui se ramène à l'équation $2x - 3y - 2z = 0$, équation du plan $\mathrm{Vect}((3,2,0),(1,0,1))$. Ainsi A est diagonalisable sur $\mathbb{R}$ et

$$A = P\mathrm{diag}(1,2,2)P^{-1}$$

avec

$$P = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Remarquons qu'il est alors facile de calculer les puissances de A.

Théorème. Soit (u_i) une famille d'endomorphismes diagonalisables qui commutent deux à deux. Alors ils sont diagonalisables dans une même base.

Application. On suppose que la caractéristique de K n'est pas 2. Alors si $\mathrm{GL}_m(K) \simeq \mathrm{GL}_n(K)$, nécessairement $n = m$.

2.3 Trigonalisabilité

Définition. Soit $u \in \mathcal{L}(E)$, on dit que u est *trigonalisable* s'il existe une base $(e_1, \dots, e_n)$ de E tel que pour tout k, $u(e_k) \in \mathrm{Vect}(e_1, \dots, e_n)$, *i.e.* une base dans laquelle la matrice de u est triangulaire supérieure.

Remarque. Dans une forme matricielle triangulaire supérieure, les valeurs sur la diagonale sont forcément les valeurs propres de l'endomorphisme.

Théorème. Un endomorphisme est trigonalisable si et seulement si son polynôme minimal est scindé, si et seulement si son polynôme caractéristique est scindé, si et seulement s'il possède un polynôme annulateur scindé.

Corollaire. Sur un corps algébriquement clos, tout endomorphisme est trigonalisable.

Remarque. Si u est trigonalisable, son déterminant est égal au produit de ses valeurs propres, et sa trace à la somme de ses valeurs propres.

Exemple. Soit $A \in \mathcal{M}_n(\mathbb{C})$, A est nilpotente si et seulement si $\chi_A = X^n$, si et seulement si pour tout $k \geq 1$, $\mathrm{Tr}(A^k) = 0$.

Remarque. Ce résultat est aussi vrai dans un corps de caractéristique nulle.

Application. Tout sous-groupe d'exposant fini de $\mathrm{GL}_n(\mathbb{C})$ est fini.

Théorème. (formule de Gelfand)
Soit $\|\cdot\|$ une norme quelconque sur $\mathcal{M}_n(\mathbb{C})$, alors si $\rho(A)$ est le plus grand module des valeurs propres de A,

$$\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}.$$

Théorème. Soit (u_i) une famille d'endomorphismes trigonalisables qui commutent deux à deux. Alors ils sont trigonalisables dans une même base.

2.4 Décompositions de Dunford et Jordan

Théorème. (Décomposition de Jordan-Chevalley-Dunford)
Soit u un endomorphisme de $\mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Alors il existe un unique couple (d, n) d'endomorphismes de $\mathcal{L}(E)$, avec d diagonalisable et n nilpotent, $dn = nd$. De plus d et n sont des polynômes en u.

Remarque. Cela se généralise à n'importe quel endomorphisme, mais d est alors semi-simple. La preuve n'est alors pas la même puisque l'on ne peut pas écrire notre décomposition du début.

Exemple. Pour calculer une décomposition de Dunford, il faut calculer une relation de Bezout pour avoir les projecteurs sur les sous-espaces caractéristiques, on bien directement trouver une base de ces sous-espaces pour avoir la forme matricielle de l'endomorphisme que l'on peut manipuler simplement.

Notre décomposition fait apparaître un endomorphisme nilpotent, que l'on sait réduire d'après la première partie. On obtient alors le théorème de réduction de Jordan.

Théorème. (de réduction de Jordan)
Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Il existe une base de E dans laquelle la matrice de u est de la forme

$$\begin{pmatrix} \lambda_1 I_{n_1} + J_{n_1} & & (0) \\ & \ddots & \\ (0) & & \lambda_k I_{n_k} + J_{n_k}. \end{pmatrix}$$

Exemple. On considère la matrice :

$$A = \begin{pmatrix} -1 & 8/3 & -2/3 & 0 & -2/3 \\ 0 & -1 & 0 & 0 & 0 \\ 2 & 7 & -1 & 0 & 4 \\ 2 & 10/3 & -1/3 & 2 & 11/3 \\ 1 & 2 & 0 & 0 & 1 \end{pmatrix}.$$

En développant selon la quatrième colonne et avec un peu de calcul, on trouve $\chi_A = X^5 - 2X^3 - 2X^2 - 3X^2$. On cherche des racines triviales. On trouve alors que -1 et 2 sont des racines de χ_A. Pour connaître les multiplicités, on calcule les dérivées en les racines et on trouve alors que -1 est racine double tandis que 2 est racine simple. On factorise alors χ_A par $(X+1)^2(X-2)$, et on trouve finalement :

$$\chi_A = (X^2+1)(X+1)^2(X-2).$$

Ainsi, A n'est pas trigonalisable sur $\mathbb{R}$. Sur $\mathbb{C}$, elle n'est pas diagonalisable ; on peut vérifier que $\dim(E_{-1}(A)) = 1 < 2$. On peut par contre trigonaliser A sur $\mathbb{C}$. On commence par le plus simple, trouver des vecteurs propres pour les valeurs propres 2, i et $-$i. Pour cela, on résout des systèmes et on trouve respectivement $(1-\mathrm{i}, 0, -2, \mathrm{i}, 1)$, $(1+\mathrm{i}, 0, -2, -\mathrm{i}, -1)$ et $(0, 0, 0, -1, 0)$. On prend ensuite un vecteur de $E_{-1}(A)$, par exemple $w := (-2, 0, -2, 0, 1)$. Enfin, on en cherche un dernier vecteur v tel que $Av = 2v + w$. On résout alors un dernier système pour obtenir $(1, -1, -2, -1, 1)$. Ainsi, si on pose

$$P = \begin{pmatrix} 1-\mathrm{i} & 1+\mathrm{i} & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ -2 & -2 & -1 & -2 & 0 \\ \mathrm{i} & -\mathrm{i} & 0 & -1 & -1 \\ -1 & -1 & 1 & 1 & 0 \end{pmatrix},$$

on a

$$A = P \begin{pmatrix} \mathrm{i} & 0 & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} P^{-1}.$$

Application. Toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable à sa transposée. Pour un corps quelconque, il faudrait passer par une clôture algébrique et utiliser que deux matrices à valeurs dans un corps et semblables dans une extension sont semblables sur le corps de base.

3 Exponentielle de matrices

On prend ici $K = \mathbb{C}$.

3.1 Généralités

Définition. Soit $A \in \mathcal{M}_n(\mathbb{C})$, la série $\sum_{k\in\mathbb{N}} \frac{A^k}{k!}$ est absolument convergente. On définit $\exp(A)$ la somme de cette série, appelée *exponentielle* de A.

Proposition. L'application exp est continue de $\mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$.

Proposition. Soient $A, B \in \mathcal{M}_n(\mathbb{C})$. Alors :

- $\exp(A)$ est un polynôme en A donc commute avec A
- Si B est inversible, $\exp(BAB^{-1}) = B\exp(A)B^{-1}$
- Si A est diagonalisable, $\exp(A)$ aussi
- $\det(\exp(A)) = \exp(\mathrm{Tr}(A))$

Proposition. La fonction $t \mapsto \exp(tA)y_0$ est l'unique solution du problème de Cauchy $y' = Ay$ avec la donnée initiale $y(0) = y_0$.

Corollaire. Si A et B sont deux matrices qui commutent, alors $\exp(A + B) = \exp(A)\exp(B)$. En particulier, $\exp(A)$ est inversible, d'inverse $\exp(-A)$.

Application. Si $A = D+N$ est la décomposition de Dunford de A, alors $\exp(A) = \exp(D)\exp(N) = \exp(D) + \exp(D)(\exp(N) - I_n)$, *i.e.* on connaît la décomposition de Dunford de $\exp(A)$.

Exemple. En gardant la matrice A que l'on a précédemment réduite sous forme de Jordan, on a

$$\exp(tA) = P \begin{pmatrix} \mathrm{e}^{\mathrm{i}t} & 0 & 0 & 0 & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}t} & 0 & 0 & 0 \\ 0 & 0 & \mathrm{e}^{-t} & t & 0 \\ 0 & 0 & 0 & \mathrm{e}^{-t} & 0 \\ 0 & 0 & 0 & 0 & \mathrm{e}^{-2t} \end{pmatrix} P^{-1}.$$

On calcule finalement ceci avec les expressions de P et P^{-1} : $\exp(tA)=$

$$\begin{pmatrix} \cos(t) + \sin(t) & \frac{11}{3}\cos(t)+3\sin(t) \\ -\frac{11}{3}\mathrm{e}^{-t}+2t & -\frac{2}{3}\cos(t) + \frac{3}{2}\mathrm{e}^{-t} & 0 & 2\sin(t) + \frac{4}{3}\cos(t) - \frac{4}{3}\mathrm{e}^{-t} \\ 0 & \mathrm{e}^{-t} & 0 & 0 & 0 \\ -2\sin(t) & -\frac{2}{3}\cos(t)-\frac{20}{3}\sin(t) \\ +\frac{2}{3}\mathrm{e}^{-t}+t & \frac{2}{3}\cos(t) + \frac{2}{3}\sin(t) + \frac{1}{3}\mathrm{e}^{-t} & 0 & \frac{2}{3}\cos(t) - \frac{10}{3}\sin(t) - \frac{2}{3}\mathrm{e}^{-t} \\ -\cos(t) + \mathrm{e}^{2t} & \frac{1}{3}\sin(t)-\frac{10}{3}\cos(t) \\ +\mathrm{e}^{-t}+\frac{7}{3}\mathrm{e}^{2t} & \frac{1}{3}\cos(t) - \frac{1}{3}\sin(t) - \frac{1}{3}\mathrm{e}^{2t} & \mathrm{e}^{2t} & -\frac{10}{3}\sin(t) + \frac{2}{3}\cos(t) - \frac{2}{3}\mathrm{e}^{-t} \\ -\sin(t) & -\frac{10}{3}\sin(t)-\frac{1}{3}\cos(t) \\ +\frac{1}{3}\mathrm{e}^{-t}-t & \frac{1}{3}\sin(t) + \frac{1}{3}\cos(t) - \frac{1}{3}\mathrm{e}^{-t} & 0 & -\frac{10}{3}\sin(t) + \frac{1}{3}\cos(t) + \frac{2}{3}\mathrm{e}^{-t}. \end{pmatrix}.$$

3.2 Logarithme matriciel

Définition. Soit M une matrice nilpotente. On définit

$$\ln(I_n - M) = \sum_{k=1}^{n-1} \frac{M^k}{k} \in K[M].$$

Proposition. Soit M une matrice nilpotente. Alors $\exp(\ln(I_n - M)) = I_n - M$.

Théorème. L'exponentielle de matrices est surjective de $\mathcal{M}_n(\mathbb{C})$ dans $\mathrm{GL}_n(\mathbb{C})$. De plus un antécédent d'une matrice inversible M par exp est un polynôme en M.

Corollaire. L'image de exp sur $\mathcal{M}_n(\mathbb{R})$ est constituée des carrés de matrices de $\mathrm{GL}_n(\mathbb{R})$.

3.3 Comportement asymptotique des solutions d'une équation différentielle

On s'intéresse au problème de Cauchy $y' = Ay$ pour $A \in \mathcal{M}_n(\mathbb{C})$ avec la condition initiale $y(0) = y_0$. On sait que ce problème admet une unique solution qui est $t \mapsto \exp(tA)y_0$. On s'intéresse au comportement de la solution en grand temps.

Définition. On note $\chi_A = \prod_\lambda (X - \lambda)^{m_\lambda}$. On définit les espaces :

- *stable* de A par $E_\mathrm{S} = \ker(P_\mathrm{S}(A))$ avec $P_\mathrm{S} = \prod_{\mathrm{Re}(\lambda)<0}(X - \lambda)^{m_\lambda}$
- *instable* de A par $E_\mathrm{S} = \ker(P_\mathrm{I}(A))$ avec $P_\mathrm{I} = \prod_{\mathrm{Re}(\lambda)>0}(X - \lambda)^{m_\lambda}$
- *neutre* de A par $E_\mathrm{S} = \ker(P_\mathrm{N}(A))$ avec $P_\mathrm{N} = \prod_{\mathrm{Re}(\lambda)=0}(X - \lambda)^{m_\lambda}$

Théorème. On a la somme directe $\mathbb{C}^n = E_\mathrm{S} \oplus E_\mathrm{I} \oplus E_\mathrm{N}$, avec

$$E_\mathrm{S} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to\infty} \exp(tA)Y_0 = 0\}, E_\mathrm{I} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to-\infty} \exp(tA)Y_0 = 0\}$$

et

$$E_\mathrm{N} = \{Y_0 \in \mathbb{C}^n, \exists C > 0, \exists p \in \mathbb{N}, \forall t \in \mathbb{R}, \|\exp(tA)Y_0\| \leq C(1 + |t|)^p\}.$$

Chapitre 20

151. Sous-espaces stables par un endomorphisme ou une famille d'endomorphismes d'un espace vectoriel de dimension finie. Applications.

Sommaire

1 Sous-espaces stables par un endomorphisme 146
1.1 Généralités 146
1.2 Polynômes d'endomorphismes 146
2 Diagonalisation, triagonalisation 147
2.1 Sous-espaces propres 147
2.2 Réduction 147
2.3 Familles d'endomorphismes qui commutent 148
3 Théorèmes de réduction 149
3.1 Sous-espaces caractéristiques et décomposition de Dunford 149
3.2 Endomorphismes nilpotents 149
3.3 Réduction de Frobenius 150

Présentation du plan

Dans tout le plan, E est un K-espace vectoriel de dimension finie n.

1 Sous-espaces stables par un endomorphisme

1.1 Généralités

Définition. Soit $u \in \mathcal{L}(E)$. On dit que F (sous-espace vectoriel de E) est *stable* par u si $u(F) \subset F$. Ainsi restreindre u à F donne un endomorphisme de F.

Exemple. Le noyau et l'image de $u \in \mathcal{L}(E)$ sont toujours stable par u.

Remarque. Sur $\mathbb{R}$ un endomorphisme admet toujours une droite ou un plan stable.

Remarque. Un endomorphisme qui stabilise tous les sous-espaces est une homothétie (et réciproquement).

Proposition. Si on a $E = \bigoplus_{i=1}^{r} F_i$ une somme directe de sous-espaces stables par $u \in \mathcal{L}(E)$, alors en prenant une base adaptée à cette décomposition, on obtient une matrice diagonale par blocs de u. Réciproquement, une telle forme matricielle donne une décomposition en somme directe de sous-espaces stables.

1.2 Polynômes d'endomorphismes

Proposition. Soit $u \in \mathcal{L}(E)$. Il existe un unique morphisme d'algèbres unitaire φ de $K[X]$ sur $\mathcal{L}(E)$ tel que $\varphi(X) = u$. Il est donné par l'expression $\varphi(\sum_k a_k X^k) = \sum_k a_k u^k =: (\sum_k a_k X^k)(u)$. On note $K[u]$ l'ensemble des polynômes en u, l'image de $K[X]$ par φ.

Remarque. Puisque E est de dimension finie, $K[u]$ est de dimension finie. En particulier, il existe au moins un polynôme $P \in K[X]$ tel que $P(u) = 0$. On appelle un tel polynôme un *polynôme annulateur* de u. Si ce polynôme annulateur a un coefficient constant non nul, alors u est inversible et son inverse est un élément de $K[u]$.

Proposition. Si u et v sont deux endomorphismes de E qui commutent, alors pour tout $P \in K[X]$, le noyau et l'image de $P(u)$ sont stables par v.

Théorème. (lemme des noyaux)
Soient P, Q deux polynômes de $K[X]$ premiers entre eux. Alors pour $u \in \mathcal{L}(E)$:

$$\ker((PQ)(u)) = \ker(P(u)) \oplus \ker(Q(u)).$$

Application. Si u possède un polynôme annulateur factorisable en un produit d'irréductibles, alors u s'écrit comme une matrice diagonale par blocs, où chaque bloc est de dimension $\dim(\ker(P_i^{m_i}(u)))$.

Proposition. (décomposition de Fitting)
Soit $u \in \mathcal{L}(E)$. Il existe un unique couple (F, G) de sous-espaces de somme directe dans E stables par u tel que la restriction de u à F est bijective et que celle de u à G soit nilpotente.

Application. Si K est fini de cardinal q, il y a $q^{d(d-1)}$ endomorphismes nilpotents sur K^d.

Proposition. Si F est stable par u alors $\chi_{u_{|F}} | \chi_u$, idem pour le polynôme minimal.

2 Diagonalisation, triagonalisation

2.1 Sous-espaces propres

Définition. Soit $u \in \mathcal{L}(E)$. Un *vecteur propre* de u pour la *valeur propre* $\lambda \in K$ est un élément $x \in E$ non nul tel que $u(x) = \lambda x$. On note $E_\lambda(u) := \ker(u - \lambda \mathrm{id}_E)$ le *sous-espace propre* de u pour la valeur propre λ. On note $\mathrm{Sp}(u) = \mathrm{Sp}_K(u)$ l'ensemble des valeurs propres de u, appelé *spectre* de u.

Proposition. Une famille de vecteurs propres associés à des valeurs propres distinctes est libre. En particulier, $u \in \mathcal{L}(E)$ possède un nombre fini de valeurs propres.

Proposition. Si deux endomorphismes commutent, tout sous-espace propre de l'un est stable par l'autre.

Proposition. Les valeurs propres de u figurent parmi les racines d'un polynôme annulateur de u. Plus précisément, les valeurs propres de u sont exactement les racines de Π_u.

Proposition. Les valeurs propres de u sont exactement les racines de χ_u. De plus, la multiplicité de λ en tant que racine de χ_u (multiplicité algébrique) est supérieure à $\dim(E_\lambda(u))$ (multiplicité géométrique).

Corollaire. La dimension de $K[u]$ est inférieure ou égale à n. Par ailleurs, $(K[u])^\times = K[u] \cap \mathrm{GL}(E)$.

2.2 Réduction

Définition. Un endomorphisme est dit *diagonalisable* s'il admet une base de vecteurs propres. Cela équivaut à l'existence d'une base dans laquelle u est diagonale, ou encore à l'identité :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u).$$

Exemple. Un projecteur et une symétrie sont diagonalisables.

Théorème. Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

- u est diagonalisable
- $\sum_{\lambda \in \mathrm{Sp}(u)} \dim(E_\lambda(u)) = n$
- χ_u est scindé sur K et les multiplicités algébrique et géométrique de u coïncident.
- u possède un polynôme annulateur scindé à racines simples
- Π_u est scindé à racines simples.

Exemple. Considérons la matrice

$$A = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0. \end{pmatrix}$$

Son polynôme caractéristique est $X^3 - 5X^2 + 8X - 4$, on vérifie que 1 en est une racine puis on factorise pour avoir $\chi_A = (X-1)(X-2)^2$. On vérifie que le sous-espace propre associé à la valeur propre 1 est $\mathrm{Vect}((1,1,-1))$ en résolvant un système. Pour la valeur propre 2, on résout aussi le système, qui se ramène à l'équation $2x - 3y - 2z = 0$, équation du plan $\mathrm{Vect}((3,2,0),(1,0,1))$. Ainsi A est diagonalisable sur $\mathbb{R}$ et

$$A = P\mathrm{diag}(1,2,2)P^{-1}$$

avec

$$P = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Remarquons qu'il est alors facile de calculer les puissances de A.

Définition. Soit $u \in \mathcal{L}(E)$, on dit que u est *trigonalisable* s'il existe une base $(e_1, \ldots, e_n)$ de E tel que pour tout k, $u(e_k) \in \mathrm{Vect}(e_1, \ldots, e_n)$, *i.e.* une base dans laquelle la matrice de u est triangulaire supérieure.

Remarque. Dans une forme matricielle triangulaire supérieure, les valeurs sur la diagonale sont forcément les valeurs propres de l'endomorphisme.

Théorème. Un endomorphisme est trigonalisable si et seulement si son polynôme minimal est scindé, si et seulement si son polynôme caractéristique est scindé, si et seulement s'il possède un polynôme annulateur scindé.

Corollaire. Sur un corps algébriquement clos, tout endomorphisme est trigonalisable.

Remarque. Si u est trigonalisable, son déterminant est égal au produit de ses valeurs propres, et sa trace à la somme de ses valeurs propres.

2.3 Familles d'endomorphismes qui commutent

Théorème. Soit (u_i) une famille d'endomorphismes diagonalisables qui commutent deux à deux. Alors ils sont diagonalisables dans une même base.

Application. On suppose que la caractéristique de K n'est pas 2. Alors si $\mathrm{GL}_m(K) \simeq \mathrm{GL}_n(K)$, nécessairement $n = m$.

Théorème. Soit (u_i) une famille d'endomorphismes trigonalisables qui commutent deux à deux. Alors ils sont trigonalisables dans une même base.

3 Théorèmes de réduction

3.1 Sous-espaces caractéristiques et décomposition de Dunford

Théorème. (Décomposition de Jordan-Chevalley-Dunford)
Soit u un endomorphisme de $\mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Alors il existe un unique couple (d, n) d'endomorphismes de $\mathcal{L}(E)$, avec d diagonalisable et n nilpotent, $dn = nd$. De plus d et n sont des polynômes en u.

Remarque. Cela se généralise à n'importe quel endomorphisme, mais d est alors semi-simple. La preuve n'est alors pas la même puisque l'on ne peut pas écrire notre décomposition du début.

Exemple. Pour calculer une décomposition de Dunford, il faut calculer une relation de Bezout pour avoir les projecteurs sur les sous-espaces caractéristiques, on bien directement trouver une base de ces sous-espaces pour avoir la forme matricielle de l'endomorphisme que l'on peut manipuler simplement.

Application. Si $A = D+N$ est la décomposition de Dunford de A, alors $\exp(A) = \exp(D)\exp(N) = \exp(D) + \exp(D)(\exp(N) - I_n)$, *i.e.* on connaît la décomposition de Dunford de $\exp(A)$.

3.2 Endomorphismes nilpotents

Remarque. Un endomorphisme est nilpotent si et seulement si son polynôme caractéristique est X^n.

Lemme. (de Fitting)
Pour u nilpotent, il existe un n_0 tel que $E = \ker(u^{n_0}) + \mathrm{Im}(u^{n_0})$.

Application. Le nombre de matrices nilpotentes sur $\mathbb{F}_q$ et de taille d est $q^{d(d-1)}$.

Théorème. (réduction des endomorphismes nilpotents)
Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent. Il existe une unique suite d'entiers $n_1 \leq \ldots \leq n_k$ et une base de E telles que dans cette base, la matrice de u soit de la forme

$$\begin{pmatrix} J_{n_1} & & (0) \\ & \ddots & \\ (0) & & J_{n_k} \end{pmatrix}$$

où pour $m \in \mathbb{N}^*$, J_m est la matrice de $\mathcal{M}_m(K)$ définie par

$$J_m = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{pmatrix}.$$

Théorème. (de réduction de Jordan)
Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Il existe une base de E dans laquelle la matrice de u est de la forme

$$\begin{pmatrix} \lambda_1 I_{n_1} + J_{n_1} & & (0) \\ & \ddots & \\ (0) & & \lambda_k I_{n_k} + J_{n_k}. \end{pmatrix}$$

Remarque. Cette théorie est notamment utile pour calculer (ou au moins étudier qualitativement) les solutions de systèmes différentiels de la formes $y' = Ay$, où A est une matrice.

3.3 Réduction de Frobenius

Définition. Soit $u \in \mathcal{L}(E)$, on dit que u est *cyclique* si l'une des propriétés suivantes (équivalentes) est satisfaite :

- il existe $x \in E$ tel que $(x, u(x), \ldots, u^{n-1}(x))$ est une base de E
- il existe une base de E dans laquelle la matrice de u est une matrice compagnon
- $\Pi_u = \chi_u$

Théorème. (réduction de Frobenius)
Soit $u \in \mathcal{L}(E)$, il existe des sous-espaces vectoriels de E $F_1, \ldots, F_r$ stables par u tels que pour tout i, la restriction u_i de u à F_i soit cyclique, de polynôme minimal P_i, avec la condition $P_{i+1}|P_i$. La suite de polynôme ne dépend que de u, $P_1 = \Pi_u$ et $P_1 \cdots P_r = \chi_u$.

Corollaire. Deux endomorphismes de E sont semblables si et seulement si ils sont les mêmes (P_i), que l'on appelle alors *invariants de similitude*.

Application. Si deux endomorphismes sur $\mathbb{R}^n$ sont semblables sur $\mathbb{C}$ alors ils le sont aussi sur $\mathbb{R}$.

Chapitre 21

152. Endomorphismes diagonalisables en dimension finie. Applications.

Sommaire

1 Endomorphismes diagonalisables 152
1.1 Définitions équivalentes 152
1.2 Codiagonalisabilité 153
2 Propriétés selon le corps de base 153
2.1 Cas du corps des complexes 153
2.2 Cas des corps finis 153
2.3 Matrices à coefficients entiers 153
3 Décomposition de Dunford 154
3.1 Le théorème 154
3.2 Exponentielle de matrices 154
4 Endomorphismes normaux 155
4.1 Le théorème spectral 155
4.2 Valeurs singulières 156
4.3 Décomposition polaire 156

Présentation du plan

Dans tout le plan, E est un K-espace vectoriel de dimension finie n.

1 Endomorphismes diagonalisables

1.1 Définitions équivalentes

Définition. Un endomorphisme est dit *diagonalisable* s'il admet une base de vecteurs propres. Cela équivaut à l'existence d'une base dans laquelle u est diagonale, ou encore à l'identité :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u).$$

Exemple. Un projecteur et une symétrie sont diagonalisables.

Théorème. Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

— u est diagonalisable

— $\sum_{\lambda \in \mathrm{Sp}(u)} \dim(E_\lambda(u)) = n$

— χ_u est scindé sur K et les multiplicités algébrique et géométrique de u coïncident.

— u possède un polynôme annulateur scindé à racines simples

— Π_u est scindé à racines simples.

Exemple. Considérons la matrice

$$A = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0. \end{pmatrix}$$

Son polynôme caractéristique est $X^3 - 5X^2 + 8X - 4$, on vérifie que 1 en est une racine puis on factorise pour avoir $\chi_A = (X-1)(X-2)^2$. On vérifie que le sous-espace propre associé à la valeur propre 1 est $\mathrm{Vect}((1,1,-1))$ en résolvant un système. Pour la valeur propre 2, on résout aussi le système, qui se ramène à l'équation $2x - 3y - 2z = 0$, équation du plan $\mathrm{Vect}((3,2,0),(1,0,1))$. Ainsi A est diagonalisable sur $\mathbb{R}$ et

$$A = P\mathrm{diag}(1,2,2)P^{-1}$$

avec

$$P = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Remarquons qu'il est alors facile de calculer les puissances de A.

Remarque. Si un endomorphisme est trigonalisable avec des valeurs propres distinctes, alors il est diagonalisable.

1.2 Codiagonalisabilité

Théorème. Soit (u_i) une famille d'endomorphismes diagonalisables qui commutent deux à deux. Alors ils sont diagonalisables dans une même base.

Remarque. La réciproque est vraie. Cependant, elle ne l'est pas pour des endomorphismes trigonalisables.

Application. On suppose que la caractéristique de K n'est pas 2. Alors si $\mathrm{GL}_m(K) \simeq \mathrm{GL}_n(K)$, nécessairement $n = m$.

2 Propriétés selon le corps de base

2.1 Cas du corps des complexes

Théorème. L'adhérence des diagonalisables de $\mathcal{M}_n(\mathbb{C})$ est $\mathcal{M}_n(\mathbb{C})$, et l'adhérence des diagonalisables de $\mathcal{M}_n(\mathbb{R})$ est l'ensemble des trigonalisables.

Application. On peut montrer le théorème de Cayley-Hamilton avec ce résultat.

Remarque. L'ensemble des diagonalisables sur $\mathbb{R}$ et $\mathbb{C}$ est connexe par arcs. Si on prend que les inversibles, ce n'est vrai que sur $\mathbb{C}$. Cela peut servir pour classifier les systèmes différentiels $y' = Ay$ par équivalence topologique.

2.2 Cas des corps finis

Proposition. Un endomorphisme $u \in \mathcal{M}_n(\mathbb{F}_q)$ est diagonalisable si et seulement si $X^q - X$ annule u.

Proposition. Le nombre de matrices diagonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|}.$$

2.3 Matrices à coefficients entiers

Proposition. Soit $M \in \mathrm{GL}_n(\mathbb{Z})$ d'ordre fini d. Alors M est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité.

Théorème. Soit M une matrice de $\mathrm{GL}_n(\mathbb{Z})$ d'ordre fini. L'ensemble des ordres possibles pour M est fini.

Théorème. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Alors son cardinal est borné par 3^{n^2}.

3 Décomposition de Dunford

3.1 Le théorème

Théorème. (Décomposition de Jordan-Chevalley-Dunford)
Soit u un endomorphisme de $\mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Alors il existe un unique couple (d, n) d'endomorphismes de $\mathcal{L}(E)$, avec d diagonalisable et n nilpotent, $dn = nd$. De plus d et n sont des polynômes en u.

Remarque. Cela se généralise à n'importe quel endomorphisme, mais d est alors semi-simple. La preuve n'est alors pas la même puisque l'on ne peut pas écrire notre décomposition du début.

Exemple. Pour calculer une décomposition de Dunford, il faut calculer une relation de Bezout pour avoir les projecteurs sur les sous-espaces caractéristiques, on bien directement trouver une base de ces sous-espaces pour avoir la forme matricielle de l'endomorphisme que l'on peut manipuler simplement.

3.2 Exponentielle de matrices

Définition. Soit $A \in \mathcal{M}_n(\mathbb{C})$, la série $\sum_{k\in\mathbb{N}} \frac{A^k}{k!}$ est absolument convergente. On définit $\exp(A)$ la somme de cette série, appelée *exponentielle* de A.

Proposition. L'application exp est continue de $\mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$.

Proposition. Soient $A, B \in \mathcal{M}_n(\mathbb{C})$. Alors :

- $\exp(A)$ est un polynôme en A donc commute avec A
- Si B est inversible, $\exp(BAB^{-1}) = B\exp(A)B^{-1}$
- Si A est diagonalisable, $\exp(A)$ aussi
- $\det(\exp(A)) = \exp(\mathrm{Tr}(A))$

Proposition. La fonction $t \mapsto \exp(tA)y_0$ est l'unique solution du problème de Cauchy $y' = Ay$ avec la donnée initiale $y(0) = y_0$.

Corollaire. Si A et B sont deux matrices qui commutent, alors $\exp(A+B) = \exp(A)\exp(B)$. En particulier, $\exp(A)$ est inversible, d'inverse $\exp(-A)$.

Application. Si $A = D+N$ est la décomposition de Dunford de A, alors $\exp(A) = \exp(D)\exp(N) = \exp(D) + \exp(D)(\exp(N) - I_n)$, *i.e.* on connaît la décomposition de Dunford de $\exp(A)$.

Théorème. L'exponentielle de matrices est surjective de $\mathcal{M}_n(\mathbb{C})$ dans $\mathrm{GL}_n(\mathbb{C})$. De plus tout antécédent d'une matrice inversible M par exp est un polynôme en M.

Corollaire. L'image de exp sur $\mathcal{M}_n(\mathbb{R})$ est constituée des carrés de matrices de $\mathrm{GL}_n(\mathbb{R})$.

On s'intéresse au problème de Cauchy $y' = Ay$ pour $A \in \mathcal{M}_n(\mathbb{C})$ avec la condition initiale $y(0) = y_0$. On sait que ce problème admet une unique solution qui est $t \mapsto \exp(tA)y_0$. On s'intéresse au comportement de la solution en grand temps.

Définition. On note $\chi_A = \prod_\lambda (X-\lambda)^{m_\lambda}$. On définit les espaces :

- *stable* de A par $E_\mathrm{S} = \ker(P_\mathrm{S}(A))$ avec $P_\mathrm{S} = \prod_{\mathrm{Re}(\lambda)<0}(X-\lambda)^{m_\lambda}$
- *instable* de A par $E_\mathrm{S} = \ker(P_\mathrm{I}(A))$ avec $P_\mathrm{I} = \prod_{\mathrm{Re}(\lambda)>0}(X-\lambda)^{m_\lambda}$
- *neutre* de A par $E_\mathrm{S} = \ker(P_\mathrm{N}(A))$ avec $P_\mathrm{N} = \prod_{\mathrm{Re}(\lambda)=0}(X-\lambda)^{m_\lambda}$

Théorème. On a la somme directe $\mathbb{C}^n = E_\mathrm{S} \oplus E_\mathrm{I} \oplus E_\mathrm{N}$, avec

$$E_\mathrm{S} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to\infty} \exp(tA)Y_0 = 0\}, E_\mathrm{I} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to-\infty} \exp(tA)Y_0 = 0\}$$

et

$$E_\mathrm{N} = \{Y_0 \in \mathbb{C}^n, \exists C > 0, \exists p \in \mathbb{N}, \forall t \in \mathbb{R}, \|\exp(tA)Y_0\| \leq C(1+|t|)^p\}.$$

4 Endomorphismes normaux

On considère ici E un espace hermitien ou euclidien.

4.1 Le théorème spectral

Définition. Un endomorphisme $u \in \mathcal{L}(E)$ est dit *normal* lorsqu'il commute à son adjoint.

Proposition. Soit $u \in \mathcal{L}(E)$. Si F est stable par u, alors son orthogonal est stable par u^*. Si de plus u est normal et F est un sous-espace propre de u, alors son orthogonal est stable par u.

Théorème. (réduction des normaux, cas hermitien)
Soit E hermitien et u normal. Alors u est diagonalisable en base orthonormée.

Remarque. Un endomorphisme u est normal si et seulement si son adjoint est un polynôme en u.

Théorème. (réduction des normaux, cas euclidien)
Soit E euclidien et u normal, il existe une base orthonormée de E telle que la matrice de u dans cette base est diagonale par blocs, avec un bloc diagonal et des blocs de taille 2×2 de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Remarque. On retrouve le théorème spectral, qui dit que tout endomorphisme réel symétrique est diagonalisable en base orthonormée, le théorème de réduction des isométries.

4.2 Valeurs singulières

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les racines des valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \mathrm{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)$.

Corollaire. On peut définir la *pseudo-inverse* de A, notée $A^\dagger$, comme $U\Sigma^\dagger V^\top$ avec $\Sigma^\dagger \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A)^{-1}, \ldots, \sigma_r(A)^{-1}, 0, \ldots, 0)$. Pour $b \in \mathbb{R}^p$, le vecteur $A^\dagger b$ est l'unique solution de $Ax = b$ au sens des moindres carrés de norme minimale.

Proposition. (Eckart-Young)
La matrice $U\Sigma_k V^\top$, diagonale dont la diagonale est constituée des k premières valeurs singulières de A, est la meilleure approximation de A de rang k pour la norme spectrale (subordonnée à la norme deux).

Corollaire. (minmax de Courant-Fischer pour les valeurs singulières)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $p \leq n$, alors si G_k est l'ensemble des sous-espaces de $\mathbb{R}^p$ de dimension k :

$$\sigma_k(A) = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Théorème. (normes de Ky Fan)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$ $(p \leq n)$. On appelle *k-ième norme de Ky Fan* la quantité $N_k(A) = \sigma_1(A) + \cdots + \sigma_k(A)$. Alors pour tout $k \in \{1, \ldots, p\}$ la k-ième norme de Ky Fan est une norme.

Remarque. La quantité $N_k(A)/N_p(A)$ permet de mesurer qu'ont les lignes de A à se compresser dans un sous-espace de dimension k. Plus précisément, la norme k de Ky Fan maximise la somme des projections des images par A d'une famille orthonormée de dimension k de $\mathbb{R}^p$ sur une famille orthonormée de dimension k de $\mathbb{R}^n$.

4.3 Décomposition polaire

Proposition. L'exponentielle matricielle réalise un homéomorphisme entre $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{S}_n^{++}(\mathbb{R})$.

Corollaire. L'application qui à $M \in \mathrm{GL}_n(\mathbb{R})$ associe sa décomposition polaire est un homéomorphisme.

Proposition. Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$ et K un compact de $\mathbb{R}^n$. Si on note λ la mesure de Lebesgue sur $\mathbb{R}^n$, alors $\lambda(AK) = |\det(A)|\lambda(K)$.

Remarque. C'est ce résultat qui donne la première intuition du jacobien comme modification élémentaire de volume.

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors G possède un produit scalaire invariant si et seulement si G est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Chapitre 22

153. Valeurs propres, vecteurs propres. Calculs exacts ou approchés d'éléments propres. Applications.

Sommaire

1 Valeurs propres, vecteurs propres 159
1.1 Valeurs propres, sous-espaces propres et polynômes annulateurs 159
1.2 Diagonalisation, trigonalisation 160
2 Recherche de valeurs propres 162
2.1 Méthodes générales 162
2.2 Rayon spectral 163
3 Matrices symétriques, hermitiennes 163
3.1 Théorème spectral, valeurs singulières 163
3.2 Formulations variationnelles 164

Présentation du plan

Dans tout le plan, E est un K-espace vectoriel de dimension finie n.

1 Valeurs propres, vecteurs propres

1.1 Valeurs propres, sous-espaces propres et polynômes annulateurs

Définition. Un vecteur $x \neq 0$ de E est un *vecteur propre* de $u \in \mathcal{L}(E)$ s'il existe $\lambda \in K$ tel que $u(x) = \lambda x$. On dit alors que λ est *valeur propre* de u. On note $E_\lambda = \ker(u - \lambda \mathrm{id}_E)$ le sous-espace propre de u associé à la valeur propre λ.

Remarque. On peut définir la même chose pour une matrice.

Proposition. Une famille de vecteurs propres associés à des valeurs propres distinctes est libre.

Application. La famille $(\mathrm{e}^{\lambda t})_{\lambda \in \mathbb{R}}$ est libre dans $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$.

Proposition. Un scalaire $\lambda \in K$ est valeur propre de u si et seulement si c'est une racine de son polynôme caractéristique, si et seulement si c'est une racine de son polynôme minimal.

Exemple. Ceci permet de calculer les valeurs propres d'une matrice 2×2 simplement.

1.2 Diagonalisation, trigonalisation

Définition. Un endomorphisme est dit *diagonalisable* s'il admet une base de vecteurs propres. Cela équivaut à l'existence d'une base dans laquelle u est diagonale, ou encore à l'identité :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u).$$

Exemple. Un projecteur et une symétrie sont diagonalisables.

Théorème. Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

- u est diagonalisable
- $\sum_{\lambda \in \mathrm{Sp}(u)} \dim(E_\lambda(u)) = n$
- χ_u est scindé sur K et les multiplicités algébrique et géométrique de u coïncident.
- u possède un polynôme annulateur scindé à racines simples
- Π_u est scindé à racines simples.

Exemple. Considérons la matrice

$$A = \begin{pmatrix} 0 & 3 & 2 \\ -2 & 5 & 2 \\ 2 & -3 & 0. \end{pmatrix}$$

Son polynôme caractéristique est $X^3 - 5X^2 + 8X - 4$, on vérifie que 1 en est une racine puis on factorise pour avoir $\chi_A = (X-1)(X-2)^2$. On vérifie que le sous-espace propre associé à la valeur propre 1 est $\mathrm{Vect}((1,1,-1))$ en résolvant un système. Pour la valeur propre 2, on résout aussi le système, qui se ramène à l'équation $2x - 3y - 2z = 0$, équation du plan $\mathrm{Vect}((3,2,0),(1,0,1))$. Ainsi A est diagonalisable sur $\mathbb{R}$ et

$$A = P\mathrm{diag}(1,2,2)P^{-1}$$

avec

$$P = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Remarquons qu'il est alors facile de calculer les puissances de A.

Définition. Soit $u \in \mathcal{L}(E)$, on dit que u est *trigonalisable* s'il existe une base $(e_1, \ldots, e_n)$ de E tel que pour tout k, $u(e_k) \in \mathrm{Vect}(e_1, \ldots, e_n)$, *i.e.* une base dans laquelle la matrice de u est triangulaire supérieure.

Remarque. Dans une forme matricielle triangulaire supérieure, les valeurs sur la diagonale sont forcément les valeurs propres de l'endomorphisme.

Théorème. Un endomorphisme est trigonalisable si et seulement si son polynôme minimal est scindé, si et seulement si son polynôme caractéristique est scindé, si et seulement s'il possède un polynôme annulateur scindé.

Corollaire. Sur un corps algébriquement clos, tout endomorphisme est trigonalisable.

Remarque. Si u est trigonalisable, son déterminant est égal au produit de ses valeurs propres, et sa trace à la somme de ses valeurs propres.

Proposition. Un endomorphisme est nilpotent si et seulement si son polynôme caractéristique est X^n ; il est nilpotent d'indice r si et seulement si son polynôme minimal est X^r.

Proposition. Soit $M \in \mathrm{GL}_n(\mathbb{Z})$ d'ordre fini d. Alors M est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité.

Théorème. Soit M une matrice de $\mathrm{GL}_n(\mathbb{Z})$ d'ordre fini. L'ensemble des ordres possibles pour M est fini.

Théorème. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Alors son cardinal est borné par 3^{n^2}.

Théorème. (réduction des endomorphismes nilpotents)
Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent. Il existe une unique suite d'entiers $n_1 \leq \ldots \leq n_k$ et une base de E telles que dans cette base, la matrice de u soit de la forme

$$\begin{pmatrix} J_{n_1} & & (0) \\ & \ddots & \\ (0) & & J_{n_k} \end{pmatrix}$$

où pour $m \in \mathbb{N}^*$, J_m est la matrice de $\mathcal{M}_m(K)$ définie par

$$J_m = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{pmatrix}.$$

Théorème. (de réduction de Jordan)
Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Il existe une base de E dans laquelle la matrice de u est de la forme

$$\begin{pmatrix} \lambda_1 I_{n_1} + J_{n_1} & & (0) \\ & \ddots & \\ (0) & & \lambda_k I_{n_k} + J_{n_k}. \end{pmatrix}$$

2 Recherche de valeurs propres

2.1 Méthodes générales

Proposition. (Lemme d'Hadamard)
Une matrice à diagonale strictement dominante est inversible.

Corollaire. (Disques de Gershgorin)
Les valeurs propres de $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{C})$ sont localisées dans $\bigcup_{i=1}^{n} \mathrm{D}(a_{i,i}, \sum_{j \neq i} |a_{i,j}|)$.

Remarque. En utilisant les matrices compagnon, cela peut aider à localiser les racines d'un polynôme.

Théorème. (méthode de la puissance)
Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $(\lambda_1, \ldots, \lambda_n)$ ses valeurs propres, on suppose que λ_1 est strictement supérieure en module à toutes les autres valeurs propres. Soit F le sous-espace caractéristique de A associé à λ_1 et soit $\mathbb{C}^n = F \oplus G$. Pour $w_0 \notin F$, on définit par récurrence $w_{n+1} = \frac{1}{\|Aw_n\|} Aw_n$. Alors

- Pour tout n, $w_n \neq 0$
- $\|Aw_n\| \to |\lambda_1|$
- $\left(\frac{\overline{\lambda_1}}{|\lambda_1|}\right) w_n \to v$ où v est vecteur propre associé à λ_1 de norme 1.
- Si la j-ième composante de v est non nulle, alors $\frac{(Aw_n)_j}{(w_n)_j}$ tend vers λ_1 quand n tend vers l'infini.

2.2 Rayon spectral

Définition. Soit $A \in \mathcal{M}_n(\mathbb{C})$, son *rayon spectral* est le plus grand module d'une valeur propre de A, noté $\rho(A)$.

Remarque. Le rayon spectral de A n'est pas forcément une valeur propre de A.

Proposition. Soit N une norme matricielle. Alors $\rho(A) \leq N(A)$.

Remarque. Le rayon spectral n'est pas une norme en général, puisqu'il ne vérifie pas l'axiome de séparation. Par contre, $A \mapsto \rho(A^*A)$ est une norme, c'est la norme induite par la norme euclidienne.

Proposition. Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $\varepsilon > 0$. Il existe une norme induite telle que $\|A\|_\varepsilon \leq \rho(A) + \varepsilon$. En particulier, le rayon spectral est l'infimum des $\|A\|$ quand $\|.\|$ parcourt l'ensemble des normes induites.

Corollaire. (formule de Gelfand)
Pour toute norme matricielle, $\rho(A) = \lim_{n\to\infty} \|A^n\|^{1/n}$.

Théorème. Soit A une matrice réelle strictement positive. Alors le rayon spectral de A est l'unique valeur propre de A de module $\rho(A)$. De plus, le sous espace propre associé est de dimension 1, engendré par un vecteur strictement positif.

Application. Soit $A > 0$ et $X_0 \geq 0$. On s'intéresse au système dynamique $X_{t+1} = AX_t$ pour $t \in \mathbb{N}^*$. Lorsque t tend vers l'infini, on a l'asymptotique suivante :

$$X_t \sim \rho(A)^t \langle v, X_0 \rangle w$$

où $w = \frac{x}{\|x\|_1}$ et v vecteur propre à gauche associé à $\rho(A)$ choisi de telle sorte que $\langle v, w \rangle = 1$.

3 Matrices symétriques, hermitiennes

3.1 Théorème spectral, valeurs singulières

Théorème. (Théorème spectral)
Une matrice symétrique réelle (resp. hermitienne) est diagonalisable en base orthonormée (resp. unitaire), avec des valeurs propres réelles.

Définition. Une matrice hermitienne est dite positive (resp. strictement) si ses valeurs propres sont (resp. strictement) positives. C'est équivalent à $\langle Ax, x \rangle \geq 0$ (resp. > 0) pour tout x.

Corollaire. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, la matrice A^*A est diagonalisable en base orthonormée.

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les racines des valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \mathrm{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)$.

Proposition. (Eckart-Young)
La matrice $U\Sigma_k V^\top$, diagonale dont la diagonale est constituée des k premières valeurs singulières de A, est la meilleure approximation de A de rang k pour la norme spectrale (subordonnée à la norme deux).

3.2 Formulations variationnelles

Proposition. Soit $S \in \mathcal{S}_p(\mathbb{R})$, alors

$$\max_{\substack{W \in \mathcal{M}_{k,p}(\mathbb{R}) \\ W^\top W = I_k}} \mathrm{Tr}(WSW^\top) = \sum_{i=1}^{k} \lambda_k,$$

avec égalité quand les lignes de W sont les k premiers vecteurs de la base orthonormée qui diagonalise S (ordonnés par ordre décroissant des valeurs propres).

Application. Cela sert en statistiques dans ce que l'on appelle l'*analyse en composantes principales*. La projection sur un sous-espace de dimension k qui maximise la variance de nos données est donnée par les k premiers axes propres de la matrice de covariance empirique.

Théorème. (minmax de Courant-Fischer)
Soit A une matrice hermitienne de taille n, dont les valeurs propres sont $\lambda_1 \geq \cdots \geq \lambda_n$. Si on note G_k l'ensemble des sous-espaces vectoriels de $\mathbb{C}^n$ de dimension k, alors

$$\lambda_k = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle.$$

Corollaire. (minmax de Courant-Fischer pour les valeurs singulières)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $p \leq n$, alors si G_k est l'ensemble des sous-espaces de $\mathbb{R}^p$ de dimension k :

$$\sigma_k(A) = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Théorème. (normes de Ky Fan)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$ $(p \leq n)$. On appelle *k-ième norme de Ky Fan* la quantité $N_k(A) = \sigma_1(A) + \cdots + \sigma_k(A)$. Alors pour tout $k \in \{1, \ldots, p\}$ la k-ième norme de Ky Fan est une norme.

Remarque. La quantité $N_k(A)/N_p(A)$ permet de mesurer qu'ont les lignes de A à se compresser dans un sous-espace de dimension k. Plus précisément, la norme k de Ky Fan maximise la somme des projections des images par A d'une famille orthonormée de dimension k de $\mathbb{R}^p$ sur une famille orthonormée de dimension k de $\mathbb{R}^n$.

Chapitre 23

155. Exponentielle de matrices. Applications.

Sommaire

1 Exponentielle de matrices 167
1.1 Définitions équivalentes 167
1.2 Injectivité et surjectivité 168
2 Exponentielle d'endomorphismes remarquables 168
2.1 Matrices symétriques 168
2.2 Matrices antisymétriques, de trace nulle 169
3 Systèmes différentiels linéaires 169
3.1 Résolution 169
3.2 Comportement qualitatif 169

Présentation du plan

1 Exponentielle de matrices

1.1 Définitions équivalentes

Définition. Soit $A \in \mathcal{M}_n(\mathbb{C})$, la série $\sum_{k\in\mathbb{N}} \frac{A^k}{k!}$ est absolument convergente. On définit $\exp(A)$ la somme de cette série, appelée *exponentielle* de A.

Proposition. L'application exp est continue de $\mathcal{M}_n(\mathbb{C})$ dans $\mathcal{M}_n(\mathbb{C})$.

Proposition. Soient $A, B \in \mathcal{M}_n(\mathbb{C})$. Alors :

— $\exp(A)$ est un polynôme en A donc commute avec A

— Si B est inversible, $\exp(BAB^{-1}) = B\exp(A)B^{-1}$

— Si A est diagonalisable, $\exp(A)$ aussi

— $\det(\exp(A)) = \exp(\mathrm{Tr}(A))$

Proposition. La fonction $t \mapsto \exp(tA)y_0$ est l'unique solution du problème de Cauchy $y' = Ay$ avec la donnée initiale $y(0) = y_0$.

Remarque. On aurait aussi pu définir l'exponentielle de A à partir de la proposition précédente.

Corollaire. Si A et B sont deux matrices qui commutent, alors $\exp(A+B) = \exp(A)\exp(B)$. En particulier, $\exp(A)$ est inversible, d'inverse $\exp(-A)$.

1.2 Injectivité et surjectivité

Remarque. L'exponentielle de matrice n'est pas injective, ni sur $\mathbb{R}$ ni sur $\mathbb{C}$.

Proposition. Si $A = D+N$ est la décomposition de Dunford de A, alors $\exp(A) = \exp(D)\exp(N) = \exp(D) + \exp(D)(\exp(N) - I_n)$, *i.e.* on connaît la décomposition de Dunford de $\exp(A)$.

Définition. Soit M une matrice nilpotente. On définit

$$\ln(I_n - M) = \sum_{k=1}^{n-1} \frac{M^k}{k} \in K[M].$$

Proposition. Soit M une matrice nilpotente. Alors $\exp(\ln(I_n - M)) = I_n - M$.

Théorème. L'exponentielle de matrices est surjective de $\mathcal{M}_n(\mathbb{C})$ dans $\mathrm{GL}_n(\mathbb{C})$. De plus un antécédent d'une matrice inversible M par exp est un polynôme en M.

Corollaire. L'image de exp sur $\mathcal{M}_n(\mathbb{R})$ est constituée des carrés de matrices de $\mathrm{GL}_n(\mathbb{R})$.

Remarque. L'exponentielle de matrices n'est donc pas surjective de $\mathcal{M}_n(\mathbb{R})$ dans $\mathrm{GL}_n(\mathbb{R})$ (car en particulier elle est à valeurs dans l'ensemble des matrices qui ont un déterminant positif).

2 Exponentielle d'endomorphismes remarquables

2.1 Matrices symétriques

Proposition. L'application $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme.

Corollaire. On peut définir une racine carrée sur $\mathcal{S}_n^{++}(\mathbb{R})$ qui est un homéomorphisme.

Corollaire. La décomposition polaire $(O, S) \mapsto SO$ est un homéomorphisme de $\mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$ dans $\mathrm{GL}_n(\mathbb{R})$.

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors G possède un produit scalaire invariant si et seulement si G est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

2.2 Matrices antisymétriques, de trace nulle

Proposition. L'application $\exp : \mathcal{A}_n(\mathbb{R}) \to \mathrm{SO}_n(\mathbb{R})$ est continue et surjective.

Proposition. L'application $\exp : \ker(\mathrm{Tr}) \to \mathrm{SL}_n(\mathbb{R})$ est continue mais n'est pas surjective. Pour être plus précis, on a $\ker(\mathrm{Tr}) = \{X \in \mathcal{M}_n(\mathbb{R}), \exp(tX) \in \mathrm{SL}_n(\mathbb{R})\ \forall t \in \mathbb{R}\}$.

Remarque. Ceci peut être lié au fait que ces groupes sont des sous variétés, et que leur dimension est la dimension de l'espace vectoriel de départ.

3 Systèmes différentiels linéaires

3.1 Résolution

On s'intéresse au problème de Cauchy $y' = Ay$ pour $A \in \mathcal{M}_n(\mathbb{C})$ avec la condition initiale $y(0) = y_0$. On sait que ce problème admet une unique solution qui est $t \mapsto \exp(tA)y_0$.

Proposition. Soit $\lambda \in \mathbb{C}$ et $m \in \mathbb{N}$. Alors pour $t \in \mathbb{R}$

$$\exp(t(\lambda I_m + J_m)) = \mathrm{e}^{t\lambda} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-1}}{(m-1)!} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t & \frac{t^2}{2} \\ \vdots & 0 & \ddots & \ddots & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Remarque. Mettre une matrice sous forme de Jordan permet de calculer les solutions d'un système différentiel linéaire à coefficients constants.

3.2 Comportement qualitatif

Définition. On note $\chi_A = \prod_\lambda (X - \lambda)^{m_\lambda}$. On définit les espaces :

- *stable* de A par $E_\mathrm{S} = \ker(P_\mathrm{S}(A))$ avec $P_\mathrm{S} = \prod_{\mathrm{Re}(\lambda)<0}(X-\lambda)^{m_\lambda}$
- *instable* de A par $E_\mathrm{I} = \ker(P_\mathrm{I}(A))$ avec $P_\mathrm{I} = \prod_{\mathrm{Re}(\lambda)>0}(X-\lambda)^{m_\lambda}$
- *neutre* de A par $E_\mathrm{N} = \ker(P_\mathrm{N}(A))$ avec $P_\mathrm{N} = \prod_{\mathrm{Re}(\lambda)=0}(X-\lambda)^{m_\lambda}$

Théorème. On a la somme directe $\mathbb{C}^n = E_\mathrm{S} \oplus E_\mathrm{I} \oplus E_\mathrm{N}$, avec

$$E_\mathrm{S} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to\infty} \exp(tA)Y_0 = 0\}, E_\mathrm{I} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to-\infty} \exp(tA)Y_0 = 0\}$$

et

$$E_\mathrm{N} = \{Y_0 \in \mathbb{C}^n, \exists C > 0, \exists p \in \mathbb{N}, \forall t \in \mathbb{R}, \|\exp(tA)Y_0\| \leq C(1+|t|)^p\}.$$

Proposition. (Stabilité des systèmes linéaires homogènes)
On considère le système différentiel $y' = Ay$, pour $A \in \mathcal{M}_n(\mathbb{R})$. Le point d'équilibre 0 est asymptotiquement stable si et seulement si les valeurs propres complexes de A ont toutes une partie réelle strictement négative, et il est stable si et seulement si toutes les valeurs propres ont une partie réelle négative, et que les valeurs propres de parties réelles nulles sont non défectives.

Proposition. (Linéarisation en première approche)
On considère le système différentiel $y' = f(y)$ sur $\mathbb{R}^n$, avec $f \in \mathcal{C}^1$. Soit y_0 un point d'équilibre de ce système et $A = \mathrm{Jac}(f)(y_0)$. Si toutes les valeurs propres de A ont une partie réelle strictement négative, alors y_0 est asymptotiquement stable.

Théorème. (Hartman-Grobman)
Avec les mêmes notations, on suppose que A n'a pas de valeur propre de partie réelle nulle, alors il existe deux voisinages de y_0 et 0 respectivement, et $h : U \to V$ un homéomorphisme qui envoie les trajectoires de $y' = f(y)$ sur celles de $x' = Ax$ en préservant le sens du temps.

Remarque. Si on est capable de classifier l'équivalence topologique des systèmes linéaires, on est capable, grâce à ce théorème, de classifier l'équivalence topologique de flots plus ou moins quelconques.

Chapitre 24

156. Endomorphismes trigonalisables. Endomorphismes nilpotents.

Sommaire

1 Endomorphismes nilpotents 171
1.1 Définitions, caractérisations 171
1.2 Noyaux itérés 172
1.3 Réduction des nilpotents 172
2 Endomorphismes trigonalisables 172
2.1 Définition, caractérisations 172
2.2 Cotrigonalisation 173
2.3 Décomposition de Dunford, Jordan 173

Présentation du plan

Soit E un K-espace vectoriel de dimension finie n.

1 Endomorphismes nilpotents

1.1 Définitions, caractérisations

Définition. Un endomorphisme est dit *nilpotent* d'indice r si $u^r = 0$ sans que u^{r-1} soit nul.

Proposition. Soit $u \in \mathcal{L}(E)$, u est nilpotent d'indice r si et seulement si $\Pi_u = X^r$, et nilpotent tout court si et seulement si $\chi_u = X^n$.

Proposition. Si K est de caractéristique nulle, alors $u \in \mathcal{L}(E)$ est nilpotent si et seulement si pour tout $k \in \mathbb{N}^*$, $\mathrm{Tr}(u^k) = 0$.

1.2 Noyaux itérés

Proposition. Soit $u \in \mathcal{L}(E)$ nilpotent d'indice r. Alors la suite $\dim(\ker(u^i))_i$ est strictement croissante jusqu'au rang r à partir duquel elle stationne.

Lemme. (de Fitting)
Pour u nilpotent, il existe un n_0 tel que $E = \ker(u^{n_0}) \oplus \mathrm{Im}(u^{n_0})$.

Application. Le nombre de matrices nilpotentes sur $\mathbb{F}_q$ et de taille d est $q^{d(d-1)}$.

1.3 Réduction des nilpotents

Théorème. (réduction des endomorphismes nilpotents)
Soit $u \in \mathcal{L}(E)$ un endomorphisme nilpotent. Il existe une unique suite d'entiers $n_1 \leq \ldots \leq n_k$ et une base de E telles que dans cette base, la matrice de u soit de la forme

$$\begin{pmatrix} J_{n_1} & & (0) \\ & \ddots & \\ (0) & & J_{n_k} \end{pmatrix}$$

où pour $m \in \mathbb{N}^*$, J_m est la matrice de $\mathcal{M}_m(K)$ définie par

$$J_m = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{pmatrix}.$$

Remarque. Ce résultat permet de mettre en bijections les orbites de nilpotents pour l'action par conjugaison et les partitions de $\{1, \ldots, n\}$.

2 Endomorphismes trigonalisables

2.1 Définition, caractérisations

Définition. Soit $u \in \mathcal{L}(E)$, on dit que u est *trigonalisable* s'il existe une base $(e_1, \ldots, e_n)$ de E tel que pour tout k, $u(e_k) \in \mathrm{Vect}(e_1, \ldots, e_n)$, *i.e.* une base dans laquelle la matrice de u est triangulaire supérieure.

Remarque. Dans une forme matricielle triangulaire supérieure, les valeurs sur la diagonale sont forcément les valeurs propres de l'endomorphisme.

Théorème. Un endomorphisme est trigonalisable si et seulement si son polynôme minimal est scindé, si et seulement si son polynôme caractéristique est scindé, si et seulement s'il possède un polynôme annulateur scindé.

Corollaire. Sur un corps algébriquement clos, tout endomorphisme est trigonalisable.

Remarque. Si u est trigonalisable, son déterminant est égal au produit de ses valeurs propres, et sa trace à la somme de ses valeurs propres.

Théorème. (formule de Gelfand)
Soit $\|\cdot\|$ une norme quelconque sur $\mathcal{M}_n(\mathbb{C})$, alors si $\rho(A)$ est le plus grand module des valeurs propres de A,

$$\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}.$$

2.2 Cotrigonalisation

Théorème. Soit (u_i) une famille d'endomorphismes trigonalisables qui commutent deux à deux. Alors ils sont trigonalisables dans une même base.

Remarque. La réciproque de ce résultat est fausse.

Proposition. Soit G un sous-groupe fini dont tous les éléments ont pour spectre $\{1\}$. Alors G est cotrigonalisable.

2.3 Décomposition de Dunford, Jordan

Théorème. (Décomposition de Jordan-Chevalley-Dunford)
Soit u un endomorphisme de $\mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Alors il existe un unique couple (d, n) d'endomorphismes de $\mathcal{L}(E)$, avec d diagonalisable et n nilpotent, $dn = nd$. De plus d et n sont des polynômes en u.

Application. Si $A = D + N$ est la décomposition de Dunford de $A \in \mathcal{M}_n(\mathbb{C})$, alors $\exp(A) = \exp(D)\exp(N) = \exp(D) + \exp(D)(\exp(N) - I_n)$, *i.e.* on connaît la décomposition de Dunford de $\exp(A)$. Ceci peut notamment permettre de montrer que l'exponentielle matricielle est surjective de $\mathcal{M}_n(\mathbb{C})$ dans $\mathrm{GL}_n(\mathbb{C})$.

On s'intéresse au problème de Cauchy $y' = Ay$ pour $A \in \mathcal{M}_n(\mathbb{C})$ avec la condition initiale $y(0) = y_0$. On sait que ce problème admet une unique solution qui est $t \mapsto \exp(tA)y_0$. On s'intéresse au comportement de la solution en grand temps.

Définition. On note $\chi_A = \prod_\lambda (X - \lambda)^{m_\lambda}$. On définit les espaces :

- *stable* de A par $E_\mathrm{S} = \ker(P_\mathrm{S}(A))$ avec $P_\mathrm{S} = \prod_{\mathrm{Re}(\lambda)<0}(X - \lambda)^{m_\lambda}$
- *instable* de A par $E_\mathrm{S} = \ker(P_\mathrm{I}(A))$ avec $P_\mathrm{I} = \prod_{\mathrm{Re}(\lambda)>0}(X - \lambda)^{m_\lambda}$
- *neutre* de A par $E_\mathrm{S} = \ker(P_\mathrm{N}(A))$ avec $P_\mathrm{N} = \prod_{\mathrm{Re}(\lambda)=0}(X - \lambda)^{m_\lambda}$

Théorème. On a la somme directe $\mathbb{C}^n = E_\mathrm{S} \oplus E_\mathrm{I} \oplus E_\mathrm{N}$, avec

$$E_\mathrm{S} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to\infty} \exp(tA)Y_0 = 0\}, E_\mathrm{I} = \{Y_0 \in \mathbb{C}^n, \lim_{t\to-\infty} \exp(tA)Y_0 = 0\}$$

et

$$E_\mathrm{N} = \{Y_0 \in \mathbb{C}^n, \exists C > 0, \exists p \in \mathbb{N}, \forall t \in \mathbb{R}, \|\exp(tA)Y_0\| \leq C(1 + |t|)^p\}.$$

Remarque. Cela se généralise à n'importe quel endomorphisme, mais d est alors semi-simple. La preuve n'est alors pas la même puisque l'on ne peut pas écrire notre décomposition du début.

Exemple. Pour calculer une décomposition de Dunford, il faut calculer une relation de Bezout pour avoir les projecteurs sur les sous-espaces caractéristiques, on bien directement trouver une base de ces sous-espaces pour avoir la forme matricielle de l'endomorphisme que l'on peut manipuler simplement.

Notre décomposition fait apparaître un endomorphisme nilpotent, que l'on sait réduire d'après la première partie. On obtient alors le théorème de réduction de Jordan.

Théorème. (de réduction de Jordan)
Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. Il existe une base de E dans laquelle la matrice de u est de la forme

$$\begin{pmatrix} \lambda_1 I_{n_1} + J_{n_1} & & (0) \\ & \ddots & \\ (0) & & \lambda_k I_{n_k} + J_{n_k}. \end{pmatrix}$$

Application. Toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable à sa transposée. Pour un corps quelconque, il faudrait passer par une clôture algébrique et utiliser que deux matrices à valeurs dans un corps et semblables dans une extension sont semblables sur le corps de base.

Remarque. La décomposition de Jordan est un moyen pratique de calculer des exponentielles matricielles, et donc de déterminer les solutions d'une équation différentielle linéaire à coefficients constants.

Chapitre 25

157. Matrices symétriques réelles, matrices hermitiennes.

Sommaire

1 Endomorphismes auto-adjoints 175
1.1 Adjoint d'un endomorphisme 175
1.2 Le théorème spectral 176
1.3 Forme quadratique associée 177
2 Cas des endomorphismes auto-adjoints positifs, définis positifs . . . 178
2.1 Décomposition polaire 178
2.2 Valeurs singulières 178
2.3 Un exemple d'intérêt, la matrice de covariance 179

Présentation du plan

Dans tout le plan, E désigne un espace euclidien ou hermitien (on précisera quand c'est nécessaire), et son produit scalaire est noté $\langle \cdot, \cdot \rangle$.

1 Endomorphismes auto-adjoints

1.1 Adjoint d'un endomorphisme

Définition. Soit $u \in \mathcal{L}(E)$. Il existe un unique $u^* \in \mathcal{L}(E)$ tel que pour tout $(x, y) \in E^2$:

$$\langle u(x), y \rangle = \langle x, u^*(y) \rangle.$$

On l'appelle *adjoint* de u.

Remarque. En version matricielle, l'adjoint d'une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est alors $A^* := \overline{A}^\top$. Dans le cas réel, c'est seulement la transposée. En particulier, pour tout u ou tout A, on a $(u^*)^* = u$ et $(A^*)^* = A$.

Remarque. L'application $u \mapsto u^*$ est donc une symétrie de $\mathcal{L}(E)$ dans lui-même.

Proposition. Un sous-espace est stable par u si et seulement si son orthogonal est stable par u^*.

Définition. Un endomorphisme est dit *auto-adjoint* s'il est égal à son adjoint. Dans le cas réel, on parle d'endomorphisme *symétrique* et dans le cas complexe d'endomorphisme *hermitien*. Matriciellement, on note $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques et $\mathcal{H}_n(\mathbb{C})$ l'ensemble des matrices hermitiennes.

Exemple. La matrice $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ est symétrique réelle. La matrice $\begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$ est symétrique complexe mais pas hermitienne. Il faudrait rajouter un signe moins devant un des i.

Définition. Un endomorphisme réel (resp. complexe) est dit *orthogonal* (resp *unitaire*) si son adjoint est égal à son inverse. On note alors respectivement $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{U}_n(\mathbb{C})$ l'ensemble des ces matrices.

Exemple. En deux dimensions, une rotation $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est orthogonale. Une symétrie orthogonale est un endomorphisme orthogonal.

Proposition. L'application $(U, H) \mapsto U^*HU$ définit une action du groupe des matrices unitaire sur l'ensemble des matrices hermitiennes. Idem pour les matrices orthogonales et les matrices symétriques.

Définition. Un endomorphisme u hermitien ou symétrique est dit *positif* (resp. *défini positif*) si et seulement si pour tout $x \in E$, $\langle u(x), x\rangle$ est positif (resp. strictement positif). On note $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrice symétriques positives, et $\mathcal{S}_n^{++}(\mathbb{R})$ l'ensemble des matrice symétriques définies positives, idem pour $\mathbb{C}$ avec un $\mathcal{H}$.

1.2 Le théorème spectral

Théorème. Un endomorphisme symétrique (resp. hermitien) se diagonalise en base orthonormée (resp. unitaire), et a des valeurs propres réelles.

Remarque. La positivité (resp. stricte) des valeurs propres de l'endomorphisme est équivalente à la (resp. stricte) positivé de l'endomorphisme en question.

Proposition. Soit $S \in \mathcal{S}_p^+(\mathbb{R})$, alors

$$\max_{\substack{W \in \mathcal{M}_{k,p}(\mathbb{R}) \\ W^\top W = I_k}} \mathrm{Tr}(WSW^\top) = \sum_{i=1}^{k} \lambda_k,$$

avec égalité quand les lignes de W sont les k premiers vecteurs de la base orthonormée qui diagonalise S (ordonnés par ordre décroissant des valeurs propres).

Théorème. (minmax de Courant-Fischer)
Soit A une matrice hermitienne de taille n, dont les valeurs propres sont $\lambda_1 \geq \cdots \geq \lambda_n$. Si on note G_k l'ensemble des sous-espaces vectoriels de $\mathbb{C}^n$ de dimension k, alors

$$\lambda_k = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x\rangle = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x\rangle.$$

1.3 Forme quadratique associée

On se restreint ici au cas euclidien.

Proposition. Soit q une forme quadratique sur E. Il existe un unique endomorphisme de E symétrique tel que $q(x) = \langle u_q(x), x\rangle$ pour tout $x \in E$. La positivité éventuellement stricte de q équivaut alors à celle de u_q.

Proposition. (orthogonalisation simultanée, inertie de Sylvester)
Si q est une forme quadratique définie positive sur E et q' une forme quadratique quelconque alors il existe une base orthonormée pour q et orthogonale pour q'. Matriciellement, cela signifie que si M, N sont deux matrices symétriques avec M définie positive, il existe P inversible telle que $P^\top MP = I_n$ et $P^\top NP = D$ avec D diagonale. De plus, D peut être choisie avec pour coefficients diagonaux des 1, -1, 0.

Remarque. La hessienne fournit un exemple d'intérêt de forme quadratique réelle, puisqu'elle permet notamment d'étudier les variations d'une fonction au voisinage d'un point d'équilibre.

Proposition. (Changement de base $\mathcal{C}^1$)
Soit A_0 une matrice symétrique réelle inversible. Alors il existe un voisinage V de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ telle que

$$A = \rho(A)^\top A_0 \rho(A) \qquad \forall A \in V.$$

Théorème. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

2 Cas des endomorphismes auto-adjoints positifs, définis positifs

On se restreint ici au cadre euclidien.

2.1 Décomposition polaire

Proposition. L'application $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme.

Corollaire. On peut définir une racine carrée sur $\mathcal{S}_n^{++}(\mathbb{R})$ qui est un homéomorphisme.

Corollaire. La décomposition polaire $(O, S) \mapsto SO$ est un homéomorphisme de $\mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$ dans $\mathrm{GL}_n(\mathbb{R})$.

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors G possède un produit scalaire invariant si et seulement si G est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Théorème. (point fixe de Kakutani)
Soit G un sous-groupe compact de $\mathrm{GL}(E)$ et K une partie non vide convexe compacte de E stable par G. Alors il existe dans K un point fixe sous G.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

2.2 Valeurs singulières

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les racines valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \mathrm{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)$.

Corollaire. On peut définir la *pseudo-inverse* de A, notée $A^\dagger$, comme $U\Sigma^\dagger V^\top$ avec $\Sigma^\dagger \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A)^{-1}, \ldots, \sigma_r(A)^{-1}, 0, \ldots, 0)$. Pour $b \in \mathbb{R}^p$, le vecteur $A^\dagger b$ est l'unique solution de $Ax = b$ au sens des moindres carrés de norme minimale.

Proposition. (Eckart-Young)
La matrice $U\Sigma_k V^\top$, diagonale dont la diagonale est constituée des k premières valeurs singulières de A, est la meilleure approximation de A de rang k pour la norme spectrale (subordonnée à la norme deux).

Corollaire. (minmax de Courant-Fischer pour les valeurs singulières)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $p \leq n$, alors si G_k est l'ensemble des sous-espaces de $\mathbb{R}^p$ de dimension k :

$$\sigma_k(A) = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Théorème. (normes de Ky Fan)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$ ($p \leq n$). On appelle *k-ième norme de Ky Fan* la quantité $N_k(A) = \sigma_1(A) + \cdots + \sigma_k(A)$. Alors pour tout $k \in \{1, \ldots, p\}$ la k-ième norme de Ky Fan est une norme.

Remarque. La quantité $N_k(A)/N_p(A)$ permet de mesurer qu'ont les lignes de A à se compresser dans un sous-espace de dimension k. Plus précisément, la norme k de Ky Fan maximise la somme des projections des images par A d'une famille orthonormée de dimension k de $\mathbb{R}^p$ sur une famille orthonormée de dimension k de $\mathbb{R}^n$.

2.3 Un exemple d'intérêt, la matrice de covariance

Proposition. Si $X = (X_1, \ldots, X_p)$ est un vecteur aléatoire réel, on définit $\mathrm{Cov}(X) := (\mathrm{Cov}(X_i, X_j)) \in \mathcal{M}_n(\mathbb{R})$ sa matrice de covariance. C'est une matrice symétrique positive, son déterminant est quelquefois appelé *variance généralisée* de X.

Définition. Si $(X^{(1)}, \ldots, X^{(n)})$ est un échantillon i.i.d de taille n de X, on définit la *matrice de covariance empirique* associée par $K_n := \frac{1}{n}\mathbf{X}^\top\mathbf{X}$ si $\mathbf{X} = (X_j^{(i)}) \in \mathcal{M}_{n,p}(\mathbb{R})$. La variance expliquée par k lignes $w_1, \ldots, w_k$ est définie par $\mathrm{Tr}(WK_nW^\top)$.

Proposition. (analyse en composantes principales)
La projection sur un sous-espace de $\mathbb{R}^p$ de dimension k qui maximise la variance expliquée est donnée par les k premiers axes propres de la matrice de covariance empirique.

Définition. Un vecteur X à valeurs dans $\mathbb{R}^p$ est dit gaussien si pour tout $x \in \mathbb{R}^d$, $\langle x, X \rangle$ est une variable gaussienne.

Proposition. Un vecteur gaussien est caractérisé par sa moyenne et sa matrice de covariance, on note alors $X \sim \mathcal{N}(\mu, \Sigma)$.

Proposition. Si P est une matrice quelconque, et $X \sim \mathcal{N}(\mu, \Sigma)$, alors $PX \sim \mathcal{N}(P\mu, P\Sigma P^\top)$.

Proposition. Soit $X \in \mathbb{R}^d$ un vecteur aléatoire quelconque centré et réduit dont la loi est invariante par multiplication par une matrice orthogonale. Alors $X \sim \mathcal{N}(0, I_d)$.

Proposition. (Formule d'Isserlis)
Soit $(X_1, \ldots, X_n)$ un vecteur gaussien. Alors

$$\mathbb{E}(X_1 \cdots X_n) = \sum_{p \in P_n^2} \prod_{\{i,j\} \in p} \mathrm{Cov}(X_i, X_j).$$

Chapitre 26

158. Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).

Sommaire

1 Endomorphismes normaux 182
1.1 Adjoint d'un endomorphisme 182
1.2 Réduction des endomorphismes normaux 182
2 Endomorphismes symétriques 183
2.1 Le théorème spectral 183
2.2 Valeurs singulières et moindres carrés 183
3 Endomorphismes orthogonaux 184
3.1 Générateurs 184
3.2 Cas de la dimension trois 185
4 Décomposition polaire 186
4.1 Résultat et premières applications 186
4.2 Sous groupes de matrices compacts 186

Présentation du plan

Dans tout le plan, E désigne un espace euclidien de dimension n.

1 Endomorphismes normaux

1.1 Adjoint d'un endomorphisme

Définition. Soit $u \in \mathcal{L}(E)$. Il existe un unique $u^* \in \mathcal{L}(E)$ tel que pour tout $(x, y) \in E^2$:

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle.$$

On l'appelle *adjoint* de u.

Remarque. En version matricielle, l'adjoint d'une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est alors $A^* = \overline{A}^\top$. En particulier, pour tout u ou tout A, on a $(u^*)^* = u$ et $(A^*)^* = A$.

Remarque. L'application $u \mapsto u^*$ est donc une symétrie de $\mathcal{L}(E)$ dans lui-même.

Proposition. Un sous-espace est stable par u si et seulement si son orthogonal est stable par u^*.

Définition. Un endomorphisme est dit *auto-adjoint* s'il est égal à son adjoint. Dans le cas réel, on parle d'endomorphisme *symétrique*. Matriciellement, on note $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques. On dit qu'il est antisymétrique si $u^* = -u$.

Exemple. La matrice $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ est symétrique réelle.

Définition. Un endomorphisme de E est dit *orthogonal* si son adjoint est égal à son inverse. On note alors $\mathrm{O}(E)$ de ces endomorphismes et $\mathrm{O}_n(\mathbb{R})$ et l'ensemble des ces matrices.

Exemple. En deux dimensions, une rotation $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est orthogonale. Une symétrie orthogonale est un endomorphisme orthogonal.

Proposition. L'application $(U, H) \mapsto U^*HU$ définit une action du groupe des matrices orthogonales sur l'ensemble des matrices symétriques.

Définition. Un endomorphisme u hermitien ou symétrique est dit *positif* (resp. *défini positif*) si et seulement si pour tout $x \in E$, $\langle u(x), x\rangle$ est positif (resp. strictement positif). On note $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrice symétriques positives, et $\mathcal{S}_n^{++}(\mathbb{R})$ l'ensemble des matrice symétriques définies positives.

1.2 Réduction des endomorphismes normaux

Définition. Un endomorphisme $u \in \mathcal{L}(E)$ est dit *normal* lorsqu'il commute à son adjoint.

Proposition. Soit $u \in \mathcal{L}(E)$. Si de plus u est normal et F est un sous-espace propre de u, alors son orthogonal est stable par u.

Théorème. (réduction des normaux, cas euclidien)
Soit E euclidien et u normal, il existe une base orthonormée de E telle que la matrice de u dans cette base est diagonale par blocs, avec un bloc diagonal et des blocs de taille 2×2 de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

2 Endomorphismes symétriques

2.1 Le théorème spectral

Théorème. Une matrice symétrique réelle se diagonalise en base orthonormée.

Remarque. La positivité (resp. stricte) des valeurs propres de l'endomorphisme est équivalente à la (resp. stricte) positivé de l'endomorphisme en question.

Proposition. Soit $S \in \mathcal{S}_p^+(\mathbb{R})$, alors

$$\max_{\substack{W \in \mathcal{M}_{k,p}(\mathbb{R}) \\ W^\top W = I_k}} \mathrm{Tr}(WSW^\top) = \sum_{i=1}^{k} \lambda_k,$$

avec égalité quand les lignes de W sont les k premiers vecteurs de la base orthonormée qui diagonalise S (ordonnés par ordre décroissant des valeurs propres).

Théorème. (minmax de Courant-Fischer)
Soit A une matrice hermitienne de taille n, dont les valeurs propres sont $\lambda_1 \geq \cdots \geq \lambda_n$. Si on note G_k l'ensemble des sous-espaces vectoriels de $\mathbb{C}^n$ de dimension k, alors

$$\lambda_k = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle.$$

2.2 Valeurs singulières et moindres carrés

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les racines valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \mathrm{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)$.

Corollaire. On peut définir la *pseudo-inverse* de A, notée $A^\dagger$, comme $U\Sigma^\dagger V^\top$ avec $\Sigma^\dagger \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A)^{-1}, \ldots, \sigma_r(A)^{-1}, 0, \ldots, 0)$. Pour $b \in \mathbb{R}^p$, le vecteur $A^\dagger b$ est l'unique solution de $Ax = b$ au sens des moindres carrés de norme minimale.

Corollaire. (minmax de Courant-Fischer pour les valeurs singulières)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $p \leq n$, alors si G_k est l'ensemble des sous-espaces de $\mathbb{R}^p$ de dimension k :

$$\sigma_k(A) = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Théorème. (normes de Ky Fan)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$ ($p \leq n$). On appelle *k-ième norme de Ky Fan* la quantité $N_k(A) = \sigma_1(A) + \cdots + \sigma_k(A)$. Alors pour tout $k \in \{1, \ldots, p\}$ la k-ième norme de Ky Fan est une norme.

Remarque. La quantité $N_k(A)/N_p(A)$ permet de mesurer qu'ont les lignes de A à se compresser dans un sous-espace de dimension k. Plus précisément, la norme k de Ky Fan maximise la somme des projections des images par A d'une famille orthonormée de dimension k de $\mathbb{R}^p$ sur une famille orthonormée de dimension k de $\mathbb{R}^n$.

3 Endomorphismes orthogonaux

3.1 Générateurs

Définition. On définit le *groupe spécial orthogonal* comme l'ensemble des matrices orthogonales de déterminant égal à un, on le note $\mathrm{SO}(E)$.

Théorème. (Réduction des isométries)
Soit $u \in \mathrm{O}(E)$. Il existe une base orthonormée de E dans laquelle la matrice de u est diagonale par blocs, avec un bloc égal à I_r, un autre bloc égal à $-I_s$ et q blocs 2×2 qui sont des matrices de rotation :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

avec $n = r + s + 2q$. En particulier, $u \in \mathrm{SO}(E)$ si et seulement si s est pair.

Définition. Une *symétrie orthogonale* $u \in \mathrm{O}(E)$ est une isométrie involutive. On a alors la somme directe orthogonale :

$$E = \ker(u - \mathrm{id}) \oplus^\perp \ker(u + \mathrm{id}) =: E_1 \oplus^\perp E_{-1}.$$

Si $\dim(E_{-1}) = 1$, on parle de réflexion, si $\dim(E_{-1}) = 2$, de renversement.

Théorème. Le groupe orthogonal est engendré par les réflexions. Plus précisément, toute isométrie vectorielle s'écrit comme produit d'au plus n réflexions.

Le groupe spécial orthogonal est engendré par les renversements, et tout élément de $\mathrm{SO}(E)$ s'écrit comme produit d'au plus n renversements.

Corollaire. Toute isométrie affine de $\mathcal{E}$ est le produit d'au plus $(n+1)$ réflexions orthogonales (une réflexion orthogonale affine est une symétrie orthogonale par rapport à un hyperplan affine).

Proposition. Le centre de $\mathrm{O}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ et le centre de $\mathrm{SO}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ si n est paire et est trivial sinon.

Proposition. Deux symétries orthogonales s et s' sont conjuguées dans $\mathrm{O}(E)$ si et seulement si elles ont même nature, *i.e.* si et seulement si $\dim(E_1(s)) = \dim(E_1(s'))$.

Corollaire. Le groupe dérivé de $\mathrm{O}(E)$ est $\mathrm{SO}(E)$, et le groupe dérivé de $\mathrm{SO}(E)$ est $\mathrm{SO}(E)$.

Corollaire. L'ensemble $\mathrm{SO}(E)$ est une partie connexe par arcs de $\mathcal{L}(E)$, et $\mathrm{O}(E)$ a deux composantes connexes par arcs, qui sont $\mathrm{SO}(E)$ et $\mathrm{O}(E) \backslash \mathrm{SO}(E)$.

3.2 Cas de la dimension trois

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a, b, c, d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_n$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

4 Décomposition polaire

4.1 Résultat et premières applications

Proposition. L'application $\exp : \mathcal{S}_n(\mathbb{R}) \to \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme.

Corollaire. On peut définir une racine carrée sur $\mathcal{S}_n^{++}(\mathbb{R})$ qui est un homéomorphisme.

Corollaire. La décomposition polaire $(O, S) \mapsto SO$ est un homéomorphisme de $\mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$ dans $\mathrm{GL}_n(\mathbb{R})$.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ qui contient $\mathrm{O}_n(\mathbb{R})$ est égal à $\mathrm{O}_n(\mathbb{R})$.

4.2 Sous groupes de matrices compacts

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors G possède un produit scalaire invariant si et seulement si G est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Théorème. (point fixe de Kakutani)
Soit G un sous-groupe compact de $\mathrm{GL}(E)$ et K une partie non vide convexe compacte de E stable par G. Alors il existe dans K un point fixe sous G.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Chapitre 27

159. Formes linéaires et dualité en dimension finie. Exemples et applications.

Sommaire

1 Formes linéaires et espace dual 188
1.1 Espace dual, bases duales 188
1.2 Bidualité, base antéduale 189
2 Orthogonalité 189
2.1 Généralités 189
2.2 Hyperplans et intersections d'hyperplans 190
3 Application transposée 191
3.1 Définition et lien avec les matrices 191
3.2 Transposée et stabilité 192
4 Dualité dans les espaces euclidiens et géométrie 193
4.1 Le théorème de Riesz en dimension finie 193
4.2 Hyperplan d'appui et convexes 193
5 Dualité dans les espaces de matrices 194
5.1 Dualité canonique 194
5.2 Formulations variationnelles 194

Présentation du plan

Dans tout le plan, K est un corps.

1 Formes linéaires et espace dual

1.1 Espace dual, bases duales

Définition. On définit l'*espace dual* de E comme l'ensemble des formes linéaires de E. On le note $E^* = \mathcal{L}(E, K)$.

Remarque. C'est un espace vectoriel de dimension n par la formule $\dim(\mathcal{L}(E,F)) = \dim(E)\dim(F)$, mais cette justification n'est pas satisfaisante puisqu'on utilise de la dualité pour la montrer. On justifiera par la suite la dimension de E^*.

Définition. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. On définit sa *base duale* $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ par

$$e_k^* = \sum_{i=1}^{n} \lambda_i e_i \mapsto \lambda_k,$$

la forme linéaire e_k^* donne la coordonnée selon e_k.

Proposition. C'est l'unique base de E^* qui vérifie $e_i^*(e_j) = \delta_{i,j}$. En particulier, on a bien $\dim(E^*) = n$.

Remarque. La notation e_i^* peut être trompeuse. En effet, pour un vecteur x quelconque, la notation u^* n'a aucun sens ; elle dépend de la base de E que l'on choisit.

Définition. Soit $E = F \oplus G$ et $(f_1, \dots, f_l, g_1, \dots, g_p)$ une base adaptée à la décomposition. Alors on définit la *projection sur F parallèlement à G* par $x \mapsto \sum_{i=1}^{l} f_i^*(x) f_i \in F$. C'est un endomorphisme (idempotent) de $\mathcal{L}(E)$.

Remarque. La dualité n'étant pas canonique dans un cadre général, on doit préciser le sous-espace parallèlement auquel on projette.

Proposition. Avec les mêmes notations, l'application $\Phi_{\mathcal{B}} = e_i \mapsto e_i^*$ est un isomorphisme de $\mathcal{L}(E, E^*)$.

Définition. Pour $\varphi \in E^*$ et $x \in E$, on définit le *crochet de dualité* de φ et x noté $\langle \varphi, x \rangle := \varphi(x)$. C'est une application bilinéaire de $E^* \times E$ dans K.

Remarque. En gros, une forme linéaire, ce n'est jamais qu'un brave vecteur ligne ! Attention cependant, il faut toujours considérer une base pour écrire un produit ligne/colonne. De plus, le crochet n'est pas, en l'état, symétrique puisque écrire $\langle x, \varphi \rangle$ n'a pour l'instant aucun sens.

Définition. Soient E et F deux espaces de dimensions respectives n et m, soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (f_1, \dots, f_m)$ des bases respectives de E et F. Soit $u \in \mathcal{L}(E,F)$. La *matrice* de u dans les bases $\mathcal{B}$ et $\mathcal{C}$ est le tableau de scalaires

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) := (f_i^*(u(e_j)))_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathcal{M}_{m,n}(K).$$

Alors faire la somme de deux endomorphismes et écrire la matrice de la somme revient à faire la somme terme à terme des deux matrices. Cela fait de $\mathcal{M}_{m,n}(K)$ un K-espace vectoriel. Comme il est clairement de dimension $m \times n$, on retrouve le fait que la dimension de $\mathcal{L}(E,F)$ est égal au produit des dimensions.

Remarque. Je l'avais oublié, mais c'est comme cela qu'on déduit proprement une matrice, puisque prendre les coordonnées dans une base, c'est juste appliquer la base duale.

1.2 Bidualité, base antéduale

Théorème. L'application

$$J : x \mapsto (\varphi \mapsto \varphi(x))$$

est un isomorphisme de $\mathcal{L}(E, (E^*)^*)$.

Remarque. Un espace est toujours canoniquement isomorphe à son bidual, mais pas à son dual ; on construit un isomorphisme canonique avec le dual que pour une base de E donnée.

Proposition. Soit $\mathcal{B}$ une base de E. Alors $J(\mathcal{B}) = (\mathcal{B}^*)^*$.

Proposition. Si $\mathcal{B}_1 = (e_1, \dots, e_n)$ et $\mathcal{B}_2 = (f_1, \dots, f_n)$ sont deux bases de E, alors

$$\mathrm{P}_{\mathcal{B}_1 \to \mathcal{B}_2} = (\mathrm{P}_{\mathcal{B}_2^* \to \mathcal{B}_1^*})^\top.$$

Corollaire. L'application $\mathcal{B} \to \mathcal{B}^*$ est une bijection.

Définition. Si $\mathcal{C}$ est une base de E^*, on définit la *base antéduale* de $\mathcal{C}$ comme l'unique antécédent de $\mathcal{C}$ par l'application précédente, qui correspond donc à $J^{-1}(\mathcal{C}^*)$.

Exemple. (Deux exemples parlants sur les espaces de polynômes)
On considère K de caractéristique nulle et $E = K_n[X]$, qui est de dimension $n+1$. On considère, pour $a \in K$, la base $((X-a)^k)_k$ pour $k \in \{0, \dots, n\}$. Sa base duale est $(P \mapsto P^{(k)}(a)/k!)_k$ d'après la formule de Taylor.

Réciproquement, la base antéduale de $(P \mapsto P(a_i))_i$ pour $(a_0, \dots, a_n)$ n points de K deux à deux distincts est la base des polynômes interpolateurs de Lagrange associée.

2 Orthogonalité

2.1 Généralités

Définition. Soit A une partie de E et B une partie de son dual. On définit

$$A^\perp := \{\varphi \in E^*, \varphi(a) = 0 \, \forall a \in A\} \text{ et } B^\circ = \{x \in E, b(x) = 0 \, \forall b \in B\}.$$

Dans les deux cas, on le lit "orthogonal de".

Remarque. On a bien ici deux définitions puisque E n'est pas canoniquement isomorphe à son dual. Cependant, on a quand même un lien entre les deux notions par bidualité : $B^\perp = J(B^\circ)$.

Proposition. Soient A_1, A_2, A des parties de E et B_1, B_2, B des parties de E^*, alors

- $A_1 \subset A_2 \Rightarrow A_2^\perp \subset A_1^\perp$
- $B_1 \subset B_2 \Rightarrow B_2^\circ \subset B_1^\circ$
- $A^\perp = \mathrm{Vect}(A)^\perp = \mathrm{Vect}(A^\perp)$
- $B^\circ = \mathrm{Vect}(B)^\circ = \mathrm{Vect}(B^\circ)$

Proposition. Soit F un sous-espace vectoriel de E, alors $\dim(F^\perp) = \dim(E) - \dim(F)$. Idem pour un sous-espace de E^* par bidualité.

Corollaire. Si F est un sous-espace vectoriel de E, alors $(F^\perp)^\circ = F$.

Proposition. Soient F et G deux sous-espaces vectoriels de E, V et W deux sous-espaces vectoriels de E^* alors

— $(F+G)^\perp = F^\perp \cap G^\perp$

— $(F \cap G)^\perp = F^\perp + G^\perp$

— $(V+W)^\circ = V^\circ \cap W^\circ$

— $(V \cap W)^\circ = V^\circ + W^\circ$

2.2 Hyperplans et intersections d'hyperplans

Théorème. Un sous-espace H de E est un hyperplan si et seulement si il existe $\varphi \in E^*\{0\}$ tel que

$$H = \{\varphi\}^\circ = \ker(\varphi).$$

Proposition. Un sous-espace est de dimension k si et seulement s'il s'écrit comme l'orthogonal de $n-k$ formes linéaires libres, si et seulement si c'est l'intersection de $n-k$ hyperplans deux à deux distincts.

Proposition. Par ailleurs si $F = \bigcap_{i=1}^q \ker(\varphi_i)$ une écriture de F comme intersection d'hyperplans, pour $q \geq n-k$, alors si $\varphi_1, \ldots, \varphi_r$ une base des $(\varphi_i)_{1\leq i\leq q}$, alors $\operatorname{codim}(F) = \operatorname{codim}((\varphi_i)^\circ_{1\leq i\leq r}) = \dim(\operatorname{Vect}(\varphi_i, 1 \leq i \leq r)) = r$. On a une inclusion et égalité des dimensions, donc

$$F = \bigcap_{i=1}^{n-k} \ker(\varphi_i).$$

Application. (Commutant d'une matrice trigonalisable)
Soit $A \in \mathcal{M}_n(K)$. On cherche la dimension de Z_A le commutant de A. Quitte à changer de base, on peut supposer A triangulaire supérieure. On cherche d'abord les éléments triangulaires supérieurs du commutant, notés Z_A^T. On a alors un système à $\frac{n(n+1)}{2}$ inconnues et autant d'équations à première vue. Cependant, les équations diagonales donnent $a_{i,i}x_i = x_i a_{i,i}$, ce qui est muet. On a donc en réalité au plus $\frac{n(n+1)}{2} - n$ équations indépendantes, donc

$$\dim(\mathrm{Z}_A) \geq \dim(\mathrm{Z}_A^T) \geq n.$$

On peut revenir au cas général par extension de corps.

Proposition. Soient (φ_i) r formes linéaires de E. Alors pour $\psi \in E^*$, $\bigcap_{i=1}^r \ker(\varphi) \subset \ker(\varphi)$ si et seulement si $\psi \in \operatorname{Vect}(\varphi_i)$.

3 Application transposée

3.1 Définition et lien avec les matrices

Définition. Soit $u \in \mathcal{L}(E,F)$. On définit sa transposée :

$$^tu : \varphi \mapsto \varphi \circ u,$$

qui est un élément de $\mathcal{L}(F^*, E^*)$.

On a alors bien ce que l'on veut.

Proposition. Soient $\mathcal{B} = (e_1, \dots, e_n), \mathcal{C} = (f_1, \dots, f_p)$ des bases respectives de E et F, soit $u \in \mathcal{L}(E,F)$.

$$\mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}(^tu) = (\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u))^\top.$$

En particulier, un endomorphisme et sa transposée ont même déterminant et même polynôme caractéristique.

Remarque. On retrouve donc une propriété de l'adjoint euclidien, puisque

$$\langle u(x), \varphi\rangle = \langle x, {}^tu(\varphi)\rangle.$$

En particulier, on retrouve aussi que transposer deux fois ne change rien, au détail près des espaces de départ et d'arrivée :

$$^t(^tu) = J \circ u \circ J^{-1},$$

ce qu'on prouve simplement avec les formes matricielles.

Corollaire. Pour u, v des endomorphismes définis dans des espaces qui rendent $u \circ v$ défini, on a

$$^t(u \circ v) = {}^tv \circ {}^tu.$$

Proposition. (lien entre image et noyau)
Soit $u \in \mathcal{L}(E,F)$ alors

- $\ker(^tu) = \mathrm{Im}(u)^\perp$
- $\mathrm{Im}(^tu) = \ker(u)^\perp$

Corollaire. Soit $u \in \mathcal{L}(E,F)$, alors le rang de u est égal au rang de son application transposée.

Remarque. On peut aussi retrouver ce résultat avec la forme matricielle, mais encore une fois logiquement on utilise ce résultat pour montrer que le rang d'une matrice c'est le rang de sa transposée.

Proposition. Soit $u \in \mathcal{L}(E,F)$, alors u est injectif (resp. surjectif) si et seulement si sa transposée est surjective (resp. injective).

Application. Soit K de cardinal q et $m \geq n$. Il y a $\prod_{i=0}^{n-1}(q^m - q^i)$ familles génératrices de K^n à m éléments.

3.2 Transposée et stabilité

Proposition. Un sous-espace vectoriel F de E est stable par $u \in \mathcal{L}(E,F)$ si et seulement si son orthogonal est stable par ${}^t u$.

Remarque. Encore une fois, cela se voit matriciellement, et c'est ce que l'on faisait sans le savoir jusqu'à présent.

Mettons cela en pratique et montrons ce résultat classique avec et sans dualité pour voir la différence.

Théorème. Un endomorphisme est trigonalisable si et seulement si son polynôme caractéristique est scindé.

Proposition. Soient $(f_i)_{i\in I}$ une famille d'endomorphismes trigonalisables qui commutent. Alors ils sont cotrigonalisables.

Application. Si $u_1, \dots, u_n$ sont nilpotents et commutent ($n = \dim(E)$) alors leur produit est nul.

Proposition. Soit $u \in \mathcal{L}(E)$ et x un vecteur cyclique pour u, c'est-à-dire tel que $\pi_{u,x} = \pi_u$ (un tel x existe toujours lais on ne le montre pas ici). Soit $F = \mathrm{Vect}(x, u(x), \dots, u^{d-1}(x))$ avec $d = \deg(\pi_u)$, c'est un sous-espace stable par u. Il existe un supplémentaire de F stable par u.

4 Dualité dans les espaces euclidiens et géométrie

4.1 Le théorème de Riesz en dimension finie

Théorème. (représentation de Riesz)
L'application

$$\Phi : x \mapsto \langle x, \cdot \rangle$$

est un isomorphisme de E.

Remarque. On a maintenant une identification canonique et naturelle entre un point et un hyperplan par $a \mapsto a^\perp$. On a par ailleurs une partition

$$E = a^\perp \cup \{x, \langle a, x\rangle > 0\} \cup \{x, \langle a, x\rangle < 0\}.$$

De même, un hyperplan affine s'écrit $\{x \langle a, x\rangle = \alpha\}$ et sépare E en deux demi-espaces. Selon moi c'est ça l'essence de la dualité dans le cadre euclidien ; identifier une direction et un hyperplan et donc deux demi-espaces. Je vais essayer de le clarifier dans le paragraphe suivant.

4.2 Hyperplan d'appui et convexes

Définition. Soit $C \subset E$ convexe. On dit que $c \in C$ est un *point extrémal* de C si $C \backslash \{c\}$ est convexe, ou de manière équivalente si $\frac{x+y}{2} = c$ entraîne $x = y = c$ pour $x, y \in C$ (c ne peut pas s'écrire comme milieu strict de points de C).

Définition. On dit que H est un hyperplan d'appui $a^\perp$ à C en $c \in C$ si $c \in H$ et si C est inclus dans $\{x \in E, \langle x - c, a\rangle \geq 0\}$ ou $\{x \in E, \langle x - c, a\rangle \leq 0\}$.

Proposition. Soit C un convexe fermé non vide. Si $c \in \partial C$, il existe un hyperplan d'appui à C en c.

Proposition. Soit $a \in C$, on suppose que a est dans un hyperplan d'appui H de C. Alors a est point extrémal de C si et seulement si a est point extrémal de $H \cap C$.

Théorème. (de Krein-Milman)
Soit C un convexe compact non vide de $\mathbb{R}^d$. Alors il est l'enveloppe convexe de ses points extrémaux.

5 Dualité dans les espaces de matrices

5.1 Dualité canonique

Proposition. L'application

$$\Psi : B \mapsto (A \mapsto \mathrm{Tr}(AB))$$

est un isomorphisme de $\mathcal{M}_{p,n}(K)$ dans $\mathcal{M}_{n,p}(K)^*$.

Remarque. La preuve que l'on retrouve communément est différente ; on montre que l'application est injective en évaluant en les $E_{i,j}$ puis on conclut par égalité des dimensions, ce qui peut sembler un peu magique. Ici, quitte à évaluer en les $E_{i,j}$ on pousse le raisonnement plus loin et on montre que ça entraîne la surjectivité en tant que tel. Par ailleurs, ce résultat n'est pas surprenant car $\mathrm{Tr}(AB) = \sum_{i,j} a_{i,j} b_{j,i}$, qui est une sorte de "produit scalaire canonique".

Application. Tout hyperplan de $\mathcal{M}_n(K)$ contient une matrice inversible, et tout hyperplan de $\mathcal{M}_n(\mathbb{R})$ contient une matrice orthogonale.

5.2 Formulations variationnelles

Proposition. (Analyse en composantes principales)
Soit $S \in \mathcal{S}_p(\mathbb{R})$, on cherche à maximiser la quantité $\mathrm{Tr}(WSW^\top)$ sur l'ensemble des matrices $W \in \mathcal{M}_{k,p}(\mathbb{R})$ telles que $WW^\top = I_k$. On note $\lambda_1 \geq \ldots \geq \lambda_p$ les valeurs propres de S. La quantité est alors maximale quand les lignes de W sont k vecteurs propres associés respectivement à $\lambda_1, \ldots, \lambda_k$, et vaut alors $\lambda_1 + \cdots + \lambda_k$.

Application. Supposons que l'on ait $\mathbf{X} \in \mathcal{M}_{n,p}(\mathbb{R})$ une matrice qui contient n observations indépendantes d'un vecteur aléatoire $X \in \mathbb{R}^p$, la somme des normes des projections de nos n points sur un sous-espace de dimension k dont une base orthonormée sont les lignes de W (ou de manière équivalente, la somme des variances expliquées par les projections sur les lignes de W) correspond à $\mathrm{Tr}(W\mathbf{X}^\top\mathbf{X}W^\top)$. Ainsi, les k premiers axes propres de $\mathbf{X}^\top\mathbf{X}$ sont les k axes qui expliquent le mieux les données $\mathbf{X}$, on les appelle les k premières *composantes principales*. On les utilise en statistique descriptives pour trouver des nouvelles variables qui nous permettront d'en réduire le nombre total.

Théorème. (valeurs singulières et Courant-Fischer)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$, il existe $P \in \mathrm{O}_n(\mathbb{R})$ et $Q \in \mathrm{O}_p(\mathbb{R})$ telles que

$$A = P\mathrm{diag}(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)Q^\top,$$

avec $\sigma_1(A) \geq \ldots \geq \sigma_r(A) \geq 0$ les *valeurs singulières* de A. Si G_d est l'ensemble des sous-espaces vectoriels de $\mathbb{R}^p$ de dimension d, alors :

$$\sigma_k(A) = \min_{W \in G_{p-k+1}} \max_{\substack{x \in W \\ \|x\|=1}} \|Ax\| = \max_{W \in G_p} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Définition. Pour $1 \leq k \leq p$, on définit la k-ième *norme de Ky Fan* de $A \in \mathcal{M}_{n,p}(\mathbb{R})$ comme la somme de ses k premières valeurs singulières, on la note $N_k(A)$.

Théorème. La norme k de Ky Fan vérifie la formulation variationnelle suivante :

$$N_k(A) = \max_{\substack{U^\top U = V^\top V = I_k \\ U \in \mathcal{M}_{n,k}(\mathbb{R}) \\ V \in \mathcal{M}_{p,k}(\mathbb{R})}} \mathrm{Tr}(U^\top AV).$$

C'est donc une norme.

Remarque. Comme on l'a fait précédemment pour les matrice symétriques, on peut ici interpréter $\mathrm{Tr}(U^\top AV)$ comme la projection sur un espace de dimension k de l'image par A d'un espace de dimension k. La quantité $\frac{N_k(A)}{N_p(A)}$ indique la quantité maximale d'information que l'on peut garder en faisant cela.

Chapitre 28

161. Espaces vectoriels et espaces affines euclidiens : distances, isométries.

Sommaire

1 Généralités : applications affines, distances, isométries 198

1.1 Distances et matrices de Gram 198

1.2 Applications affines et isométries 199

2 Étude du groupe orthogonal 201

2.1 Générateurs et réduction 201

2.2 Paramétrisation du groupe spécial orthogonal en dimension 3 par les quaternions. 202

3 Classification des isométries 203

3.1 Décomposition canonique 203

3.2 Isométries du plan 203

3.3 Isométries de l'espace 203

4 Isométries préservant des parties du plan et de l'espace 203

4.1 Groupe diédral 204

4.2 Isométries du tétraèdre et du cube 204

Présentation du plan

Plan de leçon

On considère $\mathcal{E}$ un espace affine sur E préhilbertien. On note $\mathcal{A}(\mathcal{E})$ l'ensemble des applications affines de $\mathcal{E}$ et la partie linéaire d'une application affine f est notée v_f. On note Aut(E) le groupe affine de $\mathcal{E}$.

1 Généralités : applications affines, distances, isométries

1.1 Distances et matrices de Gram

On considère ici E un espace préhilbertien, et $\mathcal{E}$ un espace affine dirigé par E.

Proposition. On peut alors définir une distance sur $\mathcal{E}$ à partir de la norme euclidienne :

$$d(M,N) = \|\overrightarrow{MN}\|.$$

Définition. On appelle distance d'un point M a un sous-espace affine $\mathcal{F} \subset \mathcal{E}$ la quantité :

$$d(M,\mathcal{F}) := \inf_{P\in\mathcal{F}} d(P,M).$$

Théorème. Lorsque $\mathcal{E}$ est de dimension finie, la distance à un sous-espace affine $\mathcal{F}$ est atteinte en un unique point P. Si on note $\mathrm{p}_G^{\perp}$ la projection orthogonale sur un sous-espace vectoriel G, et si on fixe un $A \in \mathcal{F}$, P est caractérisé par :

$$\overrightarrow{AP} = \mathrm{p}_F^{\perp}(\overrightarrow{AM}),$$

et on a alors $d = \|\mathrm{p}_{F^\perp}^{\perp}(\overrightarrow{AM})\|$.

Remarque. Ce résultat se généralise en dimension infinie dans le cas où F est un sous-espace vectoriel de dimension finie.

Exemple. Dans l'espace affine $\mathbb{R}^3$, la distance d'un point $M = (x_M, y_M, z_M)$ au plan $\mathcal{P}$ d'équation $\{ax + by + cz + d = 0\}$ se calcule par la formule :

$$d(M,\mathcal{P}) = \frac{|ax_M + by_M + cz_M + d|}{\sqrt{a^2+b^2+c^2}}.$$

Dans la suite, on se place dans le cadre vectoriel, E désigne un espace préhilbertien.

Définition. On appelle *matrice de Gram* de la famille $(x_1,\dots,x_n) \in E^n$ la matrice $(\langle x_i, x_j\rangle)_{1\le i,j\le n} \in \mathcal{M}_n(\mathbb{R})$, et son déterminant le *déterminant de Gram* de la famille, noté $G(x_1,\dots,x_n)$.

Proposition. Une matrice est une matrice de Gram si et seulement si elle est symétrique positive.

Remarque. Une matrice de Gram est définie positive si et seulement si la famille en question est liée.

Exemple. Sur $E = \mathrm{L}^2(\Omega,\mathcal{T},\mathbb{P})$, muni du produit scalaire $\langle X,Y\rangle = \mathbb{E}(XY)$, la matrice de covariance d'un vecteur aléatoire $X = (X_1,\dots,X_n)$ est la matrice de Gram du vecteur $(X_1 - \mathbb{E}(X_1),\dots,X_n - \mathbb{E}(X_n))$. On appelle variance généralisée son déterminant de Gram, noté $\mathrm{Var}(X)$.

Théorème. Soit F un sous-espace vectoriel de E qui admet une base $(e_1,\dots,e_k)$. Pour $x \in E$, on a :

$$d(x,F)^2 = \frac{G(x,e_1,\dots,e_k)}{G(e_1,\dots,e_k)}.$$

Application. L'expression $\int_0^1 (1 + a_1 x + \ldots + a_n x^n)^2 \, \mathrm{d}x$ atteint son minimum $\mu = \frac{1}{n+1}$ en un unique n-uplet.

Exemple. Dans le cas d'un vecteur gaussien centré $(X, Y) \in \mathrm{L}^2$, puisque la projection orthogonale sur $\mathrm{L}^2(\Omega, \sigma(X_1, \ldots, X_n), \mathbb{P})$ correspond à l'espérance conditionnelle (qui est combinaison linéaire des (X_i) car on est dans le monde gaussien) :

$$d(Y, \mathrm{Vect}(X_1, \ldots, X_n)) = \mathbb{E}((Y - \mathbb{E}(Y|X))^2) = \mathbb{E}(\mathrm{Var}(Y|X)) = \frac{\mathrm{Var}(Y, X)}{\mathrm{Var}(X)}.$$

Théorème. (Inégalité de Hadamard)
Soit $(x_1, \ldots, x_n)$ une famille de vecteurs non nuls de E. Alors

$$G(x_1, \ldots, x_n) \leq \prod_{i=1}^{n} \|x_i\|^2,$$

avec égalité si et seulement si la famille est orthogonale. Plus généralement,

$$|\det(x_1, \ldots, x_n)| \leq \prod_{i=1}^{n} \|x_i\|,$$

avec le même cas d'égalité.

Exemple. Pour un vecteur aléatoire X de taille n, on a $\mathrm{Var}(X) \leq \mathrm{Var}(X_1) \cdots \mathrm{Var}(X_n)$, avec égalité si et seulement si ses composantes sont décorrélées.

1.2 Applications affines et isométries

On suppose ici que E est de dimension finie (euclidien).

Définition. Une application $u \in \mathcal{E}$ est une isométrie (vectorielle) si elle préserve la norme. L'ensemble des isométries vectorielles de E forme un groupe, appelé *groupe orthogonal*, noté $\mathrm{O}(E)$.

Proposition. Une application linéaire $u \in \mathcal{L}(E)$ est une isométrie vectorielle si et seulement si sa matrice dans toute base orthonormée vérifie $MM^\top = I_n$. On dit alors que M est *orthogonale* et on note $M \in \mathrm{O}_n(\mathbb{R})$.

Exemple. Dans le plan euclidien, la rotation d'angle θ est une isométrie vectorielle. Dans la base canonique, sa matrice est

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Définition. Si $u \in \mathrm{O}(E), \det(u) = \pm 1$. On appelle *groupe spécial orthogonal*, noté $\mathrm{SO}(E)$, l'ensemble des isométries vectorielles de déterminant 1, qui forme un sous-groupe de $\mathrm{O}(E)$. On note $\mathrm{SO}_n(\mathbb{R})$ l'ensemble des matrices orthogonales de déterminant égal à 1.

Définition. Une application $f : \mathcal{E} \to \mathcal{E}$ est une *isométrie* si elle préserve les distances :

$$\forall M, N \in \mathcal{E}, d(f(M), f(N)) = d(M, N).$$

On note $\mathrm{Isom}(\mathcal{E})$ l'ensemble des isométries de $\mathcal{E}$.

Exemple. La translation de vecteur $\overrightarrow{x}$, $t_{\overrightarrow{x}} : M \mapsto M + \overrightarrow{x}$ est une isométrie, et une homothétie translation de la forme $\lambda \mathrm{id} \circ t_{\overrightarrow{x}}$ est une isométrie si et seulement si $\lambda = \pm 1$.

Exemple. Dans le plan affine euclidien, une rotation d'angle θ autour d'un point C est une isométrie.

Proposition. Une isométrie de $\mathcal{E}$ est affine. De plus, une application $f \in \mathcal{A}(\mathcal{E})$ est une isométrie si et seulement si sa partie linéaire est une isométrie vectorielle.

Définition. On note $\mathrm{Isom}^+(\mathcal{E})$ le groupe des isométries affines dont la partie linéaire est dans $\mathrm{SO}(E)$, appelées *déplacements*. Si la partie linéaire a pour déterminant -1, on parle d'anti-déplacement.

Lemme. (Principe de conjugaison)
Pour $\overrightarrow{x}$ un vecteur et f une bijection affine de $\mathcal{E}$, on a la relation :

$$f \circ t_{\overrightarrow{x}} \circ f^{-1} = t_{v_f(\overrightarrow{x})}.$$

Théorème. Le groupe des automorphismes de $\mathcal{E}$ est isomorphe à un produit semi-direct

$$\mathrm{Aut}(\mathcal{E}) \simeq E \rtimes_\alpha \mathrm{GL}(E),$$

où $\alpha : (\overrightarrow{x}, u) \mapsto u(\overrightarrow{x})$, et on a le résultat analogue pour les isométries : le groupe des isométries est isomorphe au produit semi-direct :

$$\mathrm{Isom}(\mathcal{E}) \simeq E \rtimes_\alpha \mathrm{O}(E),$$

2 Étude du groupe orthogonal

Soit E un espace euclidien de dimension n.

2.1 Générateurs et réduction

Définition. Une symétrie orthogonale $u \in \mathrm{O}(E)$ est une isométrie involutive. On a alors la somme directe orthogonale :

$$E = \ker(u - \mathrm{id}) \oplus^\perp \ker(u + \mathrm{id}) =: E_1 \oplus^\perp E_{-1}.$$

Si $\dim(E_{-1}) = 1$, on parle de réflexion, si $\dim(E_{-1}) = 2$, de renversement.

Théorème. Le groupe orthogonal est engendré par les réflexions. Plus précisément, toute isométrie vectorielle s'écrit comme produit d'au plus n réflexions.

Le groupe spécial orthogonal est engendré par les renversements, et tout élément de $\mathrm{SO}(E)$ s'écrit comme produit d'au plus n renversements.

Corollaire. Toute isométrie affine de $\mathcal{E}$ est le produit d'au plus $(n+1)$ réflexions orthogonales (une réflexion orthogonale affine est une symétrie orthogonale par rapport à un hyperplan affine).

Proposition. Le centre de $\mathrm{O}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ et le centre de $\mathrm{SO}(E)$ est égal à $\{-\mathrm{id}, \mathrm{id}\}$ si n est paire et est trivial sinon.

Proposition. Deux symétries orthogonales s et s' sont conjuguées dans $\mathrm{O}(E)$ si et seulement si elles ont même nature, *i.e.* si et seulement si $\dim(E_1(s)) = \dim(E_1(s'))$.

Corollaire. Le groupe dérivé de $\mathrm{O}(E)$ est $\mathrm{SO}(E)$, et le groupe dérivé de $\mathrm{SO}(E)$ est $\mathrm{SO}(E)$.

Théorème. (Réduction des isométries)
Soit $u \in \mathrm{O}(E)$. Il existe une base orthonormée de E dans laquelle la matrice de u est diagonale par blocs, avec un bloc égal à I_r, un autre bloc égal à $-I_s$ et q blocs 2×2 qui sont des matrices de rotation :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

avec $n = r + s + 2q$. En particulier, $u \in \mathrm{SO}(E)$ si et seulement si s est pair.

Corollaire. L'ensemble $\mathrm{SO}(E)$ est une partie connexe par arcs de $\mathcal{L}(E)$, et $\mathrm{O}(E)$ a deux composantes connexes par arcs, qui sont $\mathrm{SO}(E)$ et $\mathrm{O}(E) \backslash \mathrm{SO}(E)$.

2.2 Paramétrisation du groupe spécial orthogonal en dimension 3 par les quaternions.

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a, b, c, d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices spéciales unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_k$ et $\det(M) = 1$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

3 Classification des isométries

3.1 Décomposition canonique

Lemme. Soit $u \in \mathrm{O}(E)$. Alors $\ker(u - \mathrm{id}) = \mathrm{Im}(u - \mathrm{id})^\perp$ et E est la somme directe orthogonale de ces deux sous-espaces.

Théorème. (Décomposition canonique d'une isométrie affine) Soit $f \in \mathrm{Isom}(\mathcal{E})$. Il existe un unique couple $g \in \mathrm{Isom}(\mathcal{E}), \overrightarrow{x} \in E$ tels que :

— g admet un point fixe

— $\overrightarrow{x} \in \ker(v_f - \mathrm{id})$

— $f = t_{\overrightarrow{x}} \circ g = g \circ t_{\overrightarrow{x}}$.

Exemple. Dans le cas où v_f est une symétrie vectorielle (orthogonale), on dit que f est une *symétrie glissée* d'axe $\mathcal{D}$, où $\mathcal{D}$ est la droite passant par le point fixe de g de direction $\overrightarrow{x}$.

3.2 Isométries du plan

Théorème. Soit f une isométrie du plan affine euclidien. Alors :

— f est un déplacement, auquel cas c'est une translation, une rotation ou une symétrie centrale

— f est un anti-déplacement, auquel cas c'est une symétrie axiale ou une symétrie axiale glissée.

3.3 Isométries de l'espace

Théorème. Soit f une isométrie de l'espace affine euclidien. Alors :

— f est un déplacement, auquel cas c'est une translation, une rotation par rapport à une droite, ou une rotation par rapport à une droite composée à une translation le long de cette droite, on parle de *vissage*.

— f est un anti-déplacement, auquel cas c'est une symétrie orthogonale (éventuellement glissée) par rapport à un plan, ou bien c'est le produit d'une rotation par rapport à une droite et d'une symétrie par rapport à un plan.

4 Isométries préservant des parties du plan et de l'espace

Définition. Pour une partie X de l'espace affine $\mathbb{R}^n$, on note $\mathrm{Is}(X)$ (resp. $\mathrm{Is}^+(X)$) l'ensemble des isométries affines (resp. déplacements) f de $\mathbb{R}^n$ telles que $f(X) = X$. Ces deux ensembles forment deux sous-groupes de $\mathrm{Isom}(\mathbb{R}^n)$.

Proposition. Si X est une partie convexe, alors les éléments de $\mathrm{Is}(X)$ permutent les points extrémaux de X.

4.1 Groupe diédral

On se place ici dans $\mathbb{R}^2$.

Définition. Pour $n \geq 3$, on note D_n le groupe $\mathrm{Is}(P_n)$, où P_n est le polygone régulier à n côtés centré en 0 et dont 1 est l'un des côtés. On l'appelle *groupe diédral* d'ordre $2n$.

Théorème. D_n est un groupe de cardinal $2n$, engendré par r la rotation d'angle $\frac{2\pi}{n}$ et la symétrie orthogonale s d'axe $\mathbb{R}(1,0)$, avec la relation $(rs)^2 = \mathrm{id}$. De plus, $D_n^+ := \mathrm{Is}^+(P_n) \simeq \mathbb{Z}/n\mathbb{Z}$.

4.2 Isométries du tétraèdre et du cube

On note T le tétraèdre régulier et C le cube de $\mathbb{R}^3$.

Théorème. Le groupe $\mathrm{Is}(T)$ est isomorphe à $\mathfrak{S}_4$ et $\mathrm{Is}^+(T)$ à $\mathfrak{A}_4$.

Théorème. Le groupe $\mathrm{Is}^+(C)$ est isomorphe à $\mathfrak{S}_4$ et le groupe $\mathrm{Is}(C)$ est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

Chapitre 29

162. Systèmes d'équations linéaires ; opérations élémentaires, aspects algorithmiques et conséquences théoriques.

Sommaire

1 Généralités sur les systèmes linéaires 205
1.1 Dimension d'une intersection d'hyperplans 205
1.2 Description matricielle, système de Cramer 207
2 Résolution pratique des systèmes linéaires 208
2.1 Pivot de Gauss 208
2.2 Résolution approchée, moindres carrés 208
3 Cas des anneaux principaux 209
3.1 Échelonnement d'un système 209
3.2 Forme normale de Smith 209

Présentation du plan

Dans tout le plan, K est un corps.

1 Généralités sur les systèmes linéaires

1.1 Dimension d'une intersection d'hyperplans

Proposition. Un sous-espace est de dimension k si et seulement s'il s'écrit comme l'orthogonal de $n-k$ formes linéaires libres, si et seulement si c'est l'intersection de $n-k$ hyperplans deux à deux distincts.

Remarque. En effet, un système linéaire de la forme

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,p}x_p = 0 \\ \quad \vdots \\ a_{n,1}x_1 + \ldots + a_{n,p}x_p = 0 \end{cases}$$

où $x = (x_1, \ldots, x_p)$ est l'inconnue, peut se mettre sous la forme

$$\begin{cases} \varphi_1(x) = 0 \\ \quad \vdots \\ \varphi_n(x) = 0 \end{cases}$$

avec $\varphi_i = \sum_{j=1}^{p} a_{i,j}e_j^*$ (on voit encore l'analogie base duale/coordonnées). L'ensemble $\mathcal{S}$ des solutions de ce système est donc

$$\mathcal{S} = \bigcap_{i=1}^{n} \ker(\varphi_i),$$

c'est une intersection de n hyperplans de K^p, sa dimension est donc au plus $p - n$.

Application. (Commutant d'une matrice trigonalisable)
Soit $A \in \mathcal{M}_n(K)$. On cherche la dimension de Z_A le commutant de A. Quitte à changer de base, on peut supposer A triangulaire supérieure. On cherche d'abord les éléments triangulaires supérieurs du commutant, notés Z_A^T. On a alors un système à $\frac{n(n+1)}{2}$ inconnues et autant d'équations à première vue. Cependant, les équations diagonales donnent $a_{i,i}x_i = x_i a_{i,i}$, ce qui est muet. On a donc en réalité au plus $\frac{n(n+1)}{2} - n$ équations indépendantes, donc

$$\dim(\mathrm{Z}_A) \geq \dim(\mathrm{Z}_A^T) \geq n.$$

On peut revenir au cas général par extension de corps.

Proposition. Par ailleurs si $F = \bigcap_{i=1}^{q} \ker(\varphi_i)$ une écriture de F comme intersection d'hyperplans, pour $q \geq n-k$, alors si $\varphi_1, \ldots, \varphi_r$ une base des $(\varphi_i)_{1\leq i\leq q}$, alors $\mathrm{codim}(F) = \mathrm{codim}((\varphi_i)^\circ_{1\leq i\leq r}) = \dim(\mathrm{Vect}(\varphi_i, 1 \leq i \leq r)) = r$. On a une inclusion et égalité des dimensions, donc

$$F = \bigcap_{i=1}^{n-k} \ker(\varphi_i).$$

1.2 Description matricielle, système de Cramer

Définition. De manière équivalente, un système linéaire est une équation de la forme

$$AX = B$$

avec $A \in \mathcal{M}_{n,p}(K)$, $X \in K^p$ et $B \in K^n$.

Théorème. L'ensemble des solutions du système est non vide si et seulement si $B \in \mathrm{Im}(A)$ et dans ce cas c'est un hyperplan affine dirigé par $\ker(A)$.

Soit à résoudre le système (S) suivant :

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,q}x_q = b_1 \\ \quad \vdots \\ a_{p,1}x_1 + \ldots + a_{p,q}x_q = b_p. \end{cases}$$

Théorème. (Système de Cramer)
On suppose $p = q = n$, on a alors un système de la forme $AX = B$ avec $A \in \mathcal{M}_n(K)$. La solution s'écrit

$$\forall 1 \leq i \leq n, x_i = \frac{\det(A_1, \ldots, A_{i-1}, B, A_{i+1}, \ldots, A_n)}{\det(A)}.$$

Théorème. (de Rouché-Fontené)
On revient au cas général, $A \in \mathcal{M}_{p,q}(K)$. Soit r le rang de A et A' une sous-matrice carrée de A de taille r inversible. Son déterminant est appelé déterminant principal du système. On a de même les équations et inconnues principales. Si $A' = (a_{i,j})_{i \in I, j \in J}$, on appelle déterminants caractéristiques les déterminants d'ordre $r+1$ obtenus en ajoutant sous A' une ligne de la forme $(a_{k,j})_{j \in J}$ puis une colonne de la forme $((b_i)_{i \in I}, b_k)$ avec $k \notin J$. Ils n'existent que si $p < r$, et sont alors au nombre de $p - r$. Le système (S) admet alors des solutions si et seulement si $p = r$ ou les $p - r$ déterminants caractéristiques sont nuls. Le système est alors équivalent au système des équations principales, les inconnues principales étant déterminées par un système de Cramer à l'aide des inconnues non principales.

Exemple. Si on prend (S) égal à

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 1 \\ x_1 \qquad - x_3 - x_4 = 1 \\ -x_1 + x_2 + x_3 + 2x_4 = m \end{cases}$$

pour m réel, le système n'admet des solutions que si $m = -1$ et dans ce cas $x_1 = 1 + x_3 + x_4$ et $x_2 = -x_4$.

2 Résolution pratique des systèmes linéaires

2.1 Pivot de Gauss

Définition. Une matrice est échelonnée en ligne si et seulement si le nombre de zéros précédant la première valeur non nulle d'une ligne augmente strictement ligne par ligne jusqu'à ce qu'il ne reste éventuellement plus que des zéros.

Remarque. Résoudre un système échelonné est très simple.

Définition. La matrice $I_n + \lambda E_{i,j}$ avec $j \neq i$ et $\lambda \neq 0$ est une matrice de *transvection élémentaire*, et $I_n + (\alpha - 1)E_{i,i}$ avec $\alpha \neq 1$ est une matrice de *dilatation élémentaire*. On appelle matrice de

transposition une matrice de permutation où la permutation est une transposition non triviale. Elles correspondent à des opérations sur les lignes (multiplication à gauche) ou sur les colonnes (à droite).

Théorème. (élimination de Gauss-Jordan)
On peut se ramener à une matrice échelonnée en lignes avec des opérations élémentaires sur les lignes.

Remarque. On peut aussi faire ce travail sur les colonnes.

Remarque. Au-delà de la résolution de systèmes, le pivot de Gauss permet de calculer le rang d'une matrice, son image, son noyau, et son inverse ou déterminant lorsque celle-ci est carrée. On peut donc déterminer si une famille est libre, génératrice.

Théorème. Le groupe $\mathrm{GL}_n(\mathbb{K})$ est engendré par les transvections et les dilatations élémentaires, le groupe $\mathrm{SL}_n(\mathbb{K})$ est engendré par les transvections élémentaires.

Application. Les groupes $\mathrm{GL}_n(\mathbb{C})$ et $\mathrm{SL}_n(\mathbb{K})$ sont connexes par arcs ; mais $\mathrm{GL}_n(\mathbb{R})$ a deux composantes connexes par arcs ; qui sont les matrices de déterminant positif et négatif.

2.2 Résolution approchée, moindres carrés

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les racines valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \mathrm{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0)$.

Corollaire. On peut définir la *pseudo-inverse* de A, notée $A^\dagger$, comme $U\Sigma^\dagger V^\top$ avec $\Sigma^\dagger \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A)^{-1}, \ldots, \sigma_r(A)^{-1}, 0, \ldots, 0)$. Pour $b \in \mathbb{R}^p$, le vecteur $A^\dagger b$ est l'unique solution de $Ax = b$ au sens des moindres carrés de norme minimale.

3 Cas des anneaux principaux

Ici, A désigne un anneau principal.

3.1 Échelonnement d'un système

Théorème. (forme normale de Hermite)
Soit A un anneau principal et $M \in \mathcal{M}_{n,p}(A)$. Il existe H une unique matrice échelonnée en colonnes telle qu'il existe $P \in \mathrm{GL}_p(A)$ telle que $MP = H$. On dit que H est la *forme de Hermite* de A.

Exemple. On a

$$\begin{pmatrix} 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} -2 & 1 & 5 \\ 0 & 1 & 0 \\ 1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

Les solutions entières de $2x + 3y + 5z = 0$ sont donc les $(t - 5s, t, -t - 2s), t, s \in \mathbb{Z}$.

3.2 Forme normale de Smith

Théorème. (forme normale de Smith) Soit $M \in \mathcal{M}_{n,p}(A)$ où A est un anneau principal. Il existe $(P, Q) \in \mathrm{GL}_n(A) \times \mathrm{GL}_p(A)$ telles que

$$PMQ = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Avec $d_1 | \dots | d_r$. De plus, les (d_i) sont uniques à association près. Cette forme diagonale est appelée *forme (normale) de Smith* de la matrice M, et les (d_i) sont appelés *facteurs invariants*.

Corollaire. Soit G un groupe abélien de type fini, il existe des uniques $s, n_1 | \cdots | n_r$ tels que

$$G \simeq \mathbb{Z}^s \oplus (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z}).$$

Remarque. Avec le théorème chinois, on peut trouver le nombre de groupe abéliens d'ordre donné.

Remarque. En appliquant l'algorithme de Smith à une matrice de la forme $XI_n - M$, $M \in \mathcal{M}_n(K)$, on trouve les facteurs invariants de M. En fait, avec la théorie des modules, on peut montrer le théorème des facteurs invariants de Frobenius avec le théorème de Smith.

Chapitre 30

170. Formes quadratiques sur un espace vectoriel de dimension finie. Orthogonalité. Applications.

Sommaire

1 Formes quadratiques et orthogonalité 211

1.1 Formes quadratiques 211

1.2 Orthogonalité 212

1.3 Bases orthogonales 213

2 Classification des formes quadratiques 213

2.1 Généralités, cas d'un corps algébriquement clos 213

2.2 Loi d'inertie de Sylvester, orthogonalisation simultanée 214

2.3 Sur un corps fini 215

Présentation du plan

Dans tout le plan, K a une caractéristique différente de deux et E est un K-espace vectoriel de dimension finie.

1 Formes quadratiques et orthogonalité

1.1 Formes quadratiques

Définition. Une *forme bilinéaire symétrique* $b : E \times E \to K$ est une application qui est bilinéaire est symétrique. Une *forme quadratique* est alors une application $q : E \to K$ de la forme $x \mapsto b(x, x)$.

Exemple. Sur $\mathbb{R}^3$, $(x, y, z) \mapsto 2xy - z^2$ est une forme quadratique.

Remarque. À partir d'une forme quadratique, on peut retrouver on peut retrouver l'unique

forme bilinéaire symétrique de laquelle elle dérive :

$$b(x,y) = \frac{1}{2}(q(x+y) - q(x)).$$

Proposition. Une forme quadratique peut être définie comme un polynôme de degré deux en les coordonnées, *i.e.* une application de la forme

$$x \mapsto \sum_{i,j} a_{i,j}\ell_i(x)\ell_j(x).$$

Remarque. Si (e_i) est la base antéduale des (ℓ_i), alors $a_{i,j} = b(e_i, e_j)$. La matrice de la forme q dans la base (e_i) est alors définie comme la donnée des $(a_{i,j})$. Si x est un vecteur écrit dans la base (e_i) alors $q(x) = X^\top AX$.

Proposition. Une forme quadratique est également la donnée d'un endomorphisme de E dans son dual par

$$\varphi_b : x \mapsto (y \mapsto b(\cdot, y)).$$

Définition. On définit son *noyau* comme le noyau de φ_b. On dit qu'elle est *dégénérée* quand celui-ci est non trivial. C'est un sous-espace vectoriel de E, noté $\ker(q)$. On définit le *rang* de q comme la codimension de son noyau.

Remarque. Le rang d'une forme quadratique est le rang de sa matrice dans n'importe quelle base.

Exemple. Le produit scalaire canonique est non dégénéré sur $\mathbb{R}^n$.

Définition. Soit q une forme quadratique, on définit son *cône isotrope* comme l'ensemble des x tels que $q(x) = 0$. Ce n'est en général pas un sous-espace vectoriel. On le note $\mathcal{I}(q)$.

Remarque. On a toujours $\ker(q) \subset \mathcal{I}(q)$.

Exemple. La forme $(x,y) = x^2 - y^2$ admet des vecteurs isotropes non triviaux, mais n'est pas dégénérée.

1.2 Orthogonalité

Définition. Soit b une forme bilinéaire symétrique, et A une partie de E. On définit l'orthogonale de A (pour b ou q associée) comme l'ensemble des x tels que $b(a,x) = 0$ pour tout $a \in A$. C'est un sous-espace vectoriel de E, noté $A^{\perp_q}$.

Remarque. Le noyau correspond à l'orthogonal de tout l'espace, et le cône isotrope correspond aux vecteurs qui sont orthogonaux avec eux-mêmes.

Proposition. Soient A, B deux parties de E, alors $A^\perp = (\mathrm{Vect}(A))^\perp = \mathrm{Vect}(A^\perp)$, et $A \subset (A^\perp)^\perp$. Si $A \subset B$ alors $B^\perp \subset A^\perp$.

1.3 Bases orthogonales

Proposition. Soit F un sous-espace vectoriel de E. Alors $E = F \oplus F^{\perp_q}$ si et seulement si q_F est non dégénérée.

Définition. On dit qu'une base (e_i) de E est (q-)orthogonale si $i \neq j$ implique $b(e_i, e_j) = 0$.

Théorème. Soit q un forme quadratique sur E. Alors q admet une base q-orthogonale.

Proposition. Si la forme quadratique est anisotrope, la matrice de changement de base peut être prise triangulaire supérieure.

2 Classification des formes quadratiques

2.1 Généralités, cas d'un corps algébriquement clos

Définition. On dit que deux formes quadratiques q et q' sont *équivalentes* lorsqu'il existe f un isomorphisme de E tel que $q' = q \circ f$. Cela équivaut à ce que les matrices A et B associées à ces deux formes quadratiques soient *congruentes*, *i.e.* qu'il existe une matrice P inversible telle que $B = P^\top AP$. Cela définit donc une relation d'équivalence sur l'ensemble des formes quadratiques (ou des matrices symétriques).

Définition. Le *discriminant* d'une forme quadratique q est le déterminant de la matrice de q dans n'importe quelle base modulo les carrés de K. Il ne dépend alors pas de la base choisie.

Remarque. Le théorème d'existence d'une base orthogonale nous dit que toute forme quadratique q peut s'écrire

$$q(x) = \sum_i \alpha_i (\ell_i(x))^2,$$

où les (ℓ_i) sont des formes linéaires sur E. Dans la base antéduale de (ℓ_i), notre forme s'écrit

$$q'(y) = \sum_i \alpha_i y_i^2.$$

On voit alors que dilater chaque vecteur permet de choisir α_i dans une classe modulo les carrés de K.

Exemple. Dans $\mathbb{R}$, le discriminant d'une forme quadratique peut être choisi dans $\{\pm 1\}$.

Théorème. Dans un corps algébriquement clos, deux formes quadratiques sont équivalentes si et seulement si elles ont même rang.

2.2 Loi d'inertie de Sylvester, orthogonalisation simultanée

On considère ici $K = \mathbb{R}$.

Théorème. (loi d'inertie de Sylvester)
Soit q une forme quadratique de E avec E de dimension finie n. Il existe une base de E dans laquelle q s'écrit

$$q(x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+m}^2.$$

Les entiers p et m ne dépendent pas de la base, et (p, m) est appelée *signature* de q. Deux formes quadratiques réelles sont alors équivalentes si et seulement si elles ont la même signature.

Proposition. (algorithme de réduction de Gauss)
Soit $q(x) = \sum_i a_{i,i}x_i^2 + 2\sum_{i<j} a_{i,j}x_ix_j$ une forme quadratique. Si $a_{1,1}$ est non nul, alors

$$q(x) = a_{1,1}\left(x_1 + \sum_{i=2}^{n} \frac{a_{1,i}}{a_{1,1}}x_i\right)^2 + q'(x),$$

en mettant sous forme canonique les termes contenant du $a_{1,1}$, et où q' ne dépend pas de x_1.
Si $a_{i,i}$ est nul pour tout i, et si $a_{1,2}$ est non nul, on écrit alors q comme $2\ell_1\ell_2 + q'$ et on procède de même avec l'identité

$$4XY = (X+Y)^2 - (X-Y)^2.$$

On obtient ainsi un algorithme pour mettre une forme quadratique sous forme de somme de carrés.

Définition. On dit qu'une forme quadratique réelle est *positive* quand sa signature est de la forme $(p, 0)$ et *négative* quand elle est de la forme $(0, m)$. On rajoute *définie* si p (resp. m) vaut n. On dit que c'est un *produit scalaire* quand elle est définie positive. On dit que la matrice symétrique associée est *positive, négative, définie positive, définie négative.*

Remarque. Une forme quadratique est positive si et seulement si $q(x) \geq 0$ pour tout x et définie positive si elle est positive et si $q(x) = 0$ entraîne $x = 0$.

Proposition. (factorisation de Cholesky)
Soit M une matrice symétrique définie positive. Alors il existe L triangulaire inférieure telle que $LL^\top = M$.

Théorème. (orthogonalisation simultanée)
Soit E un espace euclidien. Il existe une bijection entre les formes quadratiques de E et les opérateurs auto-adjoints de E. De plus, il existe une base orthogonale pour q et orthonormée pour le produit scalaire ambiant. De manière équivalente, toute matrice symétrique est diagonalisable dans une base orthonormée. La matrice est alors positive si et seulement si toutes ses valeurs propres sont positives, définie positive si elles sont toutes strictement positives.

Application. Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$. Il existe $U, V \in \mathrm{O}_n(\mathbb{R}) \times \mathrm{O}_p(\mathbb{R})$ et $\sigma_1 \geq \ldots \sigma_r > 0$ avec $r = \mathrm{rg}(A)$ telles que $A = U\mathrm{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)V^\top$.

Proposition. (Changement de base $\mathcal{C}^1$)

Soit A_0 une matrice symétrique réelle inversible. Alors il existe un voisinage V de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ telle que

$$A = \rho(A)^\top A_0 \rho(A) \qquad \forall A \in V.$$

Corollaire. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle

de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

2.3 Sur un corps fini

Définition. On note $\mathbb{F}_q^2$ les carrés de $\mathbb{F}_q$ et $\mathbb{F}_q^{*2}$ les carrés non nuls de $\mathbb{F}_q$.

Proposition. Si q est une puissance de deux alors $\mathbb{F}_q^2 = \mathbb{F}_q$. Sinon, $|\mathbb{F}_q^2| = \frac{q+1}{2}$.

Proposition. Dans le cas où q n'est pas une puissance de 2, $x \in \mathbb{F}_q^{*2}$ si est seulement si $x^{\frac{q-1}{2}} = 1$.

Proposition. Pour $a, b, c \in \mathbb{F}_q$, l'équation $ax^2 + bx^2 = c$ admet des solutions dans $\mathbb{F}_q$.

Théorème. Sur un corps fini, une forme quadratique non dégénérée de discriminant $\overline{a} \in K^*/(K^*)^2$ est équivalente à $x_1^2 + \ldots + ax_n^2$, on a donc deux classes d'équivalences de formes non dégénérées.

Application. (loi de réciprocité quadratique)
Soient p, q deux nombres premiers impairs. Alors

$$\left(\frac{p}{q}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

Chapitre 31

171. Formes quadratiques réelles. Coniques. Exemples et applications.

Présentation du plan

Chapitre 32

181. Convexité dans $\mathbb{R}^n$. Applications en algèbre et en géométrie.

Sommaire

1 Barycentres 219
1.1 Définition, exemples 219
1.2 Coordonnées barycentriques 220
2 Parties convexes 221
2.1 Résultats élémentaires 221
2.2 Lien avec les fonctions convexes 222
2.3 Projection sur un convexe fermé 222
2.4 Un théorème de point fixe 223

Présentation du plan

Dans tout le plan, on se place sauf mention contraire dans l'espace affine $\mathbb{R}^n$.

1 Barycentres

1.1 Définition, exemples

Définition. Soient $A_1, \dots, A_k$ des points et $\lambda_1, \dots, \lambda_k$ des scalaires. Il existe un unique point G tel que $\sum_{i=1}^k \lambda_i \overrightarrow{GA_i} = \overrightarrow{0}$, appelé *barycentre* des $((A_i, \lambda_i))$ et noté $\mathrm{Bar}((A_i, \lambda_i))$.

Remarque. Si l'on fixe O une origine, cela revient à $\overrightarrow{OG} = \frac{1}{\sum_{i=1}^k \lambda_i} \sum_{i=1}^k \lambda_i \overrightarrow{OA_k}$. Remarquons que l'on peut toujours supposer $\sum_{i=1}^k \lambda_i = 1$.

Définition. Avec les mêmes notations, si les λ_i sont tous égaux, on dit que G est l'*isobarycentre* des A_i. En particulier, l'isobarycentre de dexu points A, B est appelé *milieu* du segment $[A, B]$.

Proposition. Le barycentre est associatif ; on peut regrouper les A_i en paquets, calculer le G_j, μ_j correspondant au barycentre d'un paquet et à la somme des λ_i sur ce paquet puis recalculer le

barycentre des G_j, et on obtient la même chose qui si on calculait le barycentre des A_i d'un seul coup.

Application. Les médianes d'un triangle ABC se coupent en l'isobarycentre de celui-ci.

Proposition. Une application est affine si et seulement si elle préserve les barycentres. De plus, une partie est un sous-espace affine si et seulement si elle est stable par barycentres. Enfin, le sous-espace affine engendré par une partie est l'ensemble des barycentres de points de cette partie.

1.2 Coordonnées barycentriques

Définition. Soient $n+1$ points de l'espace, notés $(A_i)_i$. Les *coordonnées barycentriques* d'un point P de l'espace sont les λ_i qui font de P le barycentre des A_i. Elles ne sont pas uniques, sauf si on impose à la somme des λ_i d'être égale à 1, on dit alors qu'elles sont *normalisées*.

Exemple. Dans un triangle ABC, le milieu de $[A, B]$ a pour coordonnées barycentriques $(1, 1, 0)$. De même, dans un tétraèdre $ABCD$, le milieu de la face ABC a pour coordonnées barycentriques $(1, 1, 1, 0)$.

Dans la suite, on se place dans le plan affine et on se fixe un triangle ABC (non aplati), et on note les angles en A, B, C respectivement α, β, γ.

Proposition. Les coordonnées barycentriques d'un point P sont

$$(\det(\overrightarrow{PB}, \overrightarrow{PC}), \det(\overrightarrow{PC}, \overrightarrow{PA}), \det(\overrightarrow{PA}, \overrightarrow{PB})),$$

où le déterminant est pris dans n'importe quelle base. Si on veut les coordonnées normalisées, il faut prendre le déterminant dans la base $(\overrightarrow{AB}, \overrightarrow{AC})$. Ces coordonnées correspondent à des aires.

Proposition. (coordonnées barycentriques de points remarquables)
Avec les mêmes notations, voici les coordonnées barycentriques de points remarquables du triangle :

- Centre de gravité : $(1, 1, 1)$
- Centre du cercle inscrit : $(\sin(\alpha), \sin(\beta), \sin(\gamma))$
- Centre du cercle circonscrit : $(\sin(2\alpha), \sin(2\beta), \sin(2\gamma))$
- Orthocentre : $(\tan(\alpha), \tan(\beta), \tan(\gamma))$.

2 Parties convexes

2.1 Résultats élémentaires

Définition. Une partie C est dite *convexe* lorsque pour tout $x, y \in C$ et tout $t \in [0, 1]$, $tx + (1-t)y \in C$.

Exemple. Les convexes de $\mathbb{R}$ sont les intervalles, les boules sont convexes.

Exemple. L'ensemble $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe.

Remarque. Un convexe est connexe par arcs.

Proposition. Une partie est convexe si et seulement si tout barycentre à coefficients positifs de points de C est dans C. On appelle *enveloppe convexe* de $A \subset \mathbb{R}^n$ notée $\mathrm{Conv}(A)$ l'ensemble de tels barycentres. Alors C est convexe si et seulement s'il est égal à son enveloppe convexe.

Proposition. — Un sous-espace affine est convexe

— Les images directe et réciproque d'un convexe par une application affine sont convexes.

— Un demi-espace est convexe.

— Une intersection quelconque de convexes reste convexe.

Remarque. La notion d'enveloppe convexe permet de généraliser la notion d'être entre deux points sur $\mathbb{R}$. Ainsi, on peut par exemple obtenir un analogue au théorème de Rolle pour les polynômes mais sur $\mathbb{C}$.

Théorème. (de Gauss-Lucas)
Soit $P \in \mathbb{C}[X]$ non constant. Alors les racines de P' sont dans l'enveloppe convexe de celles de P.

Théorème. (de Carathéodory)
Dans $\mathbb{R}^n$, l'enveloppe convexe d'une partie A est l'ensemble des barycentres à coefficients positifs de familles de $n+1$ points de A.

Application. L'enveloppe convexe d'un compact est compacte.

Théorème. (de Helly)
Soit $(X_1, \ldots, X_{d+2})$ une famille de convexes de $\mathbb{R}^d$. Si pour tout choix de $d+1$ convexes parmi ceux-ci, l'intersection est non vide, l'intersection de tous les (X_i) est non vide.

2.2 Lien avec les fonctions convexes

Définition. Une fonction $f : C \to \mathbb{R}$ est *convexe* si son épigraphe est convexe, *i.e.* si pour tout couple de points (x, y) et tout $t \in [0, 1]$, $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Exemple. La fonction $\log \circ \det$ est convexe sur $\mathcal{S}_n^{++}(\mathbb{R})$.

Remarque. Les fonctions convexes sont très utiles en analyse, d'une part parce qu'à leurs propriétés synthétiques s'ajoutent des propriétés de régularité forte, et d'autre part car elles ont de bonnes propriétés d'optimisation. Par exemple, avec le cas de $\log \circ \det$ on peut montrer le résultat suivant.

Théorème. Soit K un compact de $\mathbb{R}^n$ dont l'intérieur contient 0. Il existe un unique ellipsoïde centré en 0 de volume minimal contenant K.

2.3 Projection sur un convexe fermé

Théorème. (projection sur un convexe fermé)
Soit C un convexe fermé non vide de $\mathbb{R}^n$.

— Pour tout $x \in \mathbb{R}^n$, il existe un unique point $p_C(x) \in C$ qui réalise la distance à C

— Il est caractérisé par la propriété suivante :

$$p_C(x) \in C \text{ et } \mathrm{Re}(\langle x - p_C(x), y - p_C(x)\rangle) \forall y \in C$$

— p_C est une application 1-lipschitzienne

— Si C est un sous-espace vectoriel de $\mathbb{R}^n$, alors $p_C(x)$ est caractérisé par

$$p_C(x) \in C \text{ et } (x - p_C(x)) \perp F.$$

Corollaire. Tout convexe fermé est strictement contenu dans un demi-espace.

Définition. Soit $C \subset \mathbb{R}^n$ un convexe non vide. On dit que $c \in C$ est un *point extrémal de* C si $C \backslash \{c\}$ est encore convexe, ce qui équivaut au fait que $\frac{x+y}{2} = c \Rightarrow x = y = c$ pour $x, y \in C$.
On dit qu'un hyperplan affine H est un *hyperplan d'appui* à C en $c \in C$ si $c \in H$ et si C est inclus dans un demi-espace formé par H.

Proposition. Soit C un convexe fermé non vide. Alors si $c \in \partial C$, il existe un hyperplan d'appui à C en c.

Proposition. Soit $a \in C$, si a est dans un hyperplan d'appui H de C, a est point extrémal de C si et seulement si a est point extrémal de $H \cap C$.

Théorème. (de Krein-Milman)
Soit C un convexe compact non vide de $\mathbb{R}^n$, alors il est enveloppe convexe de ses points extrémaux.

Remarque. Les points extrémaux se montrent utiles lorsqu'il s'agit d'étudier les isométries laissant invariant un solide, car on peut alors se ramener aux permutations des points extrémaux. Si ceux-ci sont en nombre fini, on se ramène alors à l'étude de sous-groupes de $\mathfrak{S}_n$, ce qui est souvent plus commode.

2.4 Un théorème de point fixe

Théorème. (point fixe de Kakutani)
Soit G un sous-groupe compact de $\mathrm{GL}(E)$ et K une partie non vide convexe compacte de E stable par G. Alors il existe dans K un point fixe sous G.

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors G possède un produit scalaire invariant si et seulement si G est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Chapitre 33

190. Méthodes combinatoires, problèmes de dénombrement.

Sommaire

1 Dénombrements élémentaires 226
1.1 Applications, parties 226
1.2 Lemme des tiroirs 226
1.3 Formule du crible 227
2 Théorie des groupes et algèbre linéaire 227
2.1 Actions de groupe 227
2.2 Dénombrement de matrices dans les corps finis 228
3 Formule d'inversion de Möbius 228
3.1 Convolution arithmétique 228
3.2 La formule et des applications 229
4 Séries génératrices 229
4.1 Anneau des séries formelles 229
4.2 Suites définies par une relation de récurrence 229
5 Loi uniforme sur les permutations 230
5.1 Points fixes 230
5.2 Nombre d'inversions 230
5.3 Nombre de cycles 231
5.4 Nombre de descentes 231

Présentation du plan

1 Dénombrements élémentaires

1.1 Applications, parties

Proposition. Soit X un ensemble à m éléments et Y à n éléments. Il y a n^m applications de X dans Y. Si on se restreint aux applications injectives, il y en a alors $\frac{n!}{(n-m)!}$.

Exemple. Le cardinal du groupe des permutations de X noté $\mathfrak{S}_X$ est $|X|!$.

Proposition. Soit X un ensemble fini. L'ensemble des parties de X a pour cardinal $2^{|X|}$. Le nombre de parties de X à k éléments est égal à $\binom{n}{k} := \frac{n!}{k!(n-k!)}$ appelé *coefficient binomial* k parmi n (avec la convention que cette quantité est nulle quand k est strictement plus grand que n). On a en particulier $2^n = \sum_{k=0}^{n} \binom{n}{k}$ pour tout n.

Proposition. (formule de Pascal)
Pour $0 < k < n$, $\binom{n}{k} = \binom{n-1}{k-1} + \binom{k}{n-1}$.

Proposition. (binôme de Newton)
Soient A un anneau et a et b qui commutent dans A, alors

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

1.2 Lemme des tiroirs

Proposition. (lemme des tiroirs)
Soit X à m éléments et Y à n éléments. Pour toute surjection f de X dans Y, il existe un élément y de Y tel que $f^{-1}(\{y\})$ comporte au moins $\lceil \frac{m}{n} \rceil$. C'est toujours vrai pour $m = \infty$.

Exemple. Soient cinq points dans un carré de coté 1. Alors au moins deux de ces points sont à distance $\frac{\sqrt{2}}{2}$.

Théorème. (de Bolzano-Weierstrass) Un segment $[a, b]$ de $\mathbb{R}$ est compact.

Théorème. (approximation de Dirichlet)
Soit $x \in \mathbb{R}$ et soit $N \geq 2$ un entier. Il existe deux entiers relatifs p, q avec $q \in \mathbb{N}^*$ inférieur à N tels que $|x - \frac{p}{q}| \leq \frac{1}{qN} \leq \frac{1}{q^2}$.

Remarque. Le jeu de l'approximation diophantienne consiste à augmenter l'exposant 2, plus précisément à trouver la borne supérieure des exposants μ tels que le nombre de rationnels $\frac{p}{q}$ que l'on peut placer entre x et $x \pm \frac{1}{q^\mu}$ est infini. Ce nombre μ est appelé *mesure d'irrationalité* de x. Plus $\mu(x)$ est grand, mieux x est approché par des rationnels, et, on va le voir, moins il est rationnel.

Proposition. La mesure d'un rationnel est égale à 1 et celle d'un irrationnel est supérieure ou égale à 2. La mesure de $\sqrt{2}$ est exactement égale à 2.

1.3 Formule du crible

Proposition. (formule du crible)
Si $A_1, \ldots, A_n$ sont des parties finies d'un ensemble E, alors

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\emptyset \neq J \subset \{1,\ldots,n\}} (-1)^{|J|-1} \left| \bigcap_{j \in J} A_j \right|.$$

Remarque. La preuve de ce résultat s'adapte directement à la théorie des probabilités pour calculer la probabilité d'une union d'événements.

Application. Soient α, β non nuls des éléments de $\mathbb{F}_q$, alors la forme quadratique $(x, y) \mapsto \alpha x^2 + \beta y^2$ représente 1.

Application. Le nombre de surjections de X à m éléments vers Y à n éléments est $\sum_{i=1}^{n} (-1)^{n-i} \binom{n}{i} i^m$.

2 Théorie des groupes et algèbre linéaire

2.1 Actions de groupe

Définition. La relation $x \sim y \Leftrightarrow \exists g \in G, g \cdot x = y$ est une relation d'équivalence sur X. Les classes d'équivalences pour cette relation sont appelées *orbites*, et forment donc une partition de X.

Définition. Soit $x \in X$ le *stabilisateur* de x est le sous-groupe de G formés des g tels que $g \cdot x = x$, noté $\mathrm{Stab}(x)$.

Exemple. Les orbites pour l'action de $\mathrm{O}_n(\mathbb{R})$ sur $\mathbb{R}^n$ sont les sphères centrées en 0. La partition de $\{1, \ldots, n\}$ en orbites sous l'action des itérées d'une permutation $\sigma \in \mathfrak{S}_n$ donne sa décomposition en produit de cycles à supports disjoints. Le stabilisateur d'un $k \in \{1, \ldots, n\}$ est isomorphe à $\mathfrak{S}_{n-1}$.

Proposition. Le groupe $G/\mathrm{Stab}(x)$ est en bijection avec l'orbite de x, pour tout $x \in X$.

Application. Le cardinal de la classe de conjugaison de $\sigma \in \mathfrak{S}_n$ de type $\{k_1, \ldots, k_n\}$ est $\frac{n!}{1^{k_1} k_1! \cdots n^{k_n} k_n!}$.

Théorème. (Équation aux classes)
Si X et G sont finis, et si $x_1, \ldots, x_n$ sont des représentants des orbites de X sous l'action de G, alors

$$|X| = \sum_{i=1}^{n} \frac{|G|}{|\mathrm{Stab}(x_i)|}.$$

Application. Le centre d'un p-groupe est non trivial.

Théorème. (formule de Burnside)
Toujours dans le cas où X et G sont finis, alors le nombre d'orbites sous l'action de G est égal à

$$\frac{1}{|G|} \sum_{g \in G} |\{x \in X, g \cdot x = x\}|.$$

Application. On dispose d'assiettes de c couleurs différentes, à placer sur une table pour n convives. Le nombre de dispositions possibles des assiettes (à rotation près donc) est de $\frac{1}{n} \sum_{d|n} \varphi(d)\, c^{n/d}$.

2.2 Dénombrement de matrices dans les corps finis

Proposition. Le cardinal de $\mathrm{GL}_n(\mathbb{F}_q)$ est $(q^n-1)\cdots(q^n-q^{n-1})$.

Corollaire. Le nombre de familles génératrices de $\mathbb{F}_q^n$ à $m \geq n$ éléments est, par dualité, $\prod_{i=0}^{n-1}(q^m-q^i)$.

Proposition. Le nombre de matrices diagonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|}.$$

Proposition. Le nombre de matrices nilpotentes sur $\mathcal{M}_n(\mathbb{F}_q)$ est $q^{d(d-1)}$.

3 Formule d'inversion de Möbius

3.1 Convolution arithmétique

Définition. Sur l'ensemble $\mathbb{C}^{\mathbb{N}^*}$, on définit la loi de *convolution arithmétique*, associative et commutative :

$$f * g : n \mapsto \sum_{d|n} f(d)g\left(\frac{n}{d}\right).$$

Elle fait de $\mathbb{A} := (\mathbb{C}^{\mathbb{N}^*}, +, *)$ une $\mathbb{C}$-algèbre. Son élément neutre est δ_1.

Proposition. Une fonction f est inversible dans $\mathbb{A}$ si et seulement si $f(1) \neq 0$.

Proposition. On définit μ la *fonction de Möbius* par $\mu(1) = 1$, $\mu(n) = 0$ si n a un facteur carré, et $\mu(p_1 \cdots p_r) = (-1)^r$ si les (p_i) sont des premiers distincts.

3.2 La formule et des applications

Théorème. (formule d'inversion de Möbius)
La fonction μ est l'inverse dans $\mathbb{A}$ de la fonction constante égale à 1. En particulier, si on a une relation de la forme $f(n) = \sum_{d|n} g(d)$, alors $g(n) = \sum_{d|n} f(d)\mu\left(\frac{n}{d}\right)$.

Application. Pour tout $n \in \mathbb{N}^*$, $\varphi(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right)$.

Théorème. Soit $I(n,q)$ le nombre d'irréductibles unitaires sur le corps à q éléments. Alors

$$I(n,q) = \frac{1}{n}\sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \sim \frac{q^n}{n}.$$

Théorème. Soit $(X_n)_n, (Y_n)_n \sim \bigotimes_n \mathcal{U}([\![1,n]\!])$ deux suites indépendantes, alors

$$\mathbb{P}(X_n \wedge Y_n = 1) \xrightarrow[n\to\infty]{} \frac{6}{\pi^2}.$$

Remarque. Cette convolution se généralise aux fonctions à valeurs dans un groupe abélien, et permet alors par exemple de montrer que les polynômes cyclotomiques sont à coefficients entiers.

4 Séries génératrices

4.1 Anneau des séries formelles

Définition. Si A est un anneau commutatif unitaire, on définit $A[\![X]\!]$ comme l'anneau des suites à valeurs dans A, muni de la somme terme à terme et du produit de Cauchy. Un élément $(a_n)_{n\in\mathbb{N}}$ est alors noté $\sum_{n\in\mathbb{N}} a_n X^n$.

Proposition. Un élément $\sum_{n\in\mathbb{N}} a_n X^n$ est inversible dans $A[\![X]\!]$ si et seulement si a_0 est inversible dans A.

Exemple. Dans $A[\![X]\!]$, $(1-X)^{-1} = \sum_{n\in\mathbb{N}} X^n$.

4.2 Suites définies par une relation de récurrence

Remarque. L'intérêt des séries génératrices est de simplifier l'étude de suites définies par une relation de récurrence.

Exemple. Si F_n est le n-ième terme de la suite de Fibonacci, alors $\sum_{n\in\mathbb{N}} F_n X^n = \frac{X}{1-X-X^2}$ dans $\mathbb{R}[\![X]\!]$, ce qui permet de calculer son terme général par décomposition en éléments simples.

Remarque. Quelques fois, on a en plus besoin de l'aspect analytique des séries entières pour obtenir des résultats.

Proposition. Soit c_n le nombre de mots bien parenthésés de longueur $2n$. Alors $c_n = \frac{1}{n+1}\binom{2n}{n}$.

Application. La marche aléatoire symétrique sur $\mathbb{Z}$ revient presque sûrement à l'origine, mais met en moyenne un nombre infini de pas pour cela.

5 Loi uniforme sur les permutations

Remarque. Dénombrer revient à calculer des probabilités sous la loi uniforme. Ainsi les probabilités et par exemple les inégalités de concentration ou les théorèmes limites peuvent être utiles en dénombrement. À ce sujet, la loi uniforme sur les permutations est un domaine qui permet de montrer cette utilité sous différents aspects.

5.1 Points fixes

Proposition. Sous la loi uniforme, le nombre moyen de points fixes d'une permutation est 1.

Proposition. La proportion de permutations de $\mathfrak{S}_n$ sans point fixe tend quand n tend vers l'infini vers $\frac{1}{\mathrm{e}}$.

5.2 Nombre d'inversions

Définition. Le *nombre d'inversions* d'une permutation σ est le nombre de couples (i,j) tels que $i<j$ et $\sigma(j)>\sigma(i)$. On note I_n qui à σ associe le nombre d'inversions de $\sigma\in\mathfrak{S}_n$.

Proposition. Si σ suit une loi uniforme sur $\mathfrak{S}_{n-1}$, et k suit une loi uniforme sur $\{1,\dots,n\}$, alors la permutation obtenue en insérant n à la k-ième place dans σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. La loi de I_n est celle d'une somme de n variables indépendantes de loi $\mathcal{U}(\{0,k-1\})$ où $k \in \{1,\dots,n\}$. En particulier, $\frac{6}{n^{3/2}}(I_n - \mathbb{E}(I_n))$ converge en loi vers une loi normale $\mathcal{N}(0,1)$.

5.3 Nombre de cycles

Proposition. On va construire une permutation aléatoire selon ses cycles. On met d'abord un 1. Puis si on a i nombres utilisés, on tire une $\mathcal{B}(\frac{1}{n-i+1})$. Si elle vaut 1, on clôt le cycle. Sinon on ne touche à rien. On choisit un nombre au hasard que l'on rajoute à la suite. On obtient alors une loi uniforme sur $\mathfrak{S}_n$.

Application. On note $K_n(\sigma)$ le nombre de cycles d'une permutation $\sigma \in \mathfrak{S}_n$. Sous la loi uniforme, K_n est une somme de $\mathcal{B}(\frac{1}{k})$ indépendantes, pour $k \in \{1,\dots,n\}$. En particulier, $\frac{K_n - \ln(n)}{\sqrt{\ln(n)}}$ converge en loi vers une loi normale $\mathcal{N}(0,1)$.

5.4 Nombre de descentes

Définition. Soit $\sigma \in \mathfrak{S}_n$ on dit que σ présente une *descente au rang* i si $\sigma(i+1) < \sigma(i)$. On note $A_{n,k}$ le nombre de permutations de $\mathfrak{S}_n$ qui possèdent exactement k descentes.

Proposition. Soit $(U_1,\dots,U_n) \sim \mathcal{U}(0,1)^{\otimes n}$. Pour presque tout $\omega \in \Omega$, il existe $\sigma(\omega)$ tel que $U_{\sigma(1)} < \dots < U_{\sigma(n)}$. Alors σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. Avec les mêmes notations, la loi du nombre de descentes est la loi $\lfloor U_1 + \dots + U_n \rfloor$. De plus, pour $\alpha \in [0,1]$, alors, lorsque n tend vers $+\infty$:

$$A(n,\alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}}\, \mathrm{e}^{-6(\alpha-1/2)^2}.$$

Chapitre 34

191. Exemples d'utilisation de techniques d'algèbre en géométrie.

Sommaire

1 Théorie des groupes et géométrie 234
1.1 Actions de groupes 234
1.2 Isométries laissant stable une partie du plan ou de l'espace 235
2 L'algèbre linéaire et le déterminant 235
2.1 Déterminant et volume 235
2.2 Valeurs propres, valeurs singulières et compression des données 236
3 Corps de nombres et opérations géométriques 237
3.1 Plan complexe et similitudes 237
3.2 Quaternions et rotations de l'espace 238
4 Convexité et barycentres 239
4.1 Parties convexes de l'espace 239
4.2 Cas particulier du triangle dans le plan, coordonnées barycentriques . . 240
5 Constructibilité à la règle et au compas 241
5.1 Nombres algébriques, transcendants 241
5.2 Nombres et polygones constructibles 241

Présentation du plan

1 Théorie des groupes et géométrie

1.1 Actions de groupes

Définition. Une *action de groupe* est la donnée d'une opération $G \times X \to X$, notée avec un point et qui vérifie :

— $\forall x \in X,\ 1 \cdot x = x$

— $\forall g, h \in G, \forall x \in X, g \cdot (h \cdot x) = (gh) \cdot x$.

De manière équivalente, cela revient à se donner un morphisme de G dans $\mathfrak{S}_X$ $(g \mapsto (x \mapsto g \cdot x))$.

Définition. Une action est *fidèle* si le morphisme qui lui correspond est injectif, c'est-à-dire si $(g \cdot x = x \forall x)$ implique $(g = 1)$. Une action est *libre* si pour $g \neq 1$, $g \cdot x \neq x$ pour tout $x \in X$. Une action est *transitive* si pour tout couple (x, y) d'éléments de X, il existe $g \in G$ tel que $g \cdot x = y$.

Remarque. Si G agit sur X alors $G/\ker(\varphi)$ (où φ est l'action en question) agit fidèlement sur X.

Définition. Un *espace affine* est la donnée d'un espace de points E, d'un espace vectoriel V qui agit fidèlement, librement et transitivement sur E (en tant que groupe abélien).

Remarque. Lorsque l'on travaille sur un espace vectoriel, on donne au vecteur nul un rôle particulier. En géométrie affine, on n'a pas cela, et cela est dû à la définition même d'un espace affine. En faisant agir notre espace vectoriel sur un espace de points, on donne alors à chaque point "le même rôle" et chaque point peut être une origine.

Proposition. Soit E un espace affine euclidien, et $\mathcal{X}$ un ensemble de parties de E de la forme $\mathcal{X} = \{s(X_0), s \text{ similitude de } E, X_0 \subset E\}$. Soit φ une isométrie affine qui préserve X_0 ($\varphi(X_0) = X_0$), alors pour $X := s(X_0)$, $s \circ \varphi \circ s^{-1}$ est une isométrie affine qui préserve X. On peut donc parler du groupe des isométries affines qui laissent invariant $\mathcal{X}$ plutôt que de parler des isométries affines laissant stable X_0. Inversement, traiter le cas de X_0 permet de traiter le cas de tout élément de $\mathcal{X}$.

Définition. Une *conique* du plan est le lieu d'annulation d'un polynôme de degré deux, autrement dit l'ensemble des x, y tels que

$$ax^2 + by^2 + cxy + dx + ey + f = 0,$$

avec a, b, c, d, e, f des réels, ce qui s'écrit $q(x, y) + \ell(x, y) + f = 0$ avec q une forme quadratique et ℓ une forme linéaire.

Théorème. (classification des coniques)
Sous l'action du groupe affine, une conique est soit une ellipse (ou un cercle), soit une parabole, soit une hyperbole, soit vide, soit une droite, soit deux droites. Sous l'action du groupe des isométries, il faut rajouter deux paramètres à l'hyperbole $(x^2/a^2 - x^2/b^2 = 1)$, deux à l'ellipse/cercle $(x^2/a^2 + x^2/b^2 = 1)$ et un pour une parabole $y^2 = 2px$.

Proposition. (ellipse de Steiner)
Soit ABC un triangle non aplati, il existe une unique ellipse dans le triangle qui est tangente au triangle en chacun de ses côtés.

1.2 Isométries laissant stable une partie du plan ou de l'espace

Définition. Pour une partie X de l'espace affine $\mathbb{R}^n$, on note $\mathrm{Is}(X)$ (resp. $\mathrm{Is}^+(X)$) l'ensemble des isométries affines (resp. déplacements) f de $\mathbb{R}^n$ telles que $f(X) = X$. Ces deux ensembles forment deux sous-groupes de $\mathrm{Isom}(\mathbb{R}^n)$.

Proposition. Si X est une partie convexe, alors les éléments de $\mathrm{Is}(X)$ permutent les points extrémaux de X.

Définition. Pour $n \geq 3$, on note D_n le groupe $\mathrm{Is}(P_n)$, où P_n est le polygone régulier à n côtés centré en 0 et dont 1 est l'un des côtés. On l'appelle *groupe diédral* d'ordre $2n$.

Théorème. D_n est un groupe de cardinal $2n$, engendré par r la rotation d'angle $\frac{2\pi}{n}$ et la symétrie orthogonale s d'axe $\mathbb{R}(1,0)$, avec la relation $(rs)^2 = \mathrm{id}$. De plus, $D_n^+ := \mathrm{Is}^+(P_n) \simeq \mathbb{Z}/n\mathbb{Z}$.

On note T le tétraèdre régulier et C le cube de $\mathbb{R}^3$.

Théorème. Le groupe $\mathrm{Is}(T)$ est isomorphe à $\mathfrak{S}_4$ et $\mathrm{Is}^+(T)$ à $\mathfrak{A}_4$.

Théorème. Le groupe $\mathrm{Is}^+(C)$ est isomorphe à $\mathfrak{S}_4$ et le groupe $\mathrm{Is}(C)$ est isomorphe à $\mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}$.

2 L'algèbre linéaire et le déterminant

2.1 Déterminant et volume

Proposition. Le déterminant d'une famille de vecteurs dans la base canonique correspond au volume orienté du parallélotope engendré par cette famille de vecteurs.

Exemple. Soit ABC un triangle dans le plan affine. Alors l'aire de ABC est égale à $\frac{1}{2}|\det(\overrightarrow{AB}, \overrightarrow{BC})|$.

Proposition. Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$ et K un compact de $\mathbb{R}^n$. Si on note λ la mesure de Lebesgue sur $\mathbb{R}^n$, alors $\lambda(AK) = |\det(A)|\lambda(K)$.

Remarque. C'est ce résultat qui donne la première intuition du jacobien comme modification élémentaire de volume.

Dans la suite E est un espace préhilbertien.

Définition. Soit $(x_1, \ldots, x_n)$ une famille de vecteurs de E. La *matrice de Gram* associé à cette famille est la matrice :

$$M = (\langle x_i, x_j \rangle)_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R}).$$

Son déterminant est appelé *déterminant de Gram* de la famille $(x_1, \ldots, x_n)$.

Proposition. Les matrices de Gram de taille n sont exactement les matrices symétriques positives de taille n.

Proposition. Soit $(x_1, \ldots, x_n)$ une famille de vecteurs de E. Elle est liée si et seulement si son déterminant de Gram est nul.

Théorème. Soit V un sous-espace vectoriel de E admettant une base $(x_1, \ldots, x_n)$. Alors pour tout $u \in E$:

$$d(u, V)^2 = \frac{G(u, x_1, \ldots, x_n)}{G(x_1, \ldots, x_n)}$$

Application. L'expression $\int_0^1 (1 + a_1 + a_2 x^2 + \ldots + a_n x^n)^2 \, \mathrm{d}x$ atteint son minimum $\mu = \frac{1}{(n+1)^2}$ en un unique n-uplet.

Théorème. (Inégalité de Hadamard)
Soit $(x_1, \dots, x_n)$ une famille de vecteurs non nuls de E. Alors

$$G(x_1, \dots, x_n) \leq \prod_{i=1}^{n} \|x_i\|^2,$$

avec égalité si et seulement si la famille en question est orthogonale.
Plus généralement,

$$|\det(x_1, \dots, x_n)| \leq \prod_{i=1}^{n} \|x_i\|.$$

Remarque. En particulier, l'inégalité de Hadamard dit que le volume d'un parallélotope est maximal quand celui-ci est rectangle.

2.2 Valeurs propres, valeurs singulières et compression des données

Proposition. Soit $S \in \mathcal{S}_p(\mathbb{R})$, alors

$$\max_{\substack{W \in \mathcal{M}_{k,p}(\mathbb{R}) \\ W^\top W = I_k}} \operatorname{Tr}(WSW^\top) = \sum_{i=1}^{k} \lambda_k,$$

avec égalité quand les lignes de W sont les k premiers vecteurs de la base orthonormée qui diagonalise S (ordonnés par ordre décroissant des valeurs propres).

Application. Cela sert en statistiques dans ce que l'on appelle l'*analyse en composantes principales*. La projection sur un sous-espace de dimension k qui maximise la variance de nos données est donnée par les k premiers axes propres de la matrice de covariance empirique.

Théorème. (minmax de Courant-Fischer)
Soit A une matrice hermitienne de taille n, dont les valeurs propres sont $\lambda_1 \geq \cdots \geq \lambda_n$. Si on note G_k l'ensemble des sous-espaces vectoriels de $\mathbb{C}^n$ de dimension k, alors

$$\lambda_k = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \langle Ax, x \rangle.$$

Définition. Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, on appelle *valeurs singulières* de A les valeurs propres (non nulles) de A^*A. On les classe par ordre décroissant et on les note ainsi $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$, où $r = \operatorname{rg}(A)$.

Théorème. Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$. Il existe U et V unitaires de taille n et p respectivement telles que $A = U\Sigma V^\top$, avec $\Sigma \in \mathcal{M}_{n,p}(\mathbb{C})$ diagonale, de diagonale $(\sigma_1(A), \dots, \sigma_r(A), 0, \dots, 0)$.

Corollaire. (minmax de Courant-Fischer pour les valeurs singulières)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $p \leq n$, alors si G_k est l'ensemble des sous-espaces de $\mathbb{R}^p$ de dimension k :

$$\sigma_k(A) = \max_{V \in G_k} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \min_{V \in G_{n-k+1}} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Théorème. (normes de Ky Fan)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{C})$ $(p \leq n)$. On appelle *k-ième norme de Ky Fan* la quantité $N_k(A) = \sigma_1(A) + \cdots + \sigma_k(A)$. Alors pour tout $k \in \{1, \ldots, p\}$ la k-ième norme de Ky Fan est une norme.

Remarque. La quantité $N_k(A)/N_p(A)$ permet de mesurer qu'ont les lignes de A à se compresser dans un sous-espace de dimension k. Plus précisément, la norme k de Ky Fan maximise la somme des projections des images par A d'une famille orthonormée de dimension k de $\mathbb{R}^p$ sur une famille orthonormée de dimension k de $\mathbb{R}^n$.

3 Corps de nombres et opérations géométriques

3.1 Plan complexe et similitudes

Définition. À chaque point (x, y) du plan on associe un nombre complexe $x + \mathrm{i}y$ (appelé *affixe* de (x, y)), et réciproquement. Ceci a un intérêt géométrique puisque l'on va pouvoir encoder des transformations géométriques par des nombres, sur lesquels les opérations sont "mécaniques".

Proposition. Soient A, B, C, D quatre points du plan d'affixes a, b, c, d. Alors $d(A, B) = |a - b|$ et si $A \neq B, C \neq D$, $(\overrightarrow{AB}, \overrightarrow{CD}) = \arg\left(\frac{d-c}{b-a}\right) \mod 2\pi$.

Application. (théorème de l'angle inscrit)
Soient A, B, C trois points distincts d'un cercle de centre Ω du plan, alors $2(\overrightarrow{AB}, \overrightarrow{AC}) = (\overrightarrow{\Omega B}, \overrightarrow{\Omega C}) \mod 2\pi$.

Proposition. Les similitude directes du plan correspondent exactement aux application $z \mapsto az+b$ de $\mathbb{C}$ dans $\mathbb{C}$ avec a, b complexes, a non nul. Une telle similitude peut se réécrire $s : z \mapsto az + (1-a)\omega$ avec $\omega = \frac{b}{1-a}$ l'unique point fixe de l'application. On a alors $s(z) - \omega = |a|\mathrm{e}^{\mathrm{i}\arg(a)}(z - \omega)$ donc s est la composition de la rotation de centre Ω d'affixe ω d'angle $\theta = \arg(a)$ composée avec l'homothétie de centre Ω et de rapport $\lambda = |a|$.

Application. Soit $ABCD$ un carré du plan tel que A et B sont à coordonnées entières. Alors C et D sont à coordonnées entières.

3.2 Quaternions et rotations de l'espace

Remarque. Le groupe $\mathrm{SO}_2(\mathbb{R})$ est homéomorphe au groupe des nombre complexes de module 1 par l'application $R_\theta \mapsto \mathrm{e}^{\mathrm{i}\theta}$.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a, b, c, d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1, 1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Corollaire. Si on note $\mathrm{SU}_k(\mathbb{C})$ l'ensemble des matrices unitaires de taille n dans $\mathbb{C}$ (*i.e.* $MM^* = I_n$), on a l'isomorphisme suivant :

$$\mathrm{PSU}_2(\mathbb{C}) := \mathrm{SU}_2(\mathbb{C})/\{-\mathrm{id}, \mathrm{id}\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

4 Convexité et barycentres

4.1 Parties convexes de l'espace

Définition. Soient $A_1, \ldots, A_k$ des points et $\lambda_1, \ldots, \lambda_k$ des scalaires. Il existe un unique point G tel que $\sum_{i=1}^k \lambda_i \overrightarrow{GA_i} = \overrightarrow{0}$, appelé *barycentre* des $((A_i, \lambda_i))$ et noté $\mathrm{Bar}((A_i, \lambda_i))$.

Remarque. Si l'on fixe O une origine, cela revient à $\overrightarrow{OG} = \frac{1}{\sum_{i=1}^k \lambda_i} \sum_{i=1}^k \lambda_i \overrightarrow{OA_k}$. Remarquons que l'on peut toujours supposer $\sum_{i=1}^k \lambda_i = 1$.

Définition. Avec les mêmes notations, si les λ_i sont tous égaux, on dit que G est l'*isobarycentre* des A_i. En particulier, l'isobarycentre de dexu points A, B est appelé *milieu* du segment $[A, B]$.

Proposition. Le barycentre est associatif ; on peut regrouper les A_i en paquets, calculer le G_j, μ_j correspondant au barycentre d'un paquet et à la somme des λ_i sur ce paquet puis recalculer le barycentre des G_j, et on obtient la même chose qui si on calculait le barycentre des A_i d'un seul coup.

Définition. Une partie C est dite *convexe* lorsque pour tout $x, y \in C$ et tout $t \in [0, 1]$, $tx + (1 - t)y \in C$.

Exemple. Les convexes de $\mathbb{R}$ sont les intervalles, les boules sont convexes.

Exemple. L'ensemble $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe.

Remarque. Un convexe est connexe par arcs.

Proposition. Une partie est convexe si et seulement si tout barycentre à coefficients positifs de points de C est dans C. On appelle *enveloppe convexe* de $A \subset \mathbb{R}^n$ notée $\mathrm{Conv}(A)$ l'ensemble de tels barycentres. Alors C est convexe si et seulement s'il est égal à son enveloppe convexe.

Proposition. — Un sous-espace affine est convexe

— Les images directe et réciproque d'un convexe par une application affine sont convexes.

— Un demi-espace est convexe.

— Une intersection quelconque de convexes reste convexe.

Théorème. (de Carathéodory)
Dans $\mathbb{R}^n$, l'enveloppe convexe d'une partie A est l'ensemble des barycentres à coefficients positifs de familles de $n+1$ points de A.

Application. L'enveloppe convexe d'un compact est compacte.

Théorème. (de Helly)
Soit $(X_1, \ldots, X_{d+2})$ une famille de convexes de $\mathbb{R}^d$. Si pour tout choix de $d+1$ convexes parmi ceux-ci, l'intersection est non vide, l'intersection de tous les (X_i) est non vide.

4.2 Cas particulier du triangle dans le plan, coordonnées barycentriques

Définition. Soient $n+1$ points de l'espace, notés $(A_i)_i$. Les *coordonnées barycentriques* d'un point P de l'espace sont les λ_i qui font de P le barycentre des A_i. Elles ne sont pas uniques, sauf si on impose à la somme des λ_i d'être égale à 1, on dit alors qu'elles sont *normalisées*.

Exemple. Dans un triangle ABC, le milieu de $[A, B]$ a pour coordonnées barycentriques $(1, 1, 0)$. De même, dans un tétraèdre $ABCD$, le milieu de la face ABC a pour coordonnées barycentriques $(1, 1, 1, 0)$.

Dans la suite, on se place dans le plan affine et on se fixe un triangle ABC (non aplati), et on note les angles en A, B, C respectivement α, β, γ.

Proposition. Les coordonnées barycentriques d'un point P sont

$$(\det(\overrightarrow{PB}, \overrightarrow{PC}), \det(\overrightarrow{PC}, \overrightarrow{PA}), \det(\overrightarrow{PA}, \overrightarrow{PB})),$$

où le déterminant est pris dans n'importe quelle base. Si on veut les coordonnées normalisées, il faut prendre le déterminant dans la base $(\overrightarrow{AB}, \overrightarrow{AC})$. Ces coordonnées correspondent à des aires.

Proposition. (coordonnées barycentriques de points remarquables)
Avec les mêmes notations, voici les coordonnées barycentriques de points remarquables du triangle :

— Centre de gravité : $(1, 1, 1)$

— Centre du cercle inscrit : $(\sin(\alpha), \sin(\beta), \sin(\gamma))$

— Centre du cercle circonscrit : $(\sin(2\alpha), \sin(2\beta), \sin(2\gamma))$

— Orthocentre : $(\tan(\alpha), \tan(\beta), \tan(\gamma))$.

5 Constructibilité à la règle et au compas

5.1 Nombres algébriques, transcendants

Définition. Soit L/K une extension de corps et $\alpha \in L$. L'ensemble $I = \{P \in K[X], P(\alpha) = 0\}$ est un idéal de $K[X]$, donc principal. S'il est nul on dit que α est *transcendant* sur K et sinon on dit qu'il est *algébrique* sur K. Le générateur unitaire de I est appelé *polynôme minimal* de α. C'est un polynôme irréductible, son degré est appelé *degré* de α (c'est aussi le degré de $K[\alpha]/K$).

Exemple. Les complexes i et $\sqrt[3]{2}$ sont algébriques sur $\mathbb{Q}$.

Remarque. Dans $\mathbb{C}$, les algébriques sur $\mathbb{Q}$ sont dénombrables donc "il y a beaucoup plus de transcendants". Cependant, il est souvent difficile de montrer qu'un nombre donné est transcendant. Le premier à y parvenir est Joseph Liouville en 1844.

Proposition. Soit α un réel algébrique de degré $d > 1$ sur $\mathbb{Q}$. Il existe $A > 0$ telle que pour tout $\frac{p}{q} \in \mathbb{Q}$,

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{A}{q^d}.$$

Corollaire. Le nombre $\sum_{n=0}^{+\infty} 10^{-n!}$ est transcendant sur $\mathbb{Q}$.

Théorème. (Lindemann, ADMIS)
Les nombres e et π sont transcendants.

Proposition. Si L/K est une extension de corps, l'ensemble des algébriques sur K forme un sous-corps de L.

5.2 Nombres et polygones constructibles

Définition. On dispose d'une règle non graduée, d'un compas et d'un repère orthonormé où l'unité est placée. Un nombre est dit *constructible* si l'on peut, en un nombre fini de droites et de cercles tracés, écrire $z \in \mathbb{C}$ comme point d'intersection de telles droites et cercles.

Remarque. Un complexe $z \in \mathbb{C}$ est constructible si et seulement si ses parties réelle et imaginaire le sont.

Exemple. Tous les nombres entiers (de Gauss) sont constructibles. Le nombre $\sqrt{2}$ est constructible.

Proposition. Soient x et y deux réels constructibles non nuls. Alors $x + y$, xy, $\frac{x}{y}$ et $\sqrt{x}$ sont constructibles.

Théorème. (Wantzel)
Le réel a est constructible si et seulement s'il existe une suite finie de corps $(L_i)_{0 \leq i \leq n}$ telle que

- $L_0 = \mathbb{Q}$
- L_{i+1} est une extension de degré deux de L_i
- $a \in L_n$.

On généralise cela sans mal à un complexe constructible.

Remarque. En particulier, le degré du polynôme minimal d'un nombre constructible est une puissance de deux, ce qui garantit l'impossibilité de la duplication du cube.

Théorème. (de Gauss-Wantzel) Un polygone régulier à n côtés est constructible si et seulement si n est le produit d'une puissance de deux et d'un produit de nombres premiers de Fermat distincts.

Exemple. On n'a pas de méthode exacte pour couper une tarte en sept.

Deuxième partie

Plans de leçons d'analyse et probabilités

Chapitre 35

201. Espaces de fonctions. Exemples et applications.

Sommaire

1 Applications linéaires continues 246
1.1 Généralités 246
1.2 Formes linéaires continues 246
1.3 Théorie de Baire 247
2 Fonctions continues sur compact 247
2.1 Continuité sur un compact 247
2.2 Compacité dans l'ensemble des fonctions continues sur un compact . . . 248
2.3 Densité dans l'ensemble des fonctions continues sur un compact 248
3 Espaces de Lebesgue 249
3.1 Définition 249
3.2 Régularisation et densité 249
3.3 Transformée de Fourier 250
3.4 Compacité dans les espace de Lebesgue 251
4 Espaces de Sobolev en dimension un 251
4.1 Définition 251
4.2 Résolution de problèmes aux limites 252

Présentation du plan

1 Applications linéaires continues

1.1 Généralités

Proposition. Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Une application linéaire $u \in \mathcal{L}(E, F)$ est continue si et seulement si il existe un réel $M > 0$ tel que pour tout $x \in E$, $\|u(x)\|_F \leq M\|x\|_E$. On note $\mathcal{L}_c(E, F)$ l'espace vectoriel des applications linéaires continues de E dans F.

Exemple. Soit $E = \mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme uniforme et $K \in \mathcal{C}^0([0,1]^2, \mathbb{R})$. Si on définit $T : E \to E$ par l'expression $T(f)(x) = \int_0^1 K(x,t)f(t)\,\mathrm{d}t$, alors T est une application linéaire continue.

Exemple. La forme linéaire $f \mapsto f(0)$ est continue dans E munie de la norme uniforme, mais pas avec la norme $\|\cdot\|_1 = \int_0^1 |\cdot|$.

Exemple. L'application $P \mapsto P'$ n'est pas continue munie de la norme définie par la somme des modules des coefficients.

Proposition. Soient E, F, G trois espaces vectoriels normés munis de leur norme respective.

— On peut définir une norme sur $\mathcal{L}_c(E,F)$ par $\|T\|_{E,F} = \sup\{\|T(x)\|_F, \|x\|_E = 1\}$, appelée *norme subordonnée.*

— Si F est complet pour sa norme, alors $\mathcal{L}_c(E,F)$ est complet pour $\|\cdot\|_{E,F}$.

— Si f est linéaire continue de E dans F et g de F dans G, alors $\|g \circ f\|_{E,G} \leq \|f\|_{E,F}\|g\|_{F,G}$.

Proposition. L'espace $\mathcal{L}_c(E)$ est une algèbre de Banach pour la norme subordonnée. L'ensemble de ses inversibles en est un ouvert.

Remarque. En dimension finie, la norme subordonnée est toujours atteinte. En dimension infinie, ce n'est pas toujours le cas.

Exemple. Dans $E = \mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme uniforme, $f \mapsto \int_0^1 f$ est continue, de norme atteinte 1, et $f \mapsto \int_0^1 f(t)(1-2t)\mathrm{d}t$ est également continue, mais la norme n'est pas atteinte.

1.2 Formes linéaires continues

Définition. Pour E un espace vectoriel normé, l'ensemble des formes linéaires continues de E muni de la norme subordonnée est appelé *dual topologique* de E, et noté E'.

Proposition. Le dual topologique d'un espace vectoriel normé est un espace de Banach.

Proposition. Soit f une forme linéaire sur E. Alors $f \in E'$ si et seulement si $\ker(f)$ est fermé. Ainsi, les hyperplans fermés sont exactement les noyaux de formes linéaires continues.

Théorème. (Hahn-Banach)
Soit E un espace vectoriel normé, et G un sous-espace vectoriel de E. Soit $g \in G'$. Il existe $f \in E'$ qui prolonge g telle que $\|g\|_{G'} = \|f\|_{E'}$.

Corollaire. Soit G un sous-espace vectoriel de E. Alors G est dense dans E si et seulement si toute forme linéaire continue qui s'annule sur G est nulle.

1.3 Théorie de Baire

Théorème. (de Baire)
Soit E un espace vectoriel normé complet. Alors une union dénombrable d'ouverts denses est dense.

Théorème. (de Banach-Steinhaus)
Soient E et F deux espaces vectoriels bornés, avec E complet. Soit $(T_i) \in \mathcal{L}_c(E,F)^I$. Si $(T_i(x))_i$ est bornée pour tout x, alors la famille (T_i) est bornée en norme subordonnée.

Application. Il existe une fonction continue 2π-périodique dont la série de Fourier diverge en 0.

Théorème. (de l'application ouverte, isomorphisme de Banach)
Soient E et F deux espaces de Banach et T linéaire continue surjective de E dans F. Alors T est ouverte. En particulier, si T est bijective alors T^{-1} est continue.

Application. L'opérateur série de Fourier n'est pas surjectif.

2 Fonctions continues sur compact

On considère ici (X,d) un espace métrique compact.

2.1 Continuité sur un compact

Proposition. Toute fonction continue de (X,d) dans un espace métrique est bornée et atteint ses bornes.

Proposition. On peut donc munir $\mathcal{C}^0(X,E)$ de la norme uniforme sur X, qui en fait un espace métrique complet lorsque (E,d_E) est complet.

Proposition. (Théorème de Dini)
Une suite croissante de fonctions continues de X compact dans $\mathbb{R}$ qui converge simplement converge uniformément.

2.2 Compacité dans l'ensemble des fonctions continues sur un compact

Définition. Soit $H \subset \mathcal{C}^0(X,E)$, E métrique. On dit que H est *équicontinue* en $x_0 \in X$ lorsque

$$\forall \varepsilon > 0, \exists \eta > 0, (d(x,x_0) \leq \eta \Rightarrow d_E(f(x),f(x_0)) \leq \varepsilon).$$

H est équicontinue sur X si elle est en tout point de X.

Théorème. (d'Ascoli)
Soient (X,d) et (E,d_E) deux espaces métriques, avec X compact. Alors $H \subset \mathcal{C}(X,E)$ est relativement compact si et seulement si

— H est équicontinue sur X

— Pour tout $x \in X$, $H_x := \{f(x), f \in H\}$ est relativement compact dans H.

Remarque. Sous la première hypothèse et dans le cas où X est connexe et E normé de dimension finie, on peut seulement vérifier la deuxième hypothèse pour un x_0.

Application. Soit F un sous-espace vectoriel fermé de $\mathcal{C}^0([0,1])$, que l'on suppose inclus dans $(\mathcal{C}^1([0,1]), \|\cdot\|_{\mathcal{C}^1})$. Alors F est de dimension finie.

2.3 Densité dans l'ensemble des fonctions continues sur un compact

Théorème. (Stone-Weierstrass)
Une sous-algèbre (unitaire) de $\mathcal{C}^0(X,\mathbb{R})$ est dense dans $\mathcal{C}^0(X,\mathbb{R})$ si elle sépare les points. Pour $\mathbb{C}$ il faut rajouter la stabilité par conjugaison.

Exemple. Dans les fonctions continues sur un segment à valeurs dans $\mathbb{R}$, les polynômes sont denses. Pour les fonctions à valeurs complexes, il faut prendre les polynômes trigonométriques.

Théorème. (Riesz-Markov, ADMIS)
Soit Λ une forme linéaire continue sur $\mathcal{C}^0(X,\mathbb{R})$ (X compact). Il existe une mesure signée μ définie sur $(X,\mathcal{B}(X))$ telle que pour tout $f\in\mathcal{C}^0([0,1],\mathbb{R})$, $\Lambda(f)=\int_X f\,\mathrm{d}\mu$.

Remarque. Ce théorème, couplé au critère de densité de Hahn-Banach, est un bon outil pour montrer la densité d'un sous-espace vectoriel dans $\mathcal{C}^0([0,1],\mathbb{R})$. Cependant, il nécessite beaucoup de travail et l'intervention des mesures signées.

Définition. Soit $n\in\mathbb{N}^*$ et $\sigma:\mathbb{R}\to[0,1]$ continue. Un *réseau de neurones à une couche cachée* sur $[0,1]^n$ est une fonction de la forme :

$$G_{N,\theta}:x\mapsto\sum_{k=1}^{N}\alpha_k\,\sigma(\langle y_k,x\rangle+\beta_k),$$

où $\theta=((\alpha_k)_k,(y_k)_k,(\beta_k)_k)\in\mathbb{R}^N\times(\mathbb{R}^k)^N\times\mathbb{R}^N=:\Theta$. On note $\mathcal{NN}_n(\sigma)=\{G_{N,\theta},N\in\mathbb{N}^*,\theta\in\Theta\}\subset\mathcal{C}([0,1]^n)$.

Théorème. (d'approximation universelle, Cybenko)
Soit σ une fonction continue de $\mathbb{R}$ dans $[0,1]$. On suppose que σ est *sigmoïdale*, *i.e.* σ est croissante et $\sigma(-\infty)=0$ et $\sigma(+\infty)=1$. Alors $\mathcal{NN}_n(\sigma)$ est dense dans $(\mathcal{C}([0,1]^n),\|.\|_\infty)$.

3 Espaces de Lebesgue

3.1 Définition

Définition. Pour $(\Omega,\mathcal{F},\mu)$ un espace mesuré, on définit $\mathrm{L}^p(\Omega,\mathcal{F},\mu)$ comme

— l'ensemble des fonctions p-intégrables sur Ω définies μ-presque partout si $p\in[1,\infty[$

— l'ensemble des fonctions essentiellement bornées sur Ω définies μ-presque partout si $p=\infty$.

On garde ces notations dans la suite.

Proposition. (inégalité de Hölder)
Soit $f\in\mathrm{L}^p(\Omega)$ et $g\in\mathrm{L}^q(\Omega)$ avec $\frac{1}{p}+\frac{1}{q}=1$, alors $fg\in\mathrm{L}^1(\Omega)$ avec

$$\int_\Omega|fg|\,\mathrm{d}\mu\le\left(\int_\Omega|f|^p\,\mathrm{d}\mu\right)^{1/p}\left(\int_\Omega|g|^q\,\mathrm{d}\mu\right)^{1/q}.$$

Corollaire. Pour $p\in[1,\infty]$, si on définit $\|f\|_p=(\int_\Omega|f|^p)^{1/p}$ pour $p<\infty$ et $\|f\|_\infty=\mathrm{esssup}(|f|)$, alors on fait de $\mathrm{L}^p(\Omega)$ un espace vectoriel normé.

Théorème. (Riesz-Fischer)
Pour tout p, $\mathrm{L}^p(\Omega)$ est un espace complet.

Corollaire. Si $(f_n) \in (\mathrm{L}^p(\Omega))^{\mathbb{N}}$ converge vers f dans $\mathrm{L}^p(\Omega)$, alors elle converge presque partout simplement vers f à extraction près.

Proposition. Les translations sont continues dans $\mathrm{L}^p(\mathbb{R}^d)$ pour $p \neq \infty$.

3.2 Régularisation et densité

Définition. Soient f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{C}$. On dit qu'elles sont *convoluables* lorsque $t \mapsto f(x-t)g(t)$ est intégrable pour presque tout x. On définit alors le *produit de convolution* de f et g par la formule :

$$f * g(x) = \int_{\mathbb{R}} f(x-t)g(t)\,\mathrm{d}t.$$

Théorème. L'espace $(\mathrm{L}^1(\mathbb{R}), +, *)$ est une algèbre de Banach commutative.

Proposition. On peut définir le produit de convolution d'une fonction intégrable f et d'une fonction test φ. Alors $f * \varphi$ est $\mathcal{C}^\infty$, avec $(f * \varphi)^{(n)} = f * \varphi^{(n)}$.

Définition. Soit $(\rho_n)_n$ une suite de fonctions positives de $\mathrm{L}^1(\mathbb{R})$. Cette suite est une *approximation de l'unité* si pour tout n, $\int_{\mathbb{R}} \rho_n = 1$ et pour tout $\delta > 0$, $\lim_{n\to\infty} \int_{|x|>\delta} \rho_n(x)\,\mathrm{d}x = 0$.

Exemple. Pour φ une fonction test, la suite $(n\varphi(n\cdot))_n$ définit une approximation de l'unité.

Théorème. Soit $p \in [1, +\infty[$, $k \in \mathbb{N}$ et (φ_n) une approximation de l'unité dans $(\mathcal{D}(\mathbb{R}))^{\mathbb{N}}$. Si $f \in \mathrm{L}^p(\mathbb{R})$, $\varphi_n * f \in \mathrm{L}^p(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R})$ et $\|\varphi_n * f - f\|_p \to f$.

Corollaire. Pour $p < \infty$, l'ensemble des fonctions test est dense dans $\mathrm{L}^p(\mathbb{R})$.

3.3 Transformée de Fourier

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Exemple. Si $g : x \mapsto \mathrm{e}^{-\alpha x}$, pour $\alpha > 0$, alors $\widehat{g} : \xi \mapsto \sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\xi^2}{4\alpha}}$.

Remarque. Il arrive qu'une fonction soit dans $\mathrm{L}^1(\mathbb{R})$ sans que sa transformée de Fourier le soit pour autant. C'est le cas par exemple de la fonction porte $x \mapsto \mathbf{1}_{[-1/2,1/2]}$, donc la transformée de Fourier est $\mathrm{sinc}(\pi\cdot) \notin \mathrm{L}^1$.

Remarque. Cette définition peut être étendue à $\mathrm{L}^1(\mathbb{R}^d)$ en considérant $\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\langle x,\xi\rangle} f(x)\,\mathrm{d}x$.

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}_0^0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}_0^0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un opérateur linéaire continu.

Proposition. La transformée de Fourier est un morphisme d'algèbres de Banach de $(\mathrm{L}^1(\mathbb{R}), +, *)$ dans $(\mathcal{C}_0^0(\mathbb{R}, +, \times))$.

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Proposition. Les fonctions test sont dans la classe de Schwartz, qui est donc dense dans L^p, $p < \infty$.

Proposition. La classe de Schwartz est stable par transformée de Fourier.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x} \widehat{f}(\xi)\, \mathrm{d}\xi.$$

Par dualité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$.

3.4 Compacité dans les espace de Lebesgue

Théorème. (Riesz-Fréchet-Kolmogorov)
Soit F une partie fermée de $\mathrm{L}^p(\mathbb{R})$. Alors F est compacte si et seulement si

- F est bornée
- Les translations sont continues uniformément en F
- Pour tout $\varepsilon > 0$, $\exists R > 0$, $\forall f \in F$,

$$\int_{|x|\geq R} |f(x)|^p \, \mathrm{d}x \leq \varepsilon.$$

4 Espaces de Sobolev en dimension un

Soit I un intervalle ouvert de $\mathbb{R}$

4.1 Définition

Définition. On définit l'*espace de Sobolev* $\mathrm{H}^1(I)$ par

$$\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I); \exists g \in \mathrm{L}^p(I), \int_I u\varphi' = -\int_I g\varphi, \forall \varphi \in \mathcal{C}_{\mathrm{c}}^1(I) \right\}.$$

Pour $u \in \mathrm{H}^1(I)$, le g de la définition est unique et on le note u', car il coïncide avec la dérivée usuelle lorsque u est $\mathcal{C}^1$.

Théorème. Muni de la norme $\sqrt{\| \cdot \|_2^2 + \| \cdot' \|_2^2}$, $\mathrm{H}^1(I)$ est un espace de Hilbert séparable. De plus, si $u \in \mathrm{H}^1(I)$, alors u admet un représentant continu.

Définition. On définit $\mathrm{H}_0^1(I)$ l'adhérence des fonctions test dans $\mathrm{H}^1(I)$. C'est un espace de Hilbert séparable muni de la norme induite.

Proposition. Si $I =]a, b[$, alors $u \in \mathrm{H}_0^1(I)$ si et seulement si $u(a) = u(b) = 0$.

Théorème. (Inégalité de Poincaré)
Supposons ici que I est borné. Il existe une constante $C > 0$ telle que pour $u \in \mathrm{H}_0^1(I)$, $\|u\|_{H^1} \leq C\|u'\|_2$.

4.2 Résolution de problèmes aux limites

Proposition. (Problème de Dirichlet) Soit $f \in \mathrm{L}^2(0,1)$ ($I =]0,1[$). On cherche à résoudre $-u'' + u = f$ avec les conditions aux limites $u(0) = u(1) = 0$. On dit que $u \in \mathrm{H}_0^1$ est une *solution faible* de ce problème si pour toute fonction test φ :

$$\int_I u'\varphi' + \int_I u\varphi = \int_I f\varphi.$$

Ce problème admet une unique solution faible.

Théorème. (Problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 36

203. Utilisation de la notion de compacité.

Sommaire

1 Compacité sur un espace métrique ou vectoriel normé 254
1.1 Définition générale de la compacité 254
1.2 Propriété de Bolzano-Weierstrass 254
2 Fonctions continues sur un compact 254
2.1 Problèmes d'extremums 254
2.2 Continuité uniforme 255
2.3 Points fixes 255
2.4 Densité dans l'ensemble des fonctions continues sur un compact 255
3 Compacité en dimension finie 256
3.1 Compacité des boules 256
3.2 Topologie matricielle 256
3.3 Équations différentielles 256
4 Compacité dans les espaces de fonctions 257
4.1 Compacité dans les fonctions continues sur un compact 257
4.2 Compacité dans les espaces de Lebesgue 257
4.3 Une autre forme de compacité dans les Hilbert 258

Présentation du plan

1 Compacité sur un espace métrique ou vectoriel normé

1.1 Définition générale de la compacité

Définition. Soit (X, d) un espace métrique. Il est si *compact* si de tout recouvrement fini de X on peut extraire un sous-recouvrement fini.

Exemple. Une partie finie d'un espace métrique est compacte.

Exemple. Dans $\mathbb{R}$, $]0, 1[$ n'est pas compact.

Proposition. Toute suite d'un espace compact admet une valeur d'adhérence.

Proposition. Toute partie compacte est fermée et bornée et toute partie fermée d'un compact est compacte. Une intersection de compacts est compacte, une union finie de compacts est compacte.

1.2 Propriété de Bolzano-Weierstrass

Proposition. Un espace métrique (X, d) est compact si et seulement si toute suite de X admet une valeur d'adhérence.

Exemple. Les compacts de $\mathbb{C}^n$ pour la topologie produit sont exactement les fermés bornés.

2 Fonctions continues sur un compact

2.1 Problèmes d'extremums

Proposition. Une fonction continue sur un compact est bornée et atteint ses bornes.

Remarque. On peut donc munir $\mathcal{C}^0(X, E)$ de la norme uniforme, qui en fait un espace vectoriel normé. Si E est un Banach, alors c'est en plus un Banach.

Proposition. (Théorème de Dini)
Une suite croissante de fonctions continues de X compact dans $\mathbb{R}$ qui converge simplement converge uniformément.

Exemple. La distance entre deux compacts est strictement positive.

Théorème. (de Rolle)
Soient $a < b$ dans $\mathbb{R}$ et $f : [a, b] \to \mathbb{R}$ continue sur l'intervalle fermé, dérivable sur l'intervalle ouvert, avec $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.

Remarque. Cela n'est plus vrai en général quand f est à valeurs complexes.

Théorème. (Égalité des accroissements finis)
Soit $f \in \mathcal{C}^0([a, b], \mathbb{R})$ dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$.

Application. (théorème de Darboux)
Soit $f : I \to \mathbb{R}$ dérivable sur I. Alors f' vérifie le théorème des valeurs intermédiaires.

Théorème. Toutes les normes sont équivalentes en dimension finie. Ainsi, les compacts sont exactement les fermés bornés.

Définition. Une fonction $f : E \mapsto \mathbb{R}$ continue, où E est un espace vectoriel est dite *propre* l'image réciproque de tout compact est compacte, et *coercitive* si $f(x) \to +\infty$ quand $\|x\|$ tend vers l'infini.

Proposition. Si E est de dimension finie, f est propre si et seulement si elle est coercitive, et admet alors un minimum sur E.

2.2 Continuité uniforme

Théorème. (de Heine)
Toute fonction continue sur un compact y est uniformément continue.

Remarque. Le théorème de Heine est utile pour définir l'intégrale de Riemann.

2.3 Points fixes

Théorème. Soit $f : X \to X$ une fonction continue d'un compact dans lui même et strictement contractante, alors f admet un unique point fixe sur X.

Application. Si je suis en France et que je regarde une carte de France, il y a un point sur la carte qui coïncide avec sa position réelle.

2.4 Densité dans l'ensemble des fonctions continues sur un compact

Théorème. (Stone-Weierstrass)
Soit A une sous-algèbre de $\mathcal{C}^0(X, \mathbb{C})$, où X est compact. On suppose que A sépare les points. Alors A est dense dans $\mathcal{C}^0(X, \mathbb{C})$ pour la norme uniforme. Pour $\mathbb{C}$, il faut ajouter la stabilité par conjugaison.

Corollaire. Les polynômes et les polynômes trigonométriques sont denses dans $\mathcal{C}^0([a, b], \mathbb{R})$ et $\mathcal{C}^0([a, b], \mathbb{C})$ respectivement.

Théorème. (Riesz-Markov, ADMIS)
Soit Λ une forme linéaire continue sur $\mathcal{C}^0(X, \mathbb{R})$ (X compact). Il existe une mesure signée μ définie sur $(X, \mathcal{B}(X))$ telle que pour tout $f \in \mathcal{C}^0([0, 1], \mathbb{R})$, $\Lambda(f) = \int_X f \, \mathrm{d}\mu$.

Remarque. Ce théorème, couplé au critère de densité de Hahn-Banach, est un bon outil pour montrer la densité d'un sous-espace vectoriel dans $\mathcal{C}^0([0, 1], \mathbb{R})$. Cependant, il nécessite beaucoup de travail et l'introduction des mesures signées.

Définition. Soit $n \in \mathbb{N}^*$ et $\sigma : \mathbb{R} \to [0, 1]$ continue. Un *réseau de neurones à une couche cachée* sur $[0, 1]^n$ est une fonction de la forme :

$$G_{N,\theta} : x \mapsto \sum_{k=1}^{N} \alpha_k \, \sigma(\langle y_k, x \rangle + \beta_k),$$

où $\theta = ((\alpha_k)_k, (y_k)_k, (\beta_k)_k) \in \mathbb{R}^N \times (\mathbb{R}^k)^N \times \mathbb{R}^N =: \Theta$. On note $\mathcal{NN}_n(\sigma) = \{G_{N,\theta}, N \in \mathbb{N}^*, \theta \in \Theta\} \subset \mathcal{C}([0, 1]^n)$.

Théorème. (d'approximation universelle, Cybenko)
Soit σ une fonction continue de $\mathbb{R}$ dans $[0, 1]$. On suppose que σ est *sigmoïdale*, *i.e.* σ est croissante et $\sigma(-\infty) = 0$ et $\sigma(+\infty) = 1$. Alors $\mathcal{NN}_n(\sigma)$ est dense dans $(\mathcal{C}([0, 1]^n), \|.\|_\infty)$.

3 Compacité en dimension finie

3.1 Compacité des boules

Théorème. (de Riesz)
On a vu précédemment que les boules fermées étaient compactes en dimension finie. Cela caractérise la dimension finie ; la boule unité n'est jamais compacte en dimension infinie.

3.2 Topologie matricielle

Proposition. Le groupe orthogonal est compact.

Application. (Décomposition polaire)
L'application $(S, O) \mapsto SO$ est un homéomorphisme de $\mathcal{S}_n^{++}(\mathbb{R}) \times \mathrm{O}_n(\mathbb{R})$ dans $\mathrm{GL}_n(\mathbb{R})$.

Théorème. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_($\mathbb{R})$.

3.3 Équations différentielles

Théorème. (de sortie de tout compact)
Soit Ω un ouvert de $\mathbb{R} \times \mathbb{R}^n$, et $f : \Omega \mapsto \mathbb{R}^n$ localement lipschitzienne en la seconde variable. Soit $x :]c, d[\mapsto \mathbb{R}^n$ une solution maximale de $x' = f(t, x)$. Alors pour tout compact K de Ω, il existe un voisinage V de d dans $]c, d[$ tel que pour tout $t \in V$, $(t, x(t)) \notin K$.

Corollaire. Sous les mêmes hypothèses, si $\Omega =]a, b[\times \mathbb{R}^n$. Si x est bornée au voisinage de d, alors $d = b$.

Application. (Théorème de Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors f est un $\mathcal{C}^1$-difféomorphisme sur $\mathbb{R}^n$ si et seulement si f est propre (l'image réciproque de tout compact est compacte) et pour tout $x \in \mathbb{R}^n$, $\mathrm{d}f(x)$ est inversible.

Application. Il existe $\varepsilon > 0$ tel que pour tout $a \in \mathbb{R}^n$, pour tout $L \in \mathrm{GL}_n(\mathbb{R})$ qui vérifie $\|a\| < \varepsilon$ et $\|L - I_n\| < \varepsilon$, alors il existe un difféomorphisme f de $\mathbb{R}^n$ tel que :

— $\forall \|x\| < 1, f(x) = L(x) + a$

— $\forall \|x\| > 2, f(x) = x$.

4 Compacité dans les espaces de fonctions

4.1 Compacité dans les fonctions continues sur un compact

Définition. Soit $H \subset \mathcal{C}^0(X, E)$, E métrique. On dit que H est *équicontinue* en $x_0 \in X$ lorsque

$$\forall \varepsilon > 0, \exists \eta > 0, (d(x, x_0) \leq \eta \Rightarrow d_E(f(x), f(x_0)) \leq \varepsilon).$$

H est équicontinue sur X si elle est en tout point de X.

Théorème. (d'Ascoli)
Soient (X, d) et (E, d_E) deux espaces métriques, avec X compact. Alors $H \subset \mathcal{C}(X, E)$ est relativement compact si et seulement si

— H est équicontinue sur X

— Pour tout $x \in X$, $H_x := \{f(x), f \in H\}$ est relativement compact dans H.

Remarque. Sous la première hypothèse et dans le cas où X est connexe et E normé de dimension finie, on peut seulement vérifier la deuxième hypothèse pour un x_0.

Application. Soit F un sous-espace vectoriel fermé de $\mathcal{C}^0([0,1])$, que l'on suppose inclus dans $(\mathcal{C}^1([0,1]), \|\cdot\|_{\mathcal{C}^1})$. Alors F est de dimension finie.

4.2 Compacité dans les espaces de Lebesgue

Théorème. (Riesz-Fréchet-Kolmogorov)
Soit F une partie fermée de $\mathrm{L}^p(\mathbb{R})$. Alors F est compacte si et seulement si

— F est bornée

— Les translations sont continues uniformément en F

— Pour tout $\varepsilon > 0$, $\exists R > 0$, $\forall f \in F$,

$$\int_{|x|\geq R} |f(x)|^p \, \mathrm{d}x \leq \varepsilon.$$

4.3 Une autre forme de compacité dans les Hilbert

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}^1_0(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 37

204. Connexité. Exemples et applications.

Sommaire

1 Espaces connexes 259
1.1 Premières définitions 259
1.2 Composantes connexes 260
1.3 Connexité par arcs 261
2 Le théorème des valeurs intermédiaires 261
2.1 Le résultat 261
2.2 Points périodiques d'un système dynamique 262
3 Passage du local au global 262
3.1 Théorème d'inversion globale 262
3.2 Principe du prolongement analytique 263

Présentation du plan

Par défaut, (X, d) est un espace métrique.

1 Espaces connexes

1.1 Premières définitions

Proposition. On a équivalence entre :

— Si X est union disjointe de deux ouverts, l'un des deux est vide

— Si X est union disjointe de deux fermés, l'un des deux est vide

— Toute partie de X ouverte est fermé est vide ou X

— Toute application continue de X dans $\mathbb{Z}$ est constante.

Définition. Un espace vérifiant l'une de ces propriétés est appelé espace *connexe*.

Exemple. — Un singleton est connexe

— Le cercle unité privé d'un point est connexe, mais pas privé de deux points.

— Un connexe de $\mathbb{R}$ est forcément un intervalle.

Remarque. Le fait qu'un intervalle de $\mathbb{R}$ est connexe est un résultat qui découle de l'existence de la borne supérieure dans $\mathbb{R}$.

Proposition. (passage à la douane)
Soit A une partie de X. Si C est une partie connexe qui rencontre l'intérieur et l'extérieur de A alors elle rencontre la frontière de A.

Proposition. Si les (A_i) sont des connexes dans X et si leur intersection est non vide, alors leur union est un connexe. Si les (A_i) sont une famille finie de connexes telle que l'intersection de deux de ces connexes est toujours non vide, alors leur union est un connexe.

Remarque. Une intersection de connexes n'est pas nécessairement un connexe (l'intersection d'une droite et d'un cercle dans le plan peut être deux points disjoints).

Théorème. Soit Y un autre espace métrique. Si f est une application continue de X dans Y et que X est connexe alors $f(X)$ l'est aussi. Si $A \subset X$ et $A \subset B \subset \overline{A}$, alors B est connexe également.

Théorème. Les connexes de $\mathbb{R}$ sont exactement les intervalles.

1.2 Composantes connexes

Définition. On définit une relation d'équivalence sur X par le fait d'appartenir à un connexe commun. Les *composantes* connexes de X sont les classes d'équivalence pour cette relation d'équivalence. On note $C(x)$ la composante connexe à laquelle appartient x.

Proposition. Pour $x \in X$, $C(x)$ est le plus grand connexe contenant x.

Théorème. Les ouverts de $\mathbb{R}$ sont les réunions dénombrables d'intervalles ouverts disjoints.

1.3 Connexité par arcs

Définition. L'espace X est dit *connexe par arcs* lorsque pour tous x, y, il existe un chemin continu joignant x à y.

Exemple. S'il existe un point a tel que pour tout x, il existe un chemin continu joignant a à x, on dit que X est *étoilée* par rapport à a, et elle est alors connexe par arcs.

Proposition. Un connexe par arcs est connexe, mais la réciproque est fausse en général (graphe de $\sin(1/x)$ et segment $[-1, 1]$ sur l'axe des ordonnées).

Exemple. Soit $f : \mathbb{R} \to \mathbb{R}$ continue, son épigraphe est connexe par arcs.

Exemple. Un convexe est connexe par arcs. L'ensemble des endomorphismes normaux d'un espace euclidien est étoilé par rapport à 0 mais n'est pas convexe.

Proposition. Les groupes $\mathrm{SO}_n(\mathbb{R})$ et $\mathrm{GL}_n(\mathbb{C})$ sont connexes par arcs pour tout $n \geq 1$, tandis que $\mathrm{O}_n(\mathbb{R})$ et $\mathrm{GL}_n(\mathbb{R})$ ont deux composantes connexes par arcs, déterminées par la valeur du déterminant.

2 Le théorème des valeurs intermédiaires

2.1 Le résultat

Théorème. (des valeurs intermédiaires)
Soit f une fonction continue de I intervalle dans $\mathbb{R}$. Alors pour $a < b \in I$, $[f(a), f(b)] \subset f([a, b])$.

Exemple. Un polynôme réel de degré impair admet une racine réelle.

Corollaire. Soit f continue de $[0, 1]$ dans lui-même. Alors f admet un point fixe.

Corollaire. (Théorème de la bijection continue)
Si I est un intervalle de $\mathbb{R}$, alors f réalise un homéomorphisme de I dans $f(I)$.

Exemple. Cela permet par exemple de définir l'arcsinus.

Théorème. (de Darboux)
Si f est une fonction dérivable sur un intervalle, sa dérivée vérifie le théorème des valeurs intermédiaires.

2.2 Points périodiques d'un système dynamique

On considère $f : D \to D$ continue où D est un segment.

Définition. Un point $x \in D$ est de période p ou p-périodique lorsque :

$$f^p(x) = x \text{ et } f^k(x) \neq x \text{ pour tout } k \in \{1, \dots, p-1\}.$$

Pour $p = 1$, cela correspond à un point fixe.

Définition. Soient I et J deux segments inclus dans D. On dit que I recouvre J (par f) lorsque $J \subset f(I)$. On note alors $I \rightsquigarrow J$.

Lemme. Si $I \rightsquigarrow J$, alors il existe un segment $K \subset I$ tel que $f(K) = J$.

Proposition. Supposons que l'on ait une chaîne $I_0 \rightsquigarrow I_1 \rightsquigarrow \cdots \rightsquigarrow I_{n-1} \rightsquigarrow I_0$. Alors il existe $x \in I_0$ tel que $f^n(x) = x$, et $f^k(x) \in I_k$ pour $k \in \{0, \dots, n-1\}$.

Théorème. (de Charkovski)
Si f possède un point de période 3, alors elle possède un point à toute période.

3 Passage du local au global

3.1 Théorème d'inversion globale

Théorème. (d'inversion locale)
Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction $\mathcal{C}^1$ et $a \in \Omega$ tel que le jacobien de f en a est non nul. Alors il existe un voisinage ouvert de $f(a)$ et un voisinage ouvert de a tel que f soit un $\mathcal{C}^1$-difféomorphisme entre ces deux voisinages.

Théorème. (d'inversion globale)
On suppose les hypothèses de l'inversion locale vérifiées pour tout a avec de plus f injective. Alors f réalise un $\mathcal{C}^1$-difféomorphisme de Ω dans son image.

Remarque. Ce théorème n'est pas pratique car montrer l'injectivité revient à calculer l'inverse.

Théorème. (Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors f est un $\mathcal{C}^1$-difféomorphisme sur $\mathbb{R}^n$ si et seulement si f est propre (l'image réciproque de tout compact est compacte) et pour tout $x \in \mathbb{R}^n$, $\mathrm{d}f(x)$ est inversible.

Application. Il existe $\varepsilon > 0$ tel que pour tout $a \in \mathbb{R}^n$, pour tout $L \in \mathrm{GL}_n(\mathbb{R})$ qui vérifie $\|a\| < \varepsilon$ et $\|L - I_n\| < \varepsilon$, alors il existe un difféomorphisme f de $\mathbb{R}^n$ tel que :

— $\forall \|x\| < 1, f(x) = L(x) + a$

— $\forall \|x\| > 2, f(x) = x$.

3.2 Principe du prolongement analytique

Théorème. (formule de Cauchy)
Soit $\gamma : I \to \mathbb{C}$ un chemin fermé dans un ouvert convexe Ω, et f holomorphe sur Ω. Pour $z \in \Omega \backslash \gamma(I)$:

$$f(z)\,\mathrm{Ind}_\gamma(z) = \frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{f(\zeta)}{\zeta - z}\,\mathrm{d}\zeta.$$

Corollaire. Pour tout ouvert du plan Ω, toute fonction holomorphe sur Ω y est développable en série entière (au voisinage de chaque point).

Théorème. (Principe des zéros isolés)
Soit $f \in \mathcal{H}(\Omega)$. L'ensemble des zéros de f est isolé dans Ω.

Corollaire. (Principe du prolongement analytique)
Si deux fonctions holomorphes sur Ω sont égales sur une partie qui possède un point d'accumulation dans Ω, alors elles sont égales sur Ω.

Exemple. Il existe une unique fonction holomorphe qui prolonge Γ à $\mathbb{C}\backslash\mathbb{Z}$.

Application. (formule des compléments) Pour tout $z \in \mathbb{C}\backslash\mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Exemple. Une fonction de la classe de Schwartz non nulle ne peut pas être à la fois à support compact et de transformée de Fourier à support compact.

Chapitre 38

205. Espaces complets. Exemples et applications.

Sommaire

1 Espaces complets 265
1.1 Suites de Cauchy 265
1.2 Complétude 266
1.3 Séries dans un espace de Banach 266
2 Utilisation de la complétude 267
2.1 Prolongement d'applications 267
2.2 Théorème de point fixe 267
2.3 Théorème de Baire 267
3 Espaces de Hilbert 268
3.1 Projection sur un convexe fermé 268
3.2 Espace de Sobolev en dimension un 269
3.3 Optimisation dans un espace de Hilbert 269

Présentation du plan

1 Espaces complets

1.1 Suites de Cauchy

Définition. Soit (X, d) un espace métrique. On dit que la suite $(x_n)_{n\in\mathbb{N}^*}$ est *de Cauchy* lorsque pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que pour tout $k \in \mathbb{N}$, $d(x_{N+k} - x_N) < \varepsilon$.

Remarque. L'ensemble des suites de Cauchy sur $\mathbb{Q}$ forme un $\mathbb{Q}$-espace vectoriel.

Proposition. Toute suite convergente est de Cauchy, et toute suite de Cauchy est bornée.

Remarque. Toute suite de Cauchy ne converge pas, et c'est là justement l'intérêt de la complétude. Par exemple, si on définit, sur $\mathbb{Q}$, la suite (x_n) par $x_0 = 1$ et $x_{n+1} = 1 + \frac{1}{x_n}$, on peut montrer qu'elle est de Cauchy mais qu'elle ne converge pas sur $\mathbb{Q}$.

1.2 Complétude

Définition. Un espace métrique est dit *complet* si toute suite de Cauchy converge. Si c'est de plus un espace vectoriel normé, alors on parle d'*espace de Banach*.

Théorème. (ADMIS)
Sur $\mathbb{Q}^{\mathbb{N}}$, on munit l'ensemble des suites de Cauchy de la relation $x \sim y \Leftrightarrow x_n - y_n \to 0$. On pose alors $\mathbb{R} := \mathbb{Q}/\sim$, que l'on munit des lois + et × et de la topologie héritées de $\mathbb{Q}$. On forme alors le corps des réels usuel, qui est par construction un espace de Banach.

Exemple. Un $\mathbb{R}$-espace vectoriel de dimension finie est complet. En particulier un $\mathbb{C}$-espace vectoriel de dimension finie est complet.

Exemple. L'espace $\mathbb{R}[X]$ n'est complet pour aucune norme.

Proposition. Pour tout $p \in [1, +\infty]$, $\ell^p(\mathbb{R})$ est complet muni de la norme p.

Proposition. Pour X un espace métrique compact, $\mathcal{C}^0(X, \mathbb{R})$ muni de la norme de la convergence uniforme.

Proposition. Une intersection d'espaces complets est complet. Une union finie d'espaces complets est complet.

Proposition. Un sous-espace d'un espace complet est fermé si et seulement s'il est complet.

Exemple. L'espace vectoriel $(\mathrm{c}_0(\mathbb{R}), \|.\|_\infty)$ n'est pas un espace de Banach.

Théorème. Si un espace métrique est complet alors toute suite décroissante de fermés non vides F_n de cet espace dont le diamètre tend vers 0 est d'intersection égale à un singleton.

1.3 Séries dans un espace de Banach

Proposition. Un espace est de Banach si et seulement si toute série absolument convergente y est convergente.

Exemple. Cela permet notamment de définir l'exponentielle dans une algèbre de Banach.

Théorème. (de Riesz-Fischer)
Pour toute mesure positive μ définie sur un espace mesurable, l'espace $\mathrm{L}^p(\mu)$ est un espace de Banach ($p \in [1, \infty]$).

Remarque. On retrouve le cas $\ell^p(\mathbb{R})$ exposé précédemment.

2 Utilisation de la complétude

2.1 Prolongement d'applications

Théorème. Soient E et F des espaces vectoriels normés, avec F complet. Soit f uniformément continue de X sur F, avec X dense dans E. Alors il existe une unique prolongement continu de f de E sur F.

Application. Ce théorème permet par exemple de prolonger la transformée de Fourier ou l'espérance conditionnelle.

2.2 Théorème de point fixe

Théorème. (de point fixe de Picard)
Soit E un espace vectoriel normé, et A une partie complète de E. Soit $f : A \to A$ une application contractante. Alors

— Il existe un unique $a \in A$ tel que $f(a) = a$,

— Pour tout x_0, la suite des itérées de x_0 par f converge géométriquement vers a.

Application. (théorème de Cauchy-Lipschitz)
Soit $f : I \times \mathbb{R}^m \to \mathbb{R}^m$ une application continue globalement lipschitzienne en la seconde variable. Alors le système différentiel

$$y' = f(t, y), y(t_0) = x$$

avec $t_0 \in I$ et $x \in \mathbb{R}^m$ donnés admet une unique solution définie sur I tout entier.

Exemple. Ce théorème a aussi d'autres applications ; le théorème d'inversion locale, la méthode de Newton.

2.3 Théorème de Baire

Théorème. (de Baire)
Soit E un espace vectoriel normé complet. Alors une union dénombrable d'ouverts denses est dense.

Théorème. (de Banach-Steinhaus)
Soient E et F deux espaces vectoriels bornés, avec E complet. Soit $(T_i) \in \mathcal{L}_c(E, F)^I$. Si $(T_i(x))_i$ est bornée pour tout x, alors la famille (T_i) est bornée en norme subordonnée.

Application. Il existe une fonction continue 2π-périodique dont la série de Fourier diverge en 0.

Théorème. (de l'application ouverte, isomorphisme de Banach)
Soient E et F deux espaces de Banach et T linéaire continue surjective de E dans F. Alors T est ouverte. En particulier, si T est bijective alors T^{-1} est continue.

Application. L'opérateur série de Fourier n'est pas surjectif.

3 Espaces de Hilbert

3.1 Projection sur un convexe fermé

Définition. Un espace *de Hilbert* est un espace préhilbertien pour lequel la norme le rend complet.

Exemple. Les espaces euclidiens et hermitiens sont des espaces de Hilbert, et $\mathrm{L}^2(\Omega, \mathcal{F}, \mu)$ et $\ell^2(\mathbb{Z})$ également.

Remarque. Un sous-espace vectoriel d'un Hilbert en est un si et seulement si il est fermé. En particulier, si $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$, alors $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ est un sous-espace vectoriel fermé de $\mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

Théorème. (projection sur un convexe fermé)
Soit H un espace de Hilbert et C un convexe fermé non vide de H.

— Pour tout $x \in H$, il existe un unique point $p_C(x) \in C$ qui réalise la distance à C

— Il est caractérisé par la propriété suivante :

$$p_C(x) \in C \text{ et } \mathrm{Re}(\langle x - p_C(x), y - p_C(x)\rangle) \forall y \in C$$

— p_C est une application 1-lipschitzienne

— Si C est un sous-espace vectoriel fermé de H, alors $p_C(x)$ est caractérisé par

$$p_C(x) \in C \text{ et } (x - p_C(x)) \perp F.$$

Remarque. Seule la complétude de C est nécessaire dans la preuve.

Proposition. L'espérance conditionnelle de $X \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ sachant la tribu $\mathcal{G}$ est définie comme la projection de X sur le convexe fermé $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ et vérifie alors, pour tout Y $\mathcal{G}$-mesurable :

$$\mathbb{E}(XY) = \mathbb{E}(Y\,\mathbb{E}(X|\mathcal{G})).$$

Théorème. (supplémentaire orthogonal)
Soit F un sous-espace vectoriel fermé de H un Hilbert. Alors H est somme directe orthogonale de F et $F^\perp$.

Corollaire. Un sous-espace vectoriel F est dense dans H si et seulement si $F^\perp = 0$.

Théorème. (représentation de Riesz)
Pour tout $\varphi \in H'$, où H est un espace de Hilbert, il existe un unique a dans H tel que $\varphi = \langle \cdot, a\rangle$.

3.2 Espace de Sobolev en dimension un

Définition. On définit l'*espace de Sobolev* $\mathrm{H}^1(I)$ par

$$\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I); \exists g \in \mathrm{L}^p(I), \int_I u\varphi' = -\int_I g\varphi, \forall \varphi \in \mathcal{C}^1_{\mathrm{c}}(I) \right\}.$$

Pour $u \in \mathrm{H}^1(I)$, le g de la définition est unique et on le note u', car il coïncide avec la dérivée usuelle lorsque u est $\mathcal{C}^1$.

Théorème. Muni de la norme $\sqrt{\| \cdot \|_2^2 + \| \cdot' \|_2^2}$, $\mathrm{H}^1(I)$ est un espace de Hilbert séparable. De plus, si $u \in \mathrm{H}^1(I)$, alors u admet un représentant continu.

Définition. On définit $\mathrm{H}^1_0(I)$ l'adhérence des fonctions test dans $\mathrm{H}^1(I)$. C'est un espace de Hilbert séparable muni de la norme induite.

Proposition. Si $I =]a, b[$, alors $u \in \mathrm{H}^1_0(I)$ si et seulement si $u(a) = u(b) = 0$.

Théorème. (Inégalité de Poincaré)
Supposons ici que I est borné. Il existe une constante $C > 0$ telle que pour $u \in \mathrm{H}^1_0(I)$, $\|u\|_{\mathrm{H}^1} \leq C\|u'\|_2$.

3.3 Optimisation dans un espace de Hilbert

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 39

206. Exemples d'utilisation de la notion de dimension finie en analyse.

Sommaire

1 Utilisation de la compacité 271
1.1 Équivalence des normes 271
1.2 Application aux équations différentielles ordinaires 272
1.3 Théorème de Riesz 272
2 Analyse hilbertienne 273
2.1 Projection sur un espace de dimension finie 273
2.2 Bases hilbertiennes 273
2.3 Séries de Fourier 274
3 Calcul différentiel 274
3.1 Matrices jacobienne et hessienne 274
3.2 Optimisation 275

Présentation du plan

1 Utilisation de la compacité

1.1 Équivalence des normes

Proposition. (Théorème de Bolzano-Weierstrass)
Les compacts de $\mathbb{R}^n$ muni de $\|\cdot\|_\infty$ sont les fermés bornés.

Théorème. En dimension finie, toutes les normes sont équivalentes. Ainsi, elles définissent les mêmes ouverts, fermés, suites convergentes.

Corollaire. Tout espace vectoriel de dimension finie est complet.

Corollaire. Les compacts en dimension finie sont exactement les fermés bornés.

Proposition. En dimension finie, une application est propre si et seulement si elle est coercitive.

1.2 Application aux équations différentielles ordinaires

Proposition. (Théorème de sortie de tout compact)
Soit $y' = f(t, y)$, $y(t_0) = y_0$ un problème de Cauchy sur $I \times E$, où E est un espace de Banach. Soit $(]T^-, T^+[, y)$ la solution maximale du problème. Si $T^+ < \sup(I)$, la trajectoire de y sort de tout compact au voisinage de T^+.

Corollaire. (Explosion en temps fini)
Dans le cas où E est de dimension finie, sous les mêmes hypothèses, si la solution est bornée sur son intervalle de définition, alors elle est globale.

Exemple. Si on a une majoration de la forme $\|f(t, y)\| \leq \alpha(t)\|y\| + \beta(t)$, avec α, β deux fonctions $\mathrm{L}^1_{\mathrm{loc}}$, alors les solutions sont globales.

Application. (Théorème de Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors f est un $\mathcal{C}^1$-difféomorphisme sur $\mathbb{R}^n$ si et seulement si f est propre (l'image réciproque de tout compact est compacte) et pour tout $x \in \mathbb{R}^n$, $\mathrm{d}f(x)$ est inversible.

Application. Il existe $\varepsilon > 0$ tel que pour tout $a \in \mathbb{R}^n$, pour tout $L \in \mathrm{GL}_n(\mathbb{R})$ qui vérifie $\|a\| < \varepsilon$ et $\|L - I_n\| < \varepsilon$, alors il existe un difféomorphisme f de $\mathbb{R}^n$ tel que :

— $\forall \|x\| < 1, f(x) = L(x) + a$

— $\forall \|x\| > 2, f(x) = x$.

1.3 Théorème de Riesz

Théorème. (de Riesz)
Un espace vectoriel normé réel est de dimension finie si et seulement si sa boule unité fermée est compacte.

Remarque. Ce résultat n'est plus vrai sur un $\mathbb{Q}$-espace vectoriel. En effet, la boule unité $[-1, 1]$ n'est pas compacte.

Application. Soit F un sous-espace vectoriel fermé de $(\mathcal{C}^0([0, 1]))$, que l'on suppose inclus dans $\mathcal{C}^1([0, 1])$. Alors F est de dimension finie.

2 Analyse hilbertienne

2.1 Projection sur un espace de dimension finie

Proposition. Soit F un sous-espace vectoriel de E préhilbertien. On a toujours $F \cap F^\perp = \{0\}$. Si de plus on a une somme directe, on dit que $F^\perp$ est le supplémentaire orthogonal de F, et on a alors $(F^\perp)^\perp = F$.

Exemple. Dans $\mathbb{R}[X]$ muni de $\int_0^1 PQ$, $X\mathbb{R}[X]$ n'a pas de supplémentaire orthogonal.

Proposition. Soit F un sous-espace vectoriel de E préhilbertien réel. Alors F admet un supplémentaire orthogonal si et seulement si la distance à F est toujours atteinte. Dans ce cas, le vecteur qui réalise la distance de x à F est unique et c'est la projection orthogonale de x sur F.

Proposition. (Gram-Schmidt)
Soit E un espace hermitien ou euclidien. Alors E admet une base orthonormée.

Théorème. Soit F est un sous-espace vectoriel de dimension finie de E préhilbertien réel, on a toujours $E = F \oplus F^\perp$. De plus, si $(e_1, \dots, e_n)$ est une bas orthonormée de F, alors pour tout $x \in E$:

$$p_F(x) = \sum_{k=1}^{n} \langle e_k, x\rangle e_k \text{ et } d(x, F) = \|p_{F^\perp}(x)\|.$$

Définition. On appelle *matrice de Gram* de la famille $(x_1, \dots, x_n) \in E^n$ la matrice $(\langle x_i, x_j\rangle)_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{R})$, et son déterminant le *déterminant de Gram* de la famille, noté $G(x_1, \dots, x_n)$. Une matrice est une matrice de Gram si et seulement si elle est symétrique positive. Une matrice de Gram est définie positive si et seulement si la famille en question est liée.

Théorème. Soit F un sous-espace vectoriel de E qui admet une base $(e_1, \dots, e_k)$. Pour $x \in E$, on a :

$$d(x, F)^2 = \frac{G(x, e_1, \dots, e_k)}{G(e_1, \dots, e_k)}.$$

Application. L'expression $\int_0^1 (1 + a_1 x + \ldots + a_n x^n)^2 \,\mathrm{d}x$ atteint son minimum $\mu = \frac{1}{n+1}$ en un unique n-uplet.

2.2 Bases hilbertiennes

Définition. Une famille de vecteurs de E préhilbertien est une *base hilbertienne* si c'est une famille orthonormée totale.

Théorème. (Bessel-Parseval)
Soit E préhilbertien, et $(e_i)_{i\in I}$ orthonormée. Pour tout $x \in E$, la famille $(\langle e_i, x\rangle^2)_i$ est sommable, avec

$$\sum_{i\in I} |\langle e_i, x\rangle|^2 \leq \|x\|^2,$$

avec égalité si et seulement si la famille est une base hilbertienne.

Remarque. Un espace préhilbertien est séparable si et seulement si il admet une base hilbertienne dénombrable, donc on peut se ramener à des sommes sur $\mathbb{N}$ (tous les Hilbert sur lesquels on travaille en général sont séparables).

Théorème. Soit E un espace préhilbertien et $(e_n)_n$ une base hilbertienne de E.

- $x \mapsto (\langle x, e_n\rangle)_n$ est une isométrie de E sur $\ell^2(\mathbb{N}, \mathbb{K})$.
- Pour tout $x \in E$, $x = \sum_{n=0}^{\infty} \langle x, e_n\rangle e_n$ et cette série converge commutativement.
- L'isométrie précédente est surjective si et seulement si E est un espace de Hilbert.

2.3 Séries de Fourier

Définition. On définit la *somme partielle de Fourier* par $S_N(f) = \sum_{n=-N}^{N} c_n(f) e_n$ et la série de Fourier (convergente ou non, c'est une série de fonctions) $S(f) = \sum_{n\in\mathbb{Z}} c_n(f) e_n$.

Théorème. La famille $(e_k)_k$ est une base hilbertienne de $\mathrm{L}^2_{2\pi}$.

Corollaire. On a donc l'égalité de Bessel et la formule de Parseval, la série de Fourier converge dans L^2 et on a une isométrie linéaire bijective.

Exemple. $\zeta(4) = \frac{\pi^4}{90}$.

3 Calcul différentiel

3.1 Matrices jacobienne et hessienne

Définition. Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ une application différentiable en un point $a \in U$. La différentielle de f en a est alors un élément de $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$, que l'on peut donc voir comme un élément de $\mathcal{M}_{n,p}(\mathbb{R})$. Son expression dans la base canonique est

$$\mathrm{Jac}(f)(a) = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}.$$

Remarque. L'intérêt d'avoir une jacobienne plutôt qu'une simple application linéaire est de pouvoir calculer plus facilement les compositions, ou encore avoir accès à la notion d'élément propre.

Remarque. L'idée générale du calcul différentiel est de se ramener au cas linéaire et d'appliquer des résultats "faciles" dans ce cas là avant de remonter au cas général avec des formules d'approximation (de Taylor par exemple).

Exemple. Avec la décomposition polaire sur les matrices, on montre que pour K un compact de $\mathbb{R}^n$ et $A \in \mathcal{M}_n(\mathbb{R})$, la mesure de Lebesgue de AK est $|\det(A)|$ fois la mesure de K. On en déduit ensuite la formule de changement de variables avec la formule de Taylor-Young.

Définition. Soit $y' = f(y)$ un système différentiel. Un *point d'équilibre* y_0 est un zéro de f. Il est dit *stable* si pour tout $\varepsilon > 0$, il existe un voisinage V de y_0 tel que toute solution du problème qui part d'un point de V reste proche de y_0 à ε près. Il est dit *asymptotiquement stable* s'il est stable et s'il existe un voisinage de y_0 tel que toutes les solutions qui partent de ce voisinage convergent vers y_0.

Proposition. (Stabilité des systèmes linéaires homogènes)
On considère le système différentiel $y' = Ay$, pour $A \in \mathcal{M}_n(\mathbb{R})$. Le point d'équilibre 0 est asymptotiquement stable si et seulement si les valeurs propres complexes de A ont toutes une partie réelle strictement négative, et il est stable si et seulement si toutes les valeurs propres ont une partie réelle négative, et que les valeurs propres de parties réelles nulles sont non défectives.

Proposition. (Linéarisation en première approche)
On considère le système différentiel $y' = f(y)$ sur $\mathbb{R}^n$, avec $f \in \mathcal{C}^1$. Soit y_0 un point d'équilibre de ce système et $A = \mathrm{Jac}(f)(y_0)$. Si toutes les valeurs propres de A ont une partie réelle strictement négative, alors y_0 est asymptotiquement stable.

Théorème. (Hartman-Grobman)
Avec les mêmes notations, on suppose que A n'a pas de valeur propre de partie réelle nulle, alors il existe deux voisinages de y_0 et 0 respectivement, et $h : U \to V$ un homéomorphisme qui envoie les trajectoires de $y' = f(y)$ sur celles de $x' = Ax$ en préservant le sens du temps.

Remarque. Si on est capable de classifier l'équivalence topologique des systèmes linéaires, on est capable, grâce à ce théorème, de classifier l'équivalence topologique de flots plus ou moins quelconques.

3.2 Optimisation

Soit Ω un ouvert de $\mathbb{R}^n$.

Définition. Un *point critique* de f est un $a \in \Omega$ tel que $\mathrm{d}f(a) = 0$.

Théorème.
- Si f est différentiable en a et admet un extremum local en a, alors a est un point critique de f.
- Si f est $\mathcal{C}^2$ admet un minimum local en a alors $\mathrm{H}f(a) \geq 0$.
- Si f est $\mathcal{C}^2$ admet un point critique en a de $\mathrm{H}f(a) > 0$ alors a est un minimum local de f.
- Si f est $\mathcal{C}^2$ admet un point critique en a et $\mathrm{H}f(a)$ est positive en un vecteur et négative en un autre, alors f n'admet pas d'extremum en a.

Exemple. La fonction $f(x, y, z) \mapsto x^2 + y^2 + z^3\mathrm{e}^x$ possède un point critique $(0, 0, 0)$ mais ce n'est pas un extremum.

Théorème. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

Théorème. (des extremums liés)
Soient $f, g_1, \ldots, g_p$ des fonctions réelles de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$, et X l'ensemble défini par les équations $g_1(x) = \ldots = g_p(x) = 0, x \in U$. Si la restriction de f à X admet un extremum local en a et si les différentielles $\mathrm{d}g_1(a), \ldots, \mathrm{d}g_p(a)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$ alors il existe des constantes $(\lambda_1, \ldots, \lambda_p)$ appelées *multiplicateurs de Lagrange* tels que $\mathrm{d}f(a) = \sum_{i=1}^{p} \lambda_i \mathrm{d}g_i(a)$.

Application. Le périmètre maximal d'un polygone inscrit dans un cercle est atteint pour un polygone régulier.

Chapitre 40

208. Espaces vectoriels normés, applications linéaires continues. Exemples.

Sommaire

1 Espaces vectoriels normés 277
1.1 Norme d'un espace vectoriel 278
1.2 Espaces de Banach 278
2 Application linéaires continues 279
2.1 Caractérisation 279
2.2 Formes linéaires continues 280
2.3 Théorie de Baire 281
3 Espaces de Hilbert 281
3.1 Théorie générale 281
3.2 L'exemple des espaces de Sobolev en dimension un 282
4 Parties compactes d'un espace vectoriel normé 283
4.1 Critères de compacité 283
4.2 Une autre forme de compacité dans un espace de Hilbert 284

Présentation du plan

1 Espaces vectoriels normés

On considère E un espace vectoriel réel ou complexe.

1.1 Norme d'un espace vectoriel

Définition. Une application $\|\cdot\| : E \mapsto \mathbb{R}_+$ est une *norme* lorsqu'elle vérifie les propriétés suivantes :

— *séparation.* Pour tout $x \in E$, $\|x\| = 0$ si et seulement si $x = 0$.

— *homogénéité.* Pour tout $x \in E$, $\lambda \in \mathbb{K}$: $\|\lambda x\| = |\lambda|\|x\|$.

— *inégalité triangulaire* ; Pour tout $(x, y) \in E^2$, $\|x + y\| \leq \|x\| + \|y\|$.

On dit alors que $(E, \|\cdot\|)$ est un espace vectoriel normé.

Définition. Une *boule ouverte* (resp. *fermée*), notée $\mathrm{B}(x, r)$ (resp. $\overline{\mathrm{B}}(x, r)$) est l'ensemble $\{y \in E, \|y - x\| < r\}$ (resp. $\leq$).

Remarque. Si $\|\cdot\|$ est une norme sur E, alors $(x, y) \mapsto \|x - y\|$ définit une métrique sur E, dont les boules sont les ensembles que l'on vient d'évoquer. On considère alors par défaut cette topologie dans un espace vectoriel normé.

Remarque. Le troisième axiome de la norme équivaut à la convexité des boules.

Exemple. Voici quelques exemples d'espaces vectoriels normés :

— $\mathbb{R}$ ou $\mathbb{C}$ muni du module.

— Tout espace préhilbertien muni de sa norme euclidienne.

— $\mathbb{R}^n$ ou $\mathbb{C}^n$ muni de $\|x\|_p = (\sum_{k=1}^n |x_k|^p)^{1/p}$, ou $\|x\|_\infty = \sup_k |x_k|$ selon la valeur de p.

— $\ell^p(\mathbb{N}, \mathbb{R})$ muni de $\|u\|_p = (\sum_{k=1}^\infty |u_k|^p)^{1/p}$, ou $\|u\|_\infty = \sup_k |u_k|$ selon la valeur de p.

— $\mathrm{L}^p(\Omega, \mathcal{A}, \mu)$ muni de $\|f\|_p = \left(\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{1/p}$.

— $\mathbb{R}[X]$ ou $\mathbb{C}[X]$ muni de $\|\cdot\|_p$ si on le voit en tant que fonction ou suite de coefficients.

Définition. Deux normes N_1 et N_2 sont équivalentes s'il existe des constantes A et B strictement positives telles que $AN_1 \leq N_2 \leq BN_2$. Deux normes équivalentes définissent les mêmes ouverts, les mêmes fermés, les mêmes suites convergentes etc...

Proposition. En dimension finie, les parties compactes sont exactement les parties fermées et bornées. En particulier, toutes les normes sont équivalentes en dimension finie.

1.2 Espaces de Banach

Définition. Un *espace de Banach* est un espace vectoriel normé complet pour la topologie de sa norme.

Proposition. Dans un espace de Banach, toute série absolument convergente converge et converge commutativement.

Exemple. Voici quelques exemples d'espaces de Banach :

— Tout espace vectoriel de dimension finie.

— $\ell^p(\mathbb{N}, \mathbb{R})$ muni de $\|u\|_p = (\sum_{k=1}^\infty |u_k|^p)^{1/p}$, ou $\|u\|_\infty = \sup_k |u_k|$ selon la valeur de p.

— $\mathrm{L}^p(\Omega, \mathcal{A}, \mu)$ muni de $\|f\|_p = \left(\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{1/p}$.

— $\mathcal{C}^0(X, F)$ muni de la norme uniforme si X est un espace métrique compact et si F est complet.

Exemple. L'espace $\mathbb{R}[X]$ n'est complet pour aucune norme.

Théorème. (de point fixe de Picard)
Soit E un espace vectoriel normé, et A une partie complète de E. Soit $f : A \to A$ une application contractante. Alors

— Il existe un unique $a \in A$ tel que $f(a) = a$,

— Pour tout x_0, la suite des itérées de x_0 par f converge géométriquement vers a.

Corollaire. (théorème de Cauchy-Lipschitz)
Soit $f : I \times \mathbb{R}^m \to \mathbb{R}^m$ une application continue globalement lipschitzienne en la seconde variable. Alors le système différentiel

$$y' = f(t, y), y(t_0) = x$$

avec $t_0 \in I$ et $x \in \mathbb{R}^m$ donnés admet une unique solution définie sur I tout entier.

2 Application linéaires continues

2.1 Caractérisation

Proposition. Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés. Une application linéaire $u \in \mathcal{L}(E, F)$ est continue si et seulement si il existe un réel $M > 0$ tel que pour tout $x \in E$, $\|u(x)\|_F \leq M\|x\|_E$. On note $\mathcal{L}_c(E, F)$ l'espace vectoriel des applications linéaires continues de E dans F.

Exemple. Soit $E = \mathcal{C}^0([0, 1], \mathbb{R})$ muni de la norme uniforme et $K \in \mathcal{C}^0([0, 1]^2, \mathbb{R})$. Si on définit $T : E \to E$ par l'expression $T(f)(x) = \int_0^1 K(x, t)f(t)\,\mathrm{d}t$, alors T est une application linéaire continue.

Exemple. La forme linéaire $f \mapsto f(0)$ est continue dans E munie de la norme uniforme, mais pas avec la norme $\|\cdot\|_1 = \int_0^1 |\cdot|$.

Exemple. L'application $P \mapsto P'$ n'est pas continue munie de la norme définie par la somme des modules des coefficients.

Proposition. Soient E, F, G trois espaces vectoriels normés munis de leur norme respective.

— On peut définir une norme sur $\mathcal{L}_c(E, F)$ par $\|T\|_{E,F} = \sup\{\|T(x)\|_F, \|x\|_E = 1\}$, appelée *norme subordonnée.*

— Si F est complet pour sa norme, alors $\mathcal{L}_c(E, F)$ est complet pour $\|\cdot\|_{E,F}$.

— Si f est linéaire continue de E dans F et g de F dans G, alors $\|g \circ f\|_{E,G} \leq \|f\|_{E,F}\|g\|_{F,G}$.

Proposition. L'espace $\mathcal{L}_c(E)$ est une algèbre de Banach pour la norme subordonnée. L'ensemble de ses inversibles en est un ouvert.

Remarque. En dimension finie, la norme subordonnée est toujours atteinte. En dimension infinie, ce n'est pas toujours le cas.

Exemple. Dans $E = \mathcal{C}^0([0, 1], \mathbb{R})$ muni de la norme uniforme, $f \mapsto \int_0^1 f$ est continue, de norme atteinte 1, et $f \mapsto \int_0^1 f(t)(1 - 2t)\mathrm{d}t$ est également continue, mais la norme n'est pas atteinte.

2.2 Formes linéaires continues

Définition. Pour E un espace vectoriel normé, l'ensemble des formes linéaires continues de E muni de la norme subordonnée est appelé *dual topologique* de E, et noté E'.

Proposition. Le dual topologique d'un espace vectoriel normé est un espace de Banach.

Proposition. Soit f une forme linéaire sur E. Alors $f \in E'$ si et seulement si $\ker(f)$ est fermé. Ainsi, les hyperplans fermés sont exactement les noyaux de formes linéaires continues.

Théorème. (Hahn-Banach)
Soit E un espace vectoriel normé, et G un sous-espace vectoriel de E. Soit $g \in G'$. Il existe $f \in E'$ qui prolonge g telle que $\|g\|_{G'} = \|f\|_{E'}$.

Corollaire. Soit G un sous-espace vectoriel de E. Alors G est dense dans E si et seulement si toute forme linéaire continue qui s'annule sur G est nulle.

Théorème. (Riesz-Markov, ADMIS)
Soit Λ une forme linéaire continue sur $\mathcal{C}^0(X, \mathbb{R})$ (X compact). Il existe une mesure signée μ définie sur $(X, \mathcal{B}(X))$ telle que pour tout $f \in \mathcal{C}^0([0,1], \mathbb{R})$, $\Lambda(f) = \int_X f \, \mathrm{d}\mu$.

Définition. Soit $n \in \mathbb{N}^*$ et $\sigma : \mathbb{R} \to [0,1]$ continue. Un *réseau de neurones à une couche cachée* sur $[0,1]^n$ est une fonction de la forme :

$$G_{N,\theta} : x \mapsto \sum_{k=1}^{N} \alpha_k \, \sigma(\langle y_k, x \rangle + \beta_k),$$

où $\theta = ((\alpha_k)_k, (y_k)_k, (\beta_k)_k) \in \mathbb{R}^N \times (\mathbb{R}^k)^N \times \mathbb{R}^N =: \Theta$. On note $\mathcal{NN}_n(\sigma) = \{G_{N,\theta}, N \in \mathbb{N}^*, \theta \in \Theta\} \subset \mathcal{C}([0,1]^n)$.

Théorème. (d'approximation universelle, Cybenko)
Soit σ une fonction continue de $\mathbb{R}$ dans $[0,1]$. On suppose que σ est *sigmoïdale*, *i.e.* σ est croissante et $\sigma(-\infty) = 0$ et $\sigma(+\infty) = 1$. Alors $\mathcal{NN}_n(\sigma)$ est dense dans $(\mathcal{C}([0,1]^n), \|\cdot\|_\infty)$.

2.3 Théorie de Baire

Théorème. (de Baire)
Soit E un espace vectoriel normé complet. Alors une union dénombrable d'ouverts denses est dense.

Théorème. (de Banach-Steinhaus)
Soient E et F deux espaces vectoriels bornés, avec E complet. Soit $(T_i) \in \mathcal{L}_c(E,F)^I$. Si $(T_i(x))_i$ est bornée pour tout x, alors la famille (T_i) est bornée en norme subordonnée.

Application. Il existe une fonction continue 2π-périodique dont la série de Fourier diverge en 0.

Théorème. (de l'application ouverte, isomorphisme de Banach)
Soient E et F deux espaces de Banach et T linéaire continue surjective de E dans F. Alors T est ouverte. En particulier, si T est bijective alors T^{-1} est continue.

Application. L'opérateur série de Fourier n'est pas surjectif.

3 Espaces de Hilbert

3.1 Théorie générale

Définition. Un espace *de Hilbert* est un espace préhilbertien pour lequel la norme le rend complet.

Exemple. Les espaces euclidiens et hermitiens sont des espaces de Hilbert, et $\mathrm{L}^2(\Omega, \mathcal{F}, \mu)$ et $\ell^2(\mathbb{Z})$ également.

Remarque. Un sous-espace vectoriel d'un Hilbert en est un si et seulement si il est fermé. En particulier, si $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$, alors $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ est un sous-espace vectoriel fermé de $\mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

Théorème. (projection sur un convexe fermé)
Soit H un espace de Hilbert et C un convexe fermé non vide de H.

— Pour tout $x \in H$, il existe un unique point $p_C(x) \in C$ qui réalise la distance à C

— Il est caractérisé par la propriété suivante :

$$p_C(x) \in C \text{ et } \mathrm{Re}(\langle x - p_C(x), y - p_C(x)\rangle) \forall y \in C$$

— p_C est une application 1-lipschitzienne

— Si C est un sous-espace vectoriel fermé de H, alors $p_C(x)$ est caractérisé par

$$p_C(x) \in C \text{ et } (x - p_C(x)) \perp F.$$

Remarque. Seule la complétude de C est nécessaire dans la preuve.

Proposition. L'espérance conditionnelle de $X \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ sachant la tribu $\mathcal{G}$ est définie comme la projection de X sur le convexe fermé $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ et vérifie alors, pour tout Y $\mathcal{G}$-mesurable :

$$\mathbb{E}(XY) = \mathbb{E}(Y\,\mathbb{E}(X|\mathcal{G})).$$

Théorème. (supplémentaire orthogonal)
Soit F un sous-espace vectoriel fermé de H un Hilbert. Alors H est somme directe orthogonale de F et $F^\perp$.

Corollaire. Un sous-espace vectoriel F est dense dans H si et seulement si $F^\perp = 0$.

Théorème. (représentation de Riesz)
Pour tout $\varphi \in H'$, où H est un espace de Hilbert, il existe un unique a dans H tel que $\varphi = \langle \cdot, a\rangle$.

3.2 L'exemple des espaces de Sobolev en dimension un

Définition. On définit l'*espace de Sobolev* $\mathrm{H}^1(I)$ par

$$\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I); \exists g \in \mathrm{L}^p(I), \int_I u\varphi' = -\int_I g\varphi, \forall \varphi \in \mathcal{C}^1_{\mathrm{c}}(I) \right\}.$$

Pour $u \in \mathrm{H}^1(I)$, le g de la définition est unique et on le note u', car il coïncide avec la dérivée usuelle lorsque u est $\mathcal{C}^1$.

Théorème. Muni de la norme $\sqrt{\| \cdot \|_2^2 + \| \cdot' \|_2^2}$, $\mathrm{H}^1(I)$ est un espace de Hilbert séparable. De plus, si $u \in \mathrm{H}^1(I)$, alors u admet un représentant continu.

Définition. On définit $\mathrm{H}_0^1(I)$ l'adhérence des fonctions test dans $\mathrm{H}^1(I)$. C'est un espace de Hilbert séparable muni de la norme induite.

Proposition. Si $I =]a, b[$, alors $u \in \mathrm{H}_0^1(I)$ si et seulement si $u(a) = u(b) = 0$.

Théorème. (Inégalité de Poincaré)
Supposons ici que I est borné. Il existe une constante $C > 0$ telle que pour $u \in \mathrm{H}_0^1(I)$, $\|u\|_{H^1} \leq C\|u'\|_2$.

Proposition. (problème de Dirichlet)
Soit $f \in \mathrm{L}^2(0,1)$ ($I =]0,1[$). On cherche à résoudre $-u'' + u = f$ avec les conditions aux limites $u(0) = u(1) = 0$. On dit que $u \in \mathrm{H}_0^1$ est une *solution faible* de ce problème si pour toute fonction test φ :

$$\int_I u'\varphi' + \int_I u\varphi = \int_I f\varphi.$$

Ce problème admet une unique solution faible.

Théorème. (problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

4 Parties compactes d'un espace vectoriel normé

4.1 Critères de compacité

Théorème. La boule unité d'un espace vectoriel normé est compacte si et seulement si celui-ci est de dimension finie.

Définition. Soit $H \subset \mathcal{C}^0(X, E)$, E métrique. On dit que H est *équicontinue* en $x_0 \in X$ lorsque

$$\forall \varepsilon > 0, \exists \eta > 0, (d(x, x_0) \leq \eta \Rightarrow d_E(f(x), f(x_0)) \leq \varepsilon).$$

H est équicontinue sur X si elle est en tout point de X.

Théorème. (d'Ascoli)
Soient (X, d) et (E, d_E) deux espaces métriques, avec X compact. Alors $H \subset \mathcal{C}(X, E)$ est relativement compact si et seulement si

— H est équicontinue sur X

— Pour tout $x \in X$, $H_x := \{f(x), f \in H\}$ est relativement compact dans H.

Remarque. Sous la première hypothèse et dans le cas où X est connexe et E normé de dimension finie, on peut seulement vérifier la deuxième hypothèse pour un x_0.

Application. Soit F un sous-espace vectoriel fermé de $\mathcal{C}^0([0,1])$, que l'on suppose inclus dans $(\mathcal{C}^1([0,1]), \|\cdot\|_{\mathcal{C}^1})$. Alors F est de dimension finie.

4.2 Une autre forme de compacité dans un espace de Hilbert

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaire au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}^1_0(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 41

209. Approximation d'une fonction par des fonctions régulières. Exemples d'applications.

Sommaire

1 Approximation de fonctions continues sur un compact 285
1.1 Densité d'une sous-algèbre 285
1.2 Densité d'un sous-espace vectoriel 286
2 Approximation dans les espaces de Lebesgue 286
2.1 Théorèmes généraux 286
2.2 Convolution sur la droite réelle, transformée de Fourier 287
2.3 Espaces de Lebesgue périodiques et séries de Fourier 288

Présentation du plan

1 Approximation de fonctions continues sur un compact

1.1 Densité d'une sous-algèbre

Théorème. (Stone-Weierstrass)
Soit A une sous-algèbre de $\mathcal{C}^0(X, \mathbb{C})$, où X est compact. On suppose que A sépare les points. Alors A est dense dans $\mathcal{C}^0(X, \mathbb{C})$ pour la norme uniforme. Pour $\mathbb{C}$, il faut ajouter la stabilité par conjugaison.

Corollaire. Les polynômes et les polynômes trigonométriques sont denses dans $\mathcal{C}^0([a, b], \mathbb{R})$ et $\mathcal{C}^0([a, b], \mathbb{C})$ respectivement.

Proposition. (théorème de Fejér)
On retrouver directement le corollaire précédent pour les polynômes trigonométriques. Si K_N est le noyau de Fejér (défini ci-après), alors $\|K_N * f - f\|_\infty \to 0$.

Théorème. (théorème taubérien fort)
Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1. On suppose que la somme admet une limite en 1 par valeurs inférieures. Si $a_n = \mathrm{O}(1/n)$, alors cette limite est la somme des (a_n).

1.2 Densité d'un sous-espace vectoriel

Théorème. (Hahn-Banach)
Soit E un espace vectoriel normé, et G un sous-espace vectoriel de E. Soit $g \in G'$. Il existe $f \in E'$ qui prolonge g telle que $\|g\|_{G'} = \|f\|_{E'}$.

Corollaire. Soit G un sous-espace vectoriel de E. Alors G est dense dans E si et seulement si toute forme linéaire continue qui s'annule sur G est nulle.

Théorème. (Riesz-Markov, ADMIS)
Soit Λ une forme linéaire continue sur $\mathcal{C}^0(X, \mathbb{R})$ (X compact). Il existe une mesure signée μ définie sur $(X, \mathcal{B}(X))$ telle que pour tout $f \in \mathcal{C}^0([0,1], \mathbb{R})$, $\Lambda(f) = \int_X f \,\mathrm{d}\mu$.

Définition. Soit $n \in \mathbb{N}^*$ et $\sigma : \mathbb{R} \to [0,1]$ continue. Un *réseau de neurones à une couche cachée* sur $[0,1]^n$ est une fonction de la forme :

$$G_{N,\theta} : x \mapsto \sum_{k=1}^{N} \alpha_k \, \sigma(\langle y_k, x \rangle + \beta_k),$$

où $\theta = ((\alpha_k)_k, (y_k)_k, (\beta_k)_k) \in \mathbb{R}^N \times (\mathbb{R}^k)^N \times \mathbb{R}^N =: \Theta$. On note $\mathcal{NN}_n(\sigma) = \{G_{N,\theta}, N \in \mathbb{N}^*, \theta \in \Theta\} \subset \mathcal{C}([0,1]^n)$.

Théorème. (d'approximation universelle, Cybenko)
Soit σ une fonction continue de $\mathbb{R}$ dans $[0,1]$. On suppose que σ est *sigmoïdale*, *i.e.* σ est croissante et $\sigma(-\infty) = 0$ et $\sigma(+\infty) = 1$. Alors $\mathcal{NN}_n(\sigma)$ est dense dans $(\mathcal{C}([0,1]^n), \|\cdot\|_\infty)$.

2 Approximation dans les espaces de Lebesgue

2.1 Théorèmes généraux

Théorème. Soit $p \in [1, \infty]$. Les fonctions étagées sont denses dans $\mathrm{L}^p(\Omega, \mathcal{A}, \mu)$.

Remarque. Ce résultat est surtout utile en probabilités, pour montrer le lemme de Doob par exemple.

Théorème. (Réciproque de Lebesgue)
Soit $(f_n)_n$ une suite de fonctions L^p qui converge dans L^p vers f. Alors à extraction près, elle converge presque partout vers f.

2.2 Convolution sur la droite réelle, transformée de Fourier

Définition. Soient f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{C}$. On dit qu'elles sont *convoluables* lorsque $t \mapsto f(x-t)g(t)$ est intégrable pour presque tout x. On définit alors le *produit de convolution* de f et g par la formule :

$$f * g(x) = \int_{\mathbb{R}} f(x-t)g(t) \,\mathrm{d}t.$$

Théorème. L'espace $(\mathrm{L}^1(\mathbb{R}), +, *)$ est une algèbre de Banach commutative.

Proposition. On peut définir le produit de convolution d'une fonction intégrable f et d'une fonction test φ. Alors $f * \varphi$ est $\mathcal{C}^\infty$, avec $(f * \varphi)^{(n)} = f * \varphi^{(n)}$.

Définition. Soit $(\rho_n)_n$ une suite de fonctions positives de $\mathrm{L}^1(\mathbb{R})$. Cette suite est une *approximation de l'unité* si pour tout n, $\int_{\mathbb{R}} \rho_n = 1$ et pour tout $\delta > 0$, $\lim_{n\to\infty} \int_{|x|>\delta} \rho_n(x)\,\mathrm{d}x = 0$.

Proposition. Les translations sont continues dans L^p, $p \neq \infty$.

Exemple. Pour φ une fonction test, la suite $(n\varphi(n\cdot))_n$ définit une approximation de l'unité.

Théorème. Soit $p \in [1, +\infty[$, $k \in \mathbb{N}$ et (φ_n) une approximation de l'unité dans $(\mathcal{D}(\mathbb{R}))^{\mathbb{N}}$. Si $f \in \mathrm{L}^p(\mathbb{R})$, $\varphi_n * f \in \mathrm{L}^p(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R})$ et $\|\varphi_n * f - f\|_p \to f$.

Corollaire. Pour $p < \infty$, l'ensemble des fonctions test est dense dans $\mathrm{L}^p(\mathbb{R})$.

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}_0^0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}_0^0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un opérateur linéaire continu, et un morphisme d'algèbres de Banach de $(\mathrm{L}^1(\mathbb{R}), +, *)$ dans $(\mathcal{C}_0^0(\mathbb{R}, +, \times))$.

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Proposition. Les fonctions test sont dans la classe de Schwartz, qui est donc dense dans L^p, $p < \infty$.

Proposition. La classe de Schwartz est stable par transformée de Fourier.

Théorème. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x} \widehat{f}(\xi)\,\mathrm{d}\xi.$$

Par dualité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$.

2.3 Espaces de Lebesgue périodiques et séries de Fourier

Définition. On note $\mathrm{L}^p_{2\pi}$ l'ensemble des fonctions p-intégrables sur $[-\pi, \pi]$ et (presque partout) périodiques sur $\mathbb{R}$.

Définition. Pour $f \in \mathrm{L}^p_{2\pi}(0, 2\pi)$ et $g \in \mathcal{C}^0([0, 2\pi])$, on définit le *produit de convolution (périodique)* de f et g par :

$$f * g : x \mapsto \frac{1}{2\pi} \int_0^{2\pi} f(t) g(x - t)\,\mathrm{d}t,$$

qui définit une fonction de $\mathrm{L}^p_{2\pi}$.

Remarque. Par Cauchy-Schwarz, on peut définir le produit de convolution entre L^2 et L^2, qui est alors dans L^2, et par Fubini entre L^1 et L^1, le résultat est alors dans L^1.

Définition. Soit $(\rho_n)_n$ une suite de fonctions de $\mathrm{L}^1_{2\pi} \cap \mathcal{C}^0([0, 2\pi])$. On dit que c'est une *approximation de l'unité* lorsque $\int \rho_n = 1$ pour tout n, et pour tout $\delta > 0$, $\lim_n \int_{|x|\geq\delta} \rho_n(x)\,\mathrm{d}x = 0$.

Théorème. Soit $(\rho_n)_n$ une approximation de l'unité, et $f \in \mathrm{L}^p_{2\pi}$. Alors, on a la convergence suivante dans $\mathrm{L}^p_{2\pi}$:

$$\rho_n * f \xrightarrow[n\to\infty]{} f.$$

Définition. On note $e_k : \mathbb{R} \to \mathbb{C}$ définie par $e_k(t) = \mathrm{e}^{\mathrm{i}kt}$.
Pour une fonction $f \in \mathrm{L}^1_{2\pi}$ et $n \in \mathbb{Z}$, on définit le n-ième *coefficient de Fourier* par :

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e_{-n}(t)\,\mathrm{d}t \in \mathbb{C}.$$

Proposition. Liste de propriétés des coefficients de Fourier :

- $c_n(\tau_a f) = \mathrm{e}^{\mathrm{i}na}c_n(f)$
- $c_n(e_k f) = c_{n-k}(f)$
- Si f est $\mathcal{C}^{k-1}$ et $\mathcal{C}^k$ par morceaux, $c_n(f^{(k)}) = (\mathrm{i}n)^k c_n(f)$.

Définition. Pour $N \in \mathbb{N}^*$, on définit le *noyau de Dirichlet* :

$$D_N(t) = \sum_{n=-N}^{N} \mathrm{e}^{\mathrm{i}nt} = \frac{\sin((N+1/2)t)}{\sin(t/2)}$$

ainsi que le *noyau de Fejér* :

$$K_N(t) = \frac{1}{N}\sum_{n=0}^{N-1} D_n(t) = \frac{\sin(Nt/2)^2}{N\sin(t/2)^2}.$$

On définit les sommes partielles de Fourier par : $S_N(f) = (D_N * f)$, et leur moyenne de Cesàro s'écrit alors $(K_N * f)$.

Théorème. (de Dirichlet)
Soit $f \in \mathrm{L}^1_{2\pi}$ et $x_0 \in \mathbb{R}$. On suppose que f admet des limites à gauche et à droite en x_0, et que ses taux d'accroissements à gauche et à droite en x_0 sont intégrables au voisinage de x_0. Alors la série de Fourier de f converge en x_0 avec pour limite $\frac{f(x_0^+)+f(x_0^-)}{2}$.

Remarque. Cependant, il existe des fonctions continues dont la série de Fourier diverge en 0.

Théorème. (de Fejér)
Si $p < \infty$, alors $\|K_N * f - f\|_p \to 0$. Si $p = \infty$ et que f est continue, alors $\|K_N * f - f\|_\infty \to 0$.

Lemme. (lemme de Lebesgue)
Soit $(V, \|.\|)$ un espace vectoriel normé, U un sous-espace vectoriel de V et P un projecteur linéaire sur U. Alors pour tout $v \in V$, on a

$$\|v - P(v)\| \leq (1 + \|P\|)\inf_{u\in U}\|v - u\|.$$

Proposition. (qualité de l'approximation trigonométrique)
Soit u une fonction continue 2π-périodique. Alors

$$\|u - K_n * u\|_\infty \leq 2\omega\left(\frac{\pi}{\sqrt{n}}\right),$$

où ω est le module de continuité.

Chapitre 42

213. Espaces de Hilbert. Exemples d'applications.

Sommaire

1 Espaces préhilbertiens et euclidiens 291
1.1 Produit scalaire et norme 291
1.2 Orthogonalité 292
1.3 Cas de la dimension finie 293
2 Espaces de Hilbert 294
2.1 Premiers exemples 294
2.2 Projection sur un convexe fermé 294
2.3 Bases hilbertiennes 295
3 Analyse de Fourier 295
3.1 Séries de Fourier 295
3.2 Transformée de Fourier 296
4 Espaces de Sobolev en dimension un 297
4.1 Définition 297
4.2 Résolution de problèmes avec conditions aux limites 298

Présentation du plan

1 Espaces préhilbertiens et euclidiens

1.1 Produit scalaire et norme

Définition. Un espace vectoriel réel ou complexe est dit *préhilbertien* lorsqu'on le munit d'une forme bilinéaire (sesquilinéaire dans le cas complexe) symétrique définie positive $\langle \cdot, \cdot \rangle$, appelée alors *produit scalaire*. En posant $\|x\| = \sqrt{\langle x, x \rangle}$, on fait de E un espace vectoriel normé. Si E est de dimension finie, on parle d'espace *euclidien* ou *hermitien*.

Proposition. (inégalité de Cauchy-Schwarz)

Soit $(x, y) \in E^2$, où E est un espace préhilbertien. Alors :

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Remarque. C'est cette inégalité qui permet de montrer que l'on a bien une norme.

Exemple. $\mathbb{R}^n$ et $\mathbb{C}^n$ sont respectivement euclidien et hermitien munis du produit scalaire $\sum_{k=1}^{n} u_k\overline{v_k}$.

Exemple. Les espaces $\mathrm{L}^2(\Omega)$ et $\ell^2(\mathbb{N})$ sont des espaces préhilbertiens pour $\int f\overline{g}$ et $\sum u_n\overline{v_n}$ respectivement.

Remarque. Si on dispose d'un espace préhilbertien et de sa norme, on peut obtenir son produit scalaire grâce à une *identité de polarisation*, par exemple

$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

Proposition. (identité du parallélogramme)
Soient x et y d'un espace préhilbertien, alors :

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

1.2 Orthogonalité

Définition. Deux vecteur sont dits *orthogonaux* lorsque leur produit scalaire est nul. L'*orthogonal* d'une partie A est l'ensemble des vecteurs orthogonaux à tous les vecteurs de A, noté $A^\perp$.

Définition. Une famille (e_i) est dite *orthogonale* lorsque $\langle e_i, e_j\rangle = a_{i,j}\delta_{i,j}\forall i, j$, et *orthonormée* si $a_{i,j} = 1$.

Proposition. Soient $A, B \subset E$ deux parties d'un espace préhilbertien. Alors

- $A \subset B$ implique $B^\perp \subset A^\perp$
- $A^\perp = \mathrm{Vect}(A)^\perp$ est un sous-espace vectoriel fermé de E
- $A \subset A^{\perp\perp}$
- Si A et B sont en plus des sous-espaces vectoriels, alors $(A+B)^\perp = A^\perp \cap B^\perp$ et $F^\perp + G^\perp \subset (F \cap G)^\perp$

Proposition. (Théorème de Pythagore)
Soit E un espace préhilbertien réel et deux éléments x et y de E. Ils sont orthogonaux si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Proposition. Soit F un sous-espace vectoriel de E préhilbertien. On a toujours $F \cap F^\perp = \{0\}$. Si de plus on a une somme directe, on dit que $F^\perp$ est le supplémentaire orthogonal de F, et on a alors $(F^\perp)^\perp = F$.

Exemple. Dans $\mathbb{R}[X]$ muni de $\int_0^1 PQ$, $X\mathbb{R}[X]$ n'a pas de supplémentaire orthogonal.

Proposition. Soit F un sous-espace vectoriel de E préhilbertien réel. Alors F admet un supplémentaire orthogonal si et seulement si la distance à F est toujours atteinte. Dans ce cas, le vecteur qui réalise la distance de x à F est unique et c'est la projection orthogonale de x sur F.

1.3 Cas de la dimension finie

Proposition. (Gram-Schmidt)
Soit E un espace hermitien ou euclidien. Alors E admet une base orthonormée.

Théorème. Soit F est un sous-espace vectoriel de dimension finie de E préhilbertien réel, on a toujours $E = F \oplus F^\perp$. De plus, si $(e_1, \dots, e_n)$ est une bas orthonormée de F, alors pour tout $x \in E$:

$$p_F(x) = \sum_{k=1}^{n} \langle e_k, x \rangle e_k \text{ et } d(x, F) = \|p_{F^\perp}(x)\|.$$

Définition. On appelle *matrice de Gram* de la famille $(x_1, \dots, x_n) \in E^n$ la matrice $(\langle x_i, x_j \rangle)_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R})$, et son déterminant le *déterminant de Gram* de la famille, noté $G(x_1, \dots, x_n)$. Une matrice est une matrice de Gram si et seulement si elle est symétrique positive. Une matrice de Gram est définie positive si et seulement si la famille en question est liée.

Théorème. Soit F un sous-espace vectoriel de E qui admet une base $(e_1, \dots, e_k)$. Pour $x \in E$, on a :

$$d(x, F)^2 = \frac{G(x, e_1, \dots, e_k)}{G(e_1, \dots, e_k)}.$$

Application. L'expression $\int_0^1 (1 + a_1 x + \ldots + a_n x^n)^2 \, \mathrm{d}x$ atteint son minimum $\mu = \frac{1}{n+1}$ en un unique n-uplet.

2 Espaces de Hilbert

2.1 Premiers exemples

Définition. Un espace *de Hilbert* est un espace préhilbertien pour lequel la norme le rend complet.

Exemple. Les espaces euclidiens et hermitiens sont des espaces de Hilbert, et $\mathrm{L}^2(\Omega, \mathcal{F}, \mu)$ et $\ell^2(\mathbb{Z})$ également.

Remarque. Un sous-espace vectoriel d'un Hilbert en est un si et seulement si il est fermé. En particulier, si $\mathcal{G}$ est une sous-tribu de $\mathcal{F}$, alors $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ est un sous-espace vectoriel fermé de $\mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

2.2 Projection sur un convexe fermé

Théorème. (projection sur un convexe fermé)
Soit H un espace de Hilbert et C un convexe fermé non vide de H.

- Pour tout $x \in H$, il existe un unique point $p_C(x) \in C$ qui réalise la distance à C
- Il est caractérisé par la propriété suivante :

$$p_C(x) \in C \text{ et } \mathrm{Re}(\langle x - p_C(x), y - p_C(x) \rangle) \forall y \in C$$

- p_C est une application 1-lipschitzienne

— Si C est un sous-espace vectoriel fermé de H, alors $p_C(x)$ est caractérisé par

$$p_C(x) \in C \text{ et } (x - p_C(x)) \perp F.$$

Remarque. Seule la complétude de C est nécessaire dans la preuve.

Proposition. L'espérance conditionnelle de $X \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ sachant la tribu $\mathcal{G}$ est définie comme la projection de X sur le convexe fermé $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ et vérifie alors, pour tout Y $\mathcal{G}$-mesurable :

$$\mathbb{E}(XY) = \mathbb{E}(Y\,\mathbb{E}(X|\mathcal{G})).$$

Théorème. (supplémentaire orthogonal)
Soit F un sous-espace vectoriel fermé de H un Hilbert. Alors H est somme directe orthogonale de F et $F^\perp$.

Corollaire. Un sous-espace vectoriel F est dense dans H si et seulement si $F^\perp = 0$.

Théorème. (représentation de Riesz)
Pour tout $\varphi \in H'$, où H est un espace de Hilbert, il existe un unique a dans H tel que $\varphi = \langle \cdot, a \rangle$.

2.3 Bases hilbertiennes

Définition. Une famille de vecteurs de H Hilbert est une *base hilbertienne* si c'est une famille orthonormée totale.

Théorème. (Bessel-Parseval)
Soit E préhilbertien, et $(e_i)_{i\in I}$ orthonormée. Pour tout $x \in E$, la famille $(\langle e_i, x\rangle^2)_i$ est sommable, avec

$$\sum_{i\in I} |\langle e_i, x\rangle|^2 \leq \|x\|^2,$$

avec égalité si et seulement si la famille est une base hilbertienne.

Remarque. Un espace préhilbertien est séprable si et seulement si il admet une base hilbertienne dénombrable, donc on peut se ramener à des sommes sur $\mathbb{N}$ (tous les Hilbert sur lesquels on va travailler sont séparables).

Théorème. Soit E un espace préhilbertien et $(e_n)_n$ une base hilbertienne de E.

— $x \mapsto (\langle x, e_n\rangle)_n$ est une isométrie de E sur $\ell^2(\mathbb{N}, \mathbb{K})$.

— Pour tout $x \in E$, $x = \sum_{n=0}^{\infty} \langle x, e_n\rangle e_n$ et cette série converge commutativement.

— L'isométrie précédente est surjective si et seulement si E est un espace de Hilbert.

3 Analyse de Fourier

3.1 Séries de Fourier

Définition. On note $\mathrm{L}^2_{2\pi}$ l'ensemble des fonctions intégrables sur $[-\pi, \pi]$ et (presque partout) périodiques sur $\mathbb{R}$.

Définition. On note $e_k : \mathbb{R} \to \mathbb{C}$ définie par $e_k(t) = \mathrm{e}^{\mathrm{i}kt}$.
Pour une fonction $f \in \mathrm{L}^2_{2\pi}$ et $n \in \mathbb{Z}$, on définit le n-ième *coefficient de Fourier* par :

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e_{-n}(t) \, \mathrm{d}t \in \mathbb{C}.$$

Remarque. Pour $f \in \mathrm{L}^2_{2\pi}$, on a $c_n(f) = \langle f, e_n \rangle$ (en prenant par convention l'antilinéarité à droite).

Exemple. Coefficient de Fourier d'une porte et $c_n(e_k) = \delta_{k,n}$. Coefficient de Fourier de $\cos(\alpha t)$.

Proposition. Liste de propriétés des coefficients de Fourier :

— $c_n(\tau_a f) = \mathrm{e}^{\mathrm{i}na} c_n(f)$
— $c_n(e_k f) = c_{n-k}(f)$
— Si f est $\mathcal{C}^{k-1}$ et $\mathcal{C}^k$ par morceaux, $c_n(f^{(k)}) = (\mathrm{i}n)^k c_n(f)$.

Définition. On définit la *somme partielle de Fourier* par $S_N(f) \sum_{n=-N}^{N} c_n(f) e_n$ et la série de Fourier (convergente ou non, c'est une série de fonctions) $S(f) = \sum_{n \in \mathbb{Z}} c_n(f) e_n$.

Théorème. La famille $(e_k)_k$ est une base hilbertienne de $\mathrm{L}^2_{2\pi}$.

Corollaire. On a donc l'égalité de Bessel et la formule de Parseval, la série de Fourier converge dans L^2 et on a une isométrie linéaire bijective.

Exemple. $\zeta(4) = \frac{\pi^4}{90}$.

3.2 Transformée de Fourier

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x) \, \mathrm{d}x, \quad \xi \in \mathbb{R}$$

Exemple. Si $g : x \mapsto \mathrm{e}^{-\alpha x}$, pour $\alpha > 0$, alors $\widehat{g} : \xi \mapsto \sqrt{\frac{\pi}{\alpha}} \mathrm{e}^{-\frac{\xi^2}{4\alpha}}$.

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi| \to \infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}^0_0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}^0_0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un opérateur linéaire continu.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

— $\mathcal{F}(\check{f}) = \mathcal{F}\check{(}f)$.
— $\mathcal{F}(\overline{f}) = \overline{\mathcal{F}\check{(}f)}$.
— Pour $\lambda > 0$, $\mathcal{F}(f(\lambda \cdot)) = \frac{1}{\lambda} \mathcal{F}(f) \left(\frac{\cdot}{\lambda}\right)$
— $\mathcal{F}(\tau_a f) = \mathrm{e}^{-\mathrm{i}a\cdot} \mathcal{F}(f)$ et $\mathcal{F}(\mathrm{e}^{-\mathrm{i}a\cdot}) = \tau_a(\mathcal{F}(f))$, $a \in \mathbb{R}$.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

— Si f est $\mathcal{C}^1$ sur $\mathbb{R}$, avec $f' \in \mathrm{L}^1$. Alors $\mathcal{F}(f') = \mathrm{i}\xi\mathcal{F}(f)$.

— Si xf est intégrable, alors $\mathcal{F}(f)$ est dérivable sur $\mathbb{R}$ avec $\mathcal{F}(f)' = -\mathrm{i}\mathcal{F}(xf)$.

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Théorème. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x} \widehat{f}(\xi)\,\mathrm{d}\xi.$$

Par densité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$.

Théorème. (Plancherel) L'application $\frac{1}{\sqrt{2\pi}}\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ se prolonge de manière unique en une isométrie linéaire bijective. De plus, la relation d'inversion de Fourier est satisfaite presque partout.

Exemple. La transformée de Fourier du sinus cardinal est $\pi\mathbf{1}_{[-\pi,\pi]}$.

Remarque. En général, on n'a pas de formule explicite pour la transformée de Fourier dans L^2. Cependant, si la suite de fonctions $\xi \mapsto \int_{-n}^{n} f(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x$ admet une limite presque partout, c'est $\mathcal{F}(f)$ presque partout.

Remarque. Le premier opérateur de Fourier défini est injectif mais non surjectif.

Théorème. Il existe une unique famille (P_n) de polynômes échelonnée en degré telle que $\frac{\mathrm{d}^n}{\mathrm{d}x^n}(\mathrm{e}^{-x^2}) = (-2)^n P_n(x)\mathrm{e}^{-x^2}$. La famille $(\psi_n := \frac{1}{c_n}P_n\mathrm{e}^{-\frac{\cdot^2}{2}})$, où $c_n^2 = \frac{n!}{2^n}\sqrt{\pi}$, est une base hilbertienne de $\mathrm{L}^2(\mathbb{R})$ de vecteurs propres de $\mathcal{F}$ (en tant qu'opérateur L^2).

4 Espaces de Sobolev en dimension un

Soit I un intervalle ouvert de $\mathbb{R}$

4.1 Définition

Définition. On définit l'*espace de Sobolev* $\mathrm{H}^1(I)$ par

$$\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I); \exists g \in \mathrm{L}^p(I), \int_I u\varphi' = -\int_I g\varphi, \forall \varphi \in \mathcal{C}^1_{\mathrm{c}}(I) \right\}.$$

Pour $u \in \mathrm{H}^1(I)$, le g de la définition est unique et on le note u', car il coïncide avec la dérivée usuelle lorsque u est $\mathcal{C}^1$.

Théorème. Muni de la norme $\sqrt{\|\cdot\|_2^2 + \|\cdot'\|_2^2}$, $\mathrm{H}^1(I)$ est un espace de Hilbert séparable. De plus, si $u \in \mathrm{H}^1(I)$, alors u admet un représentant continu.

Définition. On définit $\mathrm{H}^1_0(I)$ l'adhérence des fonctions test dans $\mathrm{H}^1(I)$. C'est un espace de Hilbert séparable muni de la norme induite.

Proposition. Si $I =]a, b[$, alors $u \in \mathrm{H}^1_0(I)$ si et seulement si $u(a) = u(b) = 0$.

Théorème. (Inégalité de Poincaré)
Supposons ici que I est borné. Il existe une constante $C > 0$ telle que pour $u \in \mathrm{H}^1_0(I)$, $\|u\|_{\mathrm{H}^1} \leq C\|u'\|_2$.

4.2 Résolution de problèmes avec conditions aux limites

Proposition. (Problème de Dirichlet)
Soit $f \in \mathrm{L}^2(0,1)$ ($I =]0,1[$). On cherche à résoudre $-u'' + u = f$ avec les conditions aux limites $u(0) = u(1) = 0$. On dit que $u \in \mathrm{H}^1_0$ est une *solution faible* de ce problème si pour toute fonction test φ :

$$\int_I u'\varphi' + \int_I u\varphi = \int_I f\varphi.$$

Ce problème admet une unique solution faible.

Théorème. (Problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}^1_0(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 43

214. Théorème d'inversion locale, théorème des fonctions implicites. illustrations en analyse et en géométrie.

Sommaire

1 Le théorème d'inversion locale 300
1.1 Difféomorphismes et inversion locale 300
1.2 Analyse matricielle 300
1.3 Inversion globale 301
2 Théorème des fonctions implicites 302
2.1 1. Le théorème 302
2.2 L'équation de Burgers 302
3 Sous-variétés 302
3.1 Définition équivalentes 302
3.2 Espace tangent 303
3.3 Optimisation sous contrainte 304
3.4 Sous-variétés matricielles 304

Présentation du plan

On considère E, F, G trois $\mathbb{R}$ ou $\mathbb{C}$-espaces vectoriels normés de dimension finie.

1 Le théorème d'inversion locale

1.1 Difféomorphismes et inversion locale

Proposition. Soit $f : \Omega_E \to \Omega_F$ différentiable en a et bijective, avec f^{-1} différentiable en $f(a)$, alors $\mathrm{d}f(a)$ est bijective avec

$$\mathrm{d}(f^{-1})(f(a)) = \mathrm{d}f(a)^{-1}.$$

Définition. Une application $f : \Omega_E \to \Omega_F$ est un $\mathcal{C}^k$-difféomorphisme si elle est bijective, $\mathcal{C}^k$ et que f^{-1} l'est aussi.

Proposition. Si $f \in \mathcal{C}^k(\Omega_E, \Omega_F)$ est un $\mathcal{C}^1(\Omega_E, \Omega_F)$-difféomorphisme, alors c'est un $\mathcal{C}^k(\Omega_E, \Omega_F)$-difféomorphisme.

Théorème. Soit $\varphi = U \to V$ un $\mathcal{C}^1$-difféomorphisme entre deux ouverts de $\mathbb{R}^d$. Alors pour toute fonction mesurable $f : V \to \mathbb{C}$, la composée $f \circ \varphi$ est aussi mesurable et si f est intégrable, $f \circ \varphi$ aussi, avec la relation :

$$\int_V f(y)\,\mathrm{d}y = \int_U f(\varphi(x))|\mathrm{Jac}(\varphi)(x)|\,\mathrm{d}x.$$

Exemple. (Coordonnées polaires)
Si on considère $\varphi : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$, alors on définit un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}_+^* \times]-\pi, \pi[$ dans $\mathbb{R}^2 \backslash \mathbb{R}_-$, et on a, pour f Lebesgue intégrable sur $\mathbb{R}^2$:

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \int_0^{2\pi} f(r\cos(\theta), r\sin(\theta))\, r\mathrm{d}r\,\mathrm{d}\theta.$$

Application. L'intégrale $\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x$ vaut $\sqrt{\pi}$.

Théorème. (d'inversion locale)
Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction $\mathcal{C}^1$ et $a \in \Omega$ tel que le jacobien de f en a est non nul. Alors il existe un voisinage ouvert de $f(a)$ et un voisinage ouvert de a tel que f soit un $\mathcal{C}^1$-difféomorphisme entre ces deux voisinages.

Exemple. La fonction $x \mapsto x^3$ sur $\mathbb{R}$ fournit un contre-exemple quand la différentielle est non inversible.

Exemple. La fonction $x \mapsto x + x^2\sin(\frac{\pi}{x})$ (qui s'écrase à MORT) n'est pas $\mathcal{C}^1$ sur $\mathbb{R}$ et ne réalise pas de difféomorphisme au voisinage de zéro, même si sa dérivée y est non nulle.

1.2 Analyse matricielle

Proposition. L'exponentielle sur $\mathcal{M}_n(\mathbb{C})$ est $\mathcal{C}^1$ et définit un difféomorphisme au voisinage de 0. Son inverse locale est définie par la série du logarithme.

Corollaire. L'exponentielle matricielle de $\mathcal{M}_n(\mathbb{C})$ dans $\mathrm{GL}_n(\mathbb{C})$ est surjective non injective. Son image est l'ensemble des carrés des matrices.

Proposition. (Changement de base $\mathcal{C}^1$)

Soit A_0 une matrice symétrique réelle inversible. Alors il existe un voisinage V de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ telle que

$$A = \rho(A)^\top A_0 \rho(A) \qquad \forall A \in V.$$

Corollaire. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle

de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

1.3 Inversion globale

Théorème. (d'inversion globale)
On suppose les hypothèses de l'inversion locale vérifiées pour tout a avec de plus f injective. Alors f réalise un $\mathcal{C}^1$-difféomorphisme de Ω dans son image.

Remarque. Ce théorème n'est pas pratique car montrer l'injectivité revient à calculer l'inverse.

Théorème. (Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors f est un $\mathcal{C}^1$-difféomorphisme sur $\mathbb{R}^n$ si et seulement si f est propre (l'image réciproque de tout compact est compacte) et pour tout $x \in \mathbb{R}^n$, $\mathrm{d}f(x)$ est inversible.

Application. Il existe $\varepsilon > 0$ tel que pour tout $a \in \mathbb{R}^n$, pour tout $L \in \mathrm{GL}_n(\mathbb{R})$ qui vérifie $\|a\| < \varepsilon$ et $\|L - I_n\| < \varepsilon$, alors il existe un difféomorphisme f de $\mathbb{R}^n$ tel que :

— $\forall \|x\| < 1, f(x) = L(x) + a$

— $\forall \|x\| > 2, f(x) = x$.

2 Théorème des fonctions implicites

2.1 1. Le théorème

Théorème. (Théorème des fonctions implicites)
Soit $U \subset E \times F \to G$ (U ouvert) une fonction $\mathcal{C}^1$. On suppose $f(a, b) = 0$ pour un $(a, b) \in U$. On suppose de plus que la différentielle partielle $\mathrm{d}_y f(a, b)$ (c'est-à-dire la différentielle selon la variable sur F) est bijective de F sur G. Alors il existe

— V un voisinage ouvert de a dans E

— W un voisinage ouvert de b dans F avec $V \times W \subset U$

— $\varphi \in \mathcal{C}^1(V, W)$

tels que $(x \in V, y \in W, f(x, y) = 0)$ si et seulement si $(x \in V, y = \varphi(x))$. Si de plus f est $\mathcal{C}^k$ alors φ l'est aussi.

Remarque. La différentielle de φ s'obtient en différentiant la relation $f(x, \varphi(x)) = 0$ sur V :

$$\mathrm{d}\varphi(x) = -(\mathrm{d}_y f(x, \varphi(x)))^{-1} \circ \mathrm{d}_x f(x, \varphi(x)).$$

Exemple. L'ensemble des polynômes admettant n racines simples est un ouvert de $\mathbb{R}_n[X]$.

2.2 L'équation de Burgers

Définition. Soient a et f deux fonctions $\mathcal{C}^1(\mathbb{R},\mathbb{R})$. On cherche à résoudre en u l'équation de Burgers :

$$a(u)\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}=0,$$

avec la donnée initiale $u(x,0)=f(x)$.

Proposition. Pour tout $x_0\in\mathbb{R}$, l'équation de Burgers admet une solution locale définie sur un voisinage de $(x_0,0)$.

Exemple. Pour $a(u)=u$ et $f(x)=1-x$, la solution u est définie par $u(x,y)=\frac{1-x}{1-y}$ sur $\mathbb{R}\times]-\infty,1]$.

3 Sous-variétés

3.1 Définition équivalentes

Soit $n\in\mathbb{N}^*, p\in\mathbb{N}^*\cup\{+\infty\}, k\in\{1,\dots,n\}, N\subset\mathbb{R}^n$.

Définition. La partie N est une *sous-variété* de $\mathbb{R}^n$ de dimension k et de classe $\mathcal{C}^p$ si pour tout $x^0\in N$, il existe W un voisinage dans $\mathbb{R}^n$ de x^0 tel que l'une des propriétés suivantes ait lieu :

— *carte locale.* Il existe φ un $\mathcal{C}^p$-difféomorphisme local de W dans $\mathbb{R}^n$ en x^0 tel que

$$\varphi(N\cap W)=\mathbb{R}^k\times\{0_{n-k}\}\cap\varphi(W).$$

— *équation.* Il existe $F\in\mathcal{C}^p(W,\mathbb{R}^{n-k})$ telle que $\mathrm{d}F(x^0)$ est surjective avec

$$N\cap W=F^{-1}(\{0\}).$$

— *graphe.* Il existe en changement de coordonnées $A\in\mathrm{GL}_n(\mathbb{R})$, $u\in\mathcal{C}^p(\mathbb{R}^k,\mathbb{R}^{n-k})$ telle que

$$N\cap W=\{A(z,u(z)),z\in\mathbb{R}^k\}\cap W.$$

— *nappe paramétrée.* Il existe U un voisinage de 0 dans $\mathbb{R}^k$, $j\in\mathcal{C}^p(U,\mathbb{R}^n)$ telle que $j(0)=x^0$, $\mathrm{d}j(0)$ est injective et

$$j:U\to N\cap W \text{ est bijective bicontinue.}$$

Proposition. Les quatre définitions ci-dessus sont équivalentes.

Exemple. L'ensemble $\{(x,y,z)\in\mathbb{R}^3, z^3+x^2z+y=0\}$ est une sous-variété de $\mathbb{R}^3$ de dimension 2 et de classe $\mathcal{C}^\infty$ en tant que graphe de

$$u:(x,y)\mapsto -z^3-x^2z.$$

Exemple. $\mathrm{SL}_n(\mathbb{R})$ est une sous-variété de $\mathcal{M}_n(\mathbb{R})$ de dimension n^2-1 en tant qu'image réciproque de 0 par l'application $\det-1$.

Exemple. De même, on montre que $\mathrm{O}_n(\mathbb{R})$ est une sous-variété de dimension $\frac{n(n-1)}{2}$.

3.2 Espace tangent

Définition. Soit N une sous-variété de $\mathbb{R}^n$ de dimension k. On appelle *espace tangent* à N en x^0 l'ensemble :

$$\mathrm{T}_{x^0}(N) = \{\gamma'(0), \gamma \in \mathcal{C}^1(I, \mathbb{R}^n), \gamma(0) = x^0, \gamma(I) \subset N\}.$$

C'est un sous-espace vectoriel de $\mathbb{R}^n$.

Proposition. Selon la définition, le plan tangent s'exprime ainsi :

- $\mathrm{T}_{x^0}(N) = \mathrm{d}\varphi(x^0)^{-1}(\mathbb{R}^k \times \{0^{n-k}\})$
- $\mathrm{T}_{x^0}(N) = \ker(\mathrm{d}F(x^0))$
- $\mathrm{T}_{x^0}(N) = \{A(h, \mathrm{d}u(x^0)h), h \in \mathbb{R}^k\}$
- $\mathrm{T}_{x^0}(N) = \mathrm{d}j(0)(\mathbb{R}^k)$.

Proposition. L'espace tangent à $\mathrm{SL}_n(\mathbb{R})$ de M est $M\ker(\mathrm{Tr})$.

Proposition. L'espace tangent à $\mathrm{O}_n(\mathbb{R})$ de I_n est $\mathcal{A}_n(\mathbb{R})$.

Proposition. Dans le plan, la réunion de deux droites sécantes n'est pas une sous-variété.

Remarque. Il est difficile en général de montrer qu'une partie n'est pas une sous-variété.

3.3 Optimisation sous contrainte

Proposition. Soit $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ et N une sous-variété de $\mathbb{R}^n$, et $x^0 \in N$. Si la restriction de f à N admet un extremum local en x^0 alors $\mathrm{T}_{x^0}(N) \subset \ker(\mathrm{d}g(x^0))$.

Théorème. (des extremums liés)
Soient $f, g_1, \ldots, g_p$ des fonctions réelles de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$, et X l'ensemble défini par les équations $g_1(x) = \ldots = g_p(x) = 0, x \in U$. Si la restriction de f à X admet un extremum local en a et si les différentielles $\mathrm{d}g_1(a), \ldots, \mathrm{d}g_p(a)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$ alors il existe des constantes $(\lambda_1, \ldots, \lambda_p)$ appelées *multiplicateurs de Lagrange* tels que $\mathrm{d}f(a) = \sum_{i=1}^{p} \lambda_i \mathrm{d}g_i(a)$.

Application. Le périmètre maximal d'un polygone inscrit dans un cercle est atteint pour un polygone régulier.

3.4 Sous-variétés matricielles

Théorème. (de Cartan et Von Neumann)
Tout sous-groupe fermé de $\mathrm{GL}_n(\mathbb{R})$ est une sous-variété de $\mathcal{M}_n(\mathbb{R})$.

Chapitre 44

215. Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.

Sommaire

1 Différentiabilité 306
1.1 Définition 306
1.2 Dérivées partielles 307
1.3 Ordre supérieur 308
2 Résultats fondamentaux 309
2.1 Inégalité des accroissements finis 309
2.2 Théorème d'inversion locale 310
3 Optimisation 311
3.1 Cadre général 311
3.2 Fonctions convexes 312
3.3 Optimisation sous contrainte 312

Présentation du plan

On considère E, F, G trois $\mathbb{R}$ ou $\mathbb{C}$-espaces vectoriels normés de dimension finie.

1 Différentiabilité

1.1 Définition

Définition. Soit Ω un ouvert de E, $a \in \Omega$ et $f : \Omega \to F$. On dit que f est *différentiable* en a s'il existe une application linéaire $L \in \mathcal{L}(E, F)$ telle que pour $\|h\| \to 0$:

$$\|f(a+h) - f(a) - L(h)\| = \mathrm{o}(\|h\|).$$

L'application L est alors unique, on l'appelle *différentielle* de f en a, notée $\mathrm{d}f(a)$.

Remarque. On peut étendre cette définition en dimension finie, et la plupart des résultats que l'on verra subsisteront, mais on demande alors que l'application L soit en plus continue.

Définition. Si f est une application différentiable en tout point de $\Omega \subset E$, on dit que f est différentiable sur Ω. Si de plus l'application $f \mapsto \mathrm{d}f$ est continue, on dit que f est de classe $\mathcal{C}^1$, noté $f \in \mathcal{C}^1(\Omega, F)$.

Exemple. Dans $\mathrm{GL}_n(\mathbb{R})$, l'inversion F est $\mathcal{C}^1$ avec

$$\mathrm{d}F(A)(H) = -A^{-1}HA^{-1} \qquad A \in \mathrm{GL}_n(\mathbb{R}), H \in \mathcal{M}_n(\mathbb{R})$$

Exemple. Si B est une application bilinéaire de $E \times F$ dans G, alors elle est différentiable avec :

$$\mathrm{d}f(a_1, a_2)(h_1, h_2) = f(a_1, h_2) + f(h_1, a_2).$$

On a un résultat similaire pour des applications n-linéaires.

Exemple. Soit f une fonction holomorphe sur un ouvert U de $\mathbb{C}$. On la voit comme une fonction de deux variables. Alors f est différentiable, mais ce n'est pas équivalent, comme le montre $z \mapsto \overline{z}$.

Proposition. Soient $f, g : \Omega \subset E \to F$ deux applications différentiables sur Ω. Alors pour $\lambda \in \mathbb{C}$, $\lambda f + g$ l'est aussi et $\mathrm{d}(\lambda f + g) = \lambda \mathrm{d}f + \mathrm{d}g$.

Proposition. (Fonctions composées)
Soit $f : \Omega_E \subset E \to \Omega_F \subset F$ et soit $g : \Omega_F \to G$. On suppose que f est différentiable en a et g est différentiable en $f(a)$. Alors $(g \circ f)$ est différentiable en a avec

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a).$$

Exemple. Soit $f : \Omega \to F$ une fonction différentiable sur $\Omega \subset E$ et $[x, y]$ un segment inclus dans Ω. On pose

$$u : t \mapsto f(x + t(y - x)), \quad t \in [0, 1]$$

alors u est dérivable sur $]0, 1[$ avec $u'(t) = \mathrm{d}f(x + t(y - x))(x - y)$.

Proposition. Soit $f : \Omega_E \to \Omega_F$ différentiable en a et bijective, avec f^{-1} différentiable en $f(a)$, alors $\mathrm{d}f(a)$ est bijective avec

$$\mathrm{d}(f^{-1})(f(a)) = \mathrm{d}f(a)^{-1}.$$

Définition. Dans le cas où $f = \mathbb{R}$, la différentielle en un point a est une forme linéaire, que l'on peut représenter par un vecteur, appelé *gradient* de f en a et noté $\nabla f(a)$.

Exemple. Pour $\det : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$, on a $\nabla(\det)(A) = \mathrm{Com}(A)$ pour toute matrice A.

1.2 Dérivées partielles

Définition. On dit que $f : \Omega \to F$ admet une dérivée en $a \in E$ dans la direction $v \in \mathbb{E}$ lorsque l'application $t \mapsto f(a + tv)$ est dérivable en 0.

Proposition. Si f est différentiable en a, elle admet des dérivées dans toutes les directions. La réciproque est fausse en général.

Exemple. La fonction $f : (x, y) \mapsto y^2 \ln(|x|)$ si $(x, y) \neq 0$ et $f(0, 0) = 0$ admet des dérivées selon tout vecteur en 0 mais n'est même pas continue en 0.

Définition. Soit $(e_1, \ldots, e_n)$ une base fixée de E (si E est un espace vectoriel classique, on prendra la base canonique). On appelle i-ème dérivée partielle de f au point a la dérivée directionnelle de f en a selon le vecteur e_i (si elle existe). On la note $\frac{\partial f}{\partial x_i}(a)$.

Exemple. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \mathrm{e}^x \cos(y)$. Alors f admet des dérivées partielles en tout point. Pour les calculer, on fixe une variable et on dérive selon l'autre :

$$\frac{\partial f}{\partial x}(x_0, y_0) = \mathrm{e}^{x_0} \cos(y_0), \frac{\partial f}{\partial y}(x_0, y_0) = -\mathrm{e}^{x_0} \sin(y_0).$$

Exemple. (équations de Cauchy-Riemann)
Soit f une fonction différentiable sur U un ouvert de $\mathbb{C}$. Alors si $f = u + \mathrm{i}v$ avec sur tout Ω,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ et } \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y},$$

alors f est holomorphe sur U.

Théorème. Soit $f : \Omega \subset E \to F$. Alors $f \in \mathcal{C}^1(\Omega, F)$ si et seulement si f admet des dérivées partielles selon toutes les variables, et que celles-ci sont continues sur tout Ω.

Définition. Soit $f = (f_1, \ldots, f_n) : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiable sur Ω. Pour $a \in \Omega$, la *matrice jacobienne* de f en a est la matrice de sa différentielle dans les bases canoniques :

$$\mathrm{J}f(a) = \left(\frac{\partial f_i}{\partial x_j}\right)_{1 \leq i \leq p, 1 \leq j \leq n},$$

et le *jacobien*, noté $\mathrm{Jac}(f)(a)$ est le déterminant de la matrice jacobienne.

1.3 Ordre supérieur

Théorème. (de Schwarz)
Soit $f : \Omega \to F$, où $\Omega \subset \mathbb{R}^n$ (pour simplifier). Soit $(j, k) \in \{1, \ldots, n\}^2$. On suppose que $\frac{\partial^2 f}{\partial x_k \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_k}$ (on dérive par rapport à une variable puis par rapport à l'autre) existent sur Ω et sont continues. Alors elles sont égales sur Ω, *i.e.* l'ordre dans lequel on dérive n'a pas d'importance.

Exemple. L'hypothèse de continuité est nécessaire, comme le montre $f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$ en $(0, 0)$.

Définition. Soit $f : \Omega \to F$, et $a \in \Omega$. On dit que f est *deux fois différentiable* en a lorsque f est différentiable sur un voisinage de a et si $f \mapsto \mathrm{d}f$ est différentiable en a. On note $\mathrm{d}^2 f(a)$ la différentielle seconde ainsi obtenue.

Exemple. Le déterminant est deux fois différentiable sur $\mathrm{GL}_n(\mathbb{R})$ avec

$$(\mathrm{d}^2(\det)(A)(H))(K) = \det(A) \operatorname{Tr}(A^{-1}K) \operatorname{Tr}(A^{-1}H) + \det(A) \operatorname{Tr}(-A^{-1}KA^{-1}H).$$

Proposition. Si f est deux fois différentiable en a alors $\mathrm{d}^2 f(a)$ s'identifie à une forme bilinéaire symétrique. De plus, f admet alors des dérivées partielles d'ordre deux et l'on a :

$$\mathrm{d}^2 f(a)(h, h') = \sum_{j,k=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(a)(h_j, h'_k) \qquad h, h' \in E.$$

Définition. On définit ainsi la *matrice hessienne* de f en a comme la matrice

$$\mathrm{H}f(a) := \left(\frac{\partial^2 f}{\partial x_j \partial x_k} \right) \in \mathcal{S}_n(\mathbb{R}).$$

Définition. On peut alors de proche en proche définir la différentielle d'ordre k, qui est une application k-linéaire symétrique. On peut définir par récurrence la classe $\mathcal{C}^k$ comme l'ensemble des fonctions différentiable dont la différentielle est dans $\mathcal{C}^{k-1}$, ainsi que la classe $\mathcal{C}^\infty$ comme l'intersection de toutes.

Proposition. Une fonction est $\mathcal{C}^k(\Omega)$ si et seulement si elle admet des dérivées partielles jusqu'à l'ordre n continues sur Ω.

Remarque. Les applications affines, multilinéaires, polynomiales sont $\mathcal{C}^\infty$. De plus, les classes $\mathcal{C}^k$ sont stables par composition et combinaison linéaire.

Définition. Une application $f : \Omega_E \to \Omega_F$ est un $\mathcal{C}^k$-difféomorphisme si elle est bijective, $\mathcal{C}^k$ et que f^{-1} l'est aussi.

Proposition. Si $f \in \mathcal{C}^1(\Omega_E, \Omega_F)$ est un $\mathcal{C}^1(\Omega_E, \Omega_F)$-difféomorphisme, alors c'est un $\mathcal{C}^k(\Omega_E, \Omega_F)$-difféomorphisme.

2 Résultats fondamentaux

2.1 Inégalité des accroissements finis

Théorème. (Inégalité des accroissements finis)
Soit $f : \Omega \to F$ différentiable, $[x, y]$ un segment inclus dans Ω. Alors :

$$\|f(x) - f(y)\|_F \leq \sup\{\|\mathrm{d}f(z)\|_{E,F}, z \in [x, y]\}\|x - y\|_E.$$

Corollaire. Avec les mêmes hypothèses, si f est de plus $\mathcal{C}^1$, alors f est localement lipschitzienne.

Corollaire. Avec les mêmes hypothèses que le théorème, si Ω est de plus connexe, alors f est constante si et seulement si sa différentielle est identiquement nulle.

2. Formules de Taylor

Théorème. (formule de Taylor-Young)
Soit $f : \Omega \to F$ une application n-fois différentiable en a. Alors :

$$f(a + h) = f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}^k f(a)(h, \ldots, h) + \mathrm{o}_0(\|h\|_H^n).$$

Théorème. Soit $\varphi = U \to V$ un $\mathcal{C}^1$-difféomorphisme entre deux ouverts de $\mathbb{R}^d$. Alors pour toute fonction mesurable $f : V \to \mathbb{C}$, la composée $f \circ \varphi$ est aussi mesurable et si f est intégrable, $f \circ \varphi$ aussi, avec la relation :

$$\int_V f(y)\,\mathrm{d}y = \int_U f(\varphi(x))|\mathrm{Jac}(\varphi)(x)|\,\mathrm{d}x.$$

Exemple. (Coordonnées polaires)
Si on considère $\varphi : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$, alors on définit un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}_+^* \times]-\pi, \pi[$ dans $\mathbb{R}^2 \backslash \mathbb{R}_-$, et on a, pour f Lebesgue intégrable sur $\mathbb{R}^2$:

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \int_0^{2\pi} f(r\cos(\theta), r\sin(\theta))\, r\mathrm{d}r\,\mathrm{d}\theta.$$

Application. L'intégrale $\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x$ vaut $\sqrt{\pi}$.

Théorème. (formule de Taylor avec reste intégral)
Soit $f \in \mathcal{C}^k(\Omega, F)$ avec $[a, a+h] \subset \Omega$ alors

$$f(a+h) = f(a) + \sum_{j=1}^{k-1} \frac{1}{j!}\mathrm{d}^k f(a)(h, \ldots, h) + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!}\mathrm{d}^k f(a+th)(h, \ldots, h)\,\mathrm{d}t.$$

2.2 Théorème d'inversion locale

Théorème. (d'inversion locale)
Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction $\mathcal{C}^1$ et $a \in \Omega$ tel que le jacobien de f en a est non nul. Alors il existe un voisinage ouvert de $f(a)$ et un voisinage ouvert de a tel que f soit un $\mathcal{C}^1$-difféomorphisme entre ces deux voisinages.

Exemple. La fonction $x \mapsto x^3$ sur $\mathbb{R}$ fournit un contre-exemple quand la différentielle est non inversible.

Exemple. La fonction $x \mapsto x + x^2\sin(\frac{\pi}{x})$ (qui s'écrase à MORT) n'est pas $\mathcal{C}^1$ sur $\mathbb{R}$ et ne réalise pas de difféomorphisme au voisinage de zéro, même si sa dérivée y est non nulle.

Exemple. L'exponentielle sur $\mathcal{M}_n(\mathbb{C})$ est $\mathcal{C}^1$ et définit un difféomorphisme au voisinage de 0. Son inverse locale est définie par la série du logarithme.

Proposition. (Changement de base $\mathcal{C}^1$)

Soit A_0 une matrice symétrique réelle inversible. Alors il existe un voisinage V de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ telle que

$$A = \rho(A)^\top A_0 \rho(A) \qquad \forall A \in V.$$

Corollaire. (Théorème des fonctions implicites)
Soit $U \subset E \times F \to G$ (U ouvert) une fonction $\mathcal{C}^1$. On suppose $f(a,b) = 0$ pour un $(a,b) \in U$. On suppose de plus que la différentielle partielle $\mathrm{d}_y f(a,b)$ (c'est-à-dire la différentielle selon la variable sur F) est bijective de F sur G. Alors il existe

— V un voisinage ouvert de a dans E

— W un voisinage ouvert de b dans F avec $V \times W \subset U$

— $\varphi \in \mathcal{C}^1(V, W)$

tels que $(x \in V, y \in W, f(x, y) = 0)$ si et seulement si $(x \in V, y = \varphi(x))$. Si de plus f est $\mathcal{C}^k$ alors φ l'est aussi.

Remarque. La différentielle de φ s'obtient en différentiant la relation $f(x, \varphi(x)) = 0$ sur V :

$$\mathrm{d}\varphi(x) = -(\mathrm{d}_y f(x, \varphi(x)))^{-1} \circ \mathrm{d}_x f(x, \varphi(x)).$$

Exemple. L'ensemble des polynômes admettant n racines simples est un ouvert de $\mathbb{R}_n[X]$.

Théorème. (d'inversion globale)
On suppose les hypothèses de l'inversion locale vérifiées pour tout a avec de plus f injective. Alors f réalise un $\mathcal{C}^1$-difféomorphisme de Ω dans son image.

Remarque. Ce théorème n'est pas pratique car montrer l'injectivité revient à calculer l'inverse.

Théorème. (Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors f est un $\mathcal{C}^1$-difféomorphisme sur $\mathbb{R}^n$ si et seulement si f est propre (l'image réciproque de tout compact est compacte) et pour tout $x \in \mathbb{R}^n$, $\mathrm{d}f(x)$ est inversible.

3 Optimisation

On considère une fonction $f : \Omega \subset E \to \mathbb{R}$

3.1 Cadre général

Définition. Un *point critique* de f est un $a \in \Omega$ tel que $\mathrm{d}f(a) = 0$.

Théorème.
— Si f est différentiable en a et admet un extremum local en a, alors a est un point critique de f.

— Si f est $\mathcal{C}^2$ admet un minimum local en a alors $\mathrm{H}f(a) \geq 0$.

— Si f est $\mathcal{C}^2$ admet un point critique en a de $\mathrm{H}f(a) > 0$ alors a est un minimum local de f.

— Si f est $\mathcal{C}^2$ admet un point critique en a et $\mathrm{H}f(a)$ est positive en un vecteur et négative en un autre, alors f n'admet pas d'extremum en a.

Exemple. La fonction $f(x, y, z) \mapsto x^2 + y^2 + z^3\mathrm{e}^x$ possède un point critique $(0, 0, 0)$ mais ce n'est pas un extremum.

Théorème. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

Application. (méthode de Laplace)
Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\|\to\infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x)\mathrm{e}^{-t\varphi(x)}\,\mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}}\left(\frac{2\pi}{t}\right)^{d/2}\mathrm{e}^{-t\varphi(x_0)}.$$

3.2 Fonctions convexes

On suppose ici Ω convexe.

Proposition. Si f est différentiable sur Ω, alors f est convexe sur Ω si et seulement si $f(y) - f(x) \geq \mathrm{d}f(x)(y-x)$ pour tout $x, y \in \Omega$. Si f est deux fois différentiable, on peut remplacer la condition par $\mathrm{d}^2 f(x) \geq 0$ pour tout $x \in \Omega$. Si f est strictement convexe, on peut remplacer les inégalités larges par des inégalités strictes.

Proposition. Si f est convexe sur Ω différentiable et si $c \in \Omega$ alors

— c est un minimum de f si et seulement si c est un point critique de f

— Si f est strictement convexe, alors elle admet au plus un minimum sur Ω.

3.3 Optimisation sous contrainte

Théorème. (des extremums liés)
Soient $f, g_1, \ldots, g_p$ des fonctions réelles de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$, et X l'ensemble défini par les équations $g_1(x) = \ldots = g_p(x) = 0, x \in U$. Si la restriction de f à X admet un extremum local en a et si les différentielles $\mathrm{d}g_1(a), \ldots, \mathrm{d}g_p(a)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$ alors il existe des constantes $(\lambda_1, \ldots, \lambda_p)$ appelées *multiplicateurs de Lagrange* tels que $\mathrm{d}f(a) = \sum_{i=1}^{p} \lambda_i \mathrm{d}g_i(a)$.

Application. Le périmètre maximal d'un polygone inscrit dans un cercle est atteint pour un polygone régulier.

Chapitre 45

218. Formules de Taylor. Exemples et applications.

Sommaire

1 Fonctions de la variable réelle 313
1.1 Formules de Taylor pour des fonctions de la variable réelle 313
1.2 Développements limités 315
1.3 Développement en série entière 316
2 Suites récurrentes 317
2.1 Résolution approchée d'équations 317
2.2 Asymptotique de suites récurrentes 317
3 En dimension supérieure 318
3.1 Formules de Taylor en dimension supérieure 318
3.2 Comportement d'une fonction au voisinage d'un extremum 318

Présentation du plan

1 Fonctions de la variable réelle

1.1 Formules de Taylor pour des fonctions de la variable réelle

Théorème. (de Rolle)
Soient $a < b$ dans $\mathbb{R}$ et $f : [a, b] \to \mathbb{R}$ continue sur l'intervalle fermé, dérivable sur l'intervalle ouvert, avec $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.

Remarque. Cela n'est plus vrai en général quand f est à valeurs complexes.

Théorème. (Égalité des accroissements finis)
Soit $f \in \mathcal{C}^0([a, b], \mathbb{R})$ dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$.

Application. (théorème de Darboux)
Soit $f : I \to \mathbb{R}$ dérivable sur I. Alors f' vérifie le théorème des valeurs intermédiaires.

Proposition. Soit $f :]a, b[\to \mathbb{R}$ dérivable. La fonction f est croissante si et seulement si $f' \geq 0$. Si $f' > 0$ alors f est strictement croissante.

Exemple. Il est possible que la dérivée s'annule mais que la fonction reste strictement croissante, comme le montre $x \mapsto x^3$.

Théorème. (formule de Taylor-Lagrange)
Soit $f \in \mathcal{C}^n([a, b], \mathbb{R})$ une fonction $(n + 1)$ fois dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b - a)^k + \frac{(b - a)^{n+1}}{(n + 1)!} f^{(n+1)}(c).$$

Théorème. (des accroissements finis)
Soit f et φ deux fonctions continues sur $[a, b]$ dérivables sur $]a, b[$, f à valeurs dans un espace vectoriel normé E et φ dans $\mathbb{R}$. On suppose que pour tout $t \in]a, b[$, $\|f'(t)\| \leq \varphi'(t)$. Alors $\|f(b) - f(a)\| \leq \varphi(b) - \varphi(a)$.

Corollaire. Soit f une fonction continue et $\mathcal{C}^1$ par morceaux sur $[a, b]$ (à valeurs dans un Banach). Alors :

$$\int_a^b f' = f(b) - f(a).$$

Théorème. (formule de Taylor-Young)
Soit $f \in \mathcal{C}^n(I, E)$, E espace vectoriel normé de dimension finie, $(n+1)$ fois dérivable en a. Alors :

$$f(a + h) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a)h^k}{k!} + \mathrm{o}(h^{n+1}).$$

Théorème. (Formule de Taylor avec reste intégral)
Soit $f \in \mathcal{C}^{n+1}([a, b], E)$, avec E un Banach. Alors

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)(b - a)^k}{k!} + \int_a^b \frac{(b - t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t.$$

Exemple. Soit $f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ telle que $f(0) = 0$. Alors la fonction définie par $h(x) = \frac{f(x)}{x}$ et prolongée par $f'(0)$ en 0 définit une fonction $\mathcal{C}^\infty$ sur $\mathbb{R}$.

Proposition. Soit f continue sur $[a, b]$ et dérivable sur $]a, b[$. Si f' admet une limite au bord, alors on peut la prolonger avec cette valeur pour en faire une fonction dérivable jusqu'au bord.

1.2 Développements limités

Définition. Soit $f : I \to \mathbb{R}$. On dit que f admet un *développement limité* à l'ordre n en $x_0 \in I$ si on a une expression de la forme :

$$f(x_0 + h) = P(h) + \mathrm{o}_0(h^n),$$

où P est un polynôme réel de degré au plus n.

Remarque. La formule de Taylor-Young montre qu'une fonction $\mathcal{C}^n$ admet un unique développement limité à l'ordre n (et donc d'avoir tous les dl usuels), mais la réciproque est fausse. En effet, $x \mapsto \mathrm{e}^{-1/x}\sin(\mathrm{e}^{1/x})$ prolongée en 0 par continuité est un $\mathrm{o}(x^n)$ en 0 pour tout n mais n'est pas $\mathcal{C}^1$. Par contre, la dérivabilité équivaut à l'existence d'un développement limité à l'ordre un, ce qui montre que le nombre dérivée est la pente de la tangente qui approxime le mieux la fonction au voisinage du point en question.

Remarque. La manipulation des développements limités est commode ; on peut les sommer, multiplier, composer etc... Ils sont un outil essentiel pour étudier une fonction localement.

Proposition. Développements limités usuels en annexe.

Exemple. Soit H_n le n-ième terme de la série harmonique. Alors $H_n = \ln(n) + \gamma + \frac{1}{2n} + \mathrm{o}(\frac{1}{n})$.

Lemme. Pour $n \in \mathbb{N}$ et $t > 0$:

$$\left|\mathrm{e}^{\mathrm{i}t} - 1 - \mathrm{i}t - \cdots - \frac{(\mathrm{i}t)^{n-1}}{(n-1)!}\right| \leq \frac{t^n}{n!}.$$

Lemme. Si X admet un moment d'ordre n alors sa fonction caractéristique φ est de classe $\mathcal{C}^n$, et

$$\varphi^{(n)}(\xi) = \mathrm{i}^n \int_{\mathbb{R}} x^n \mathrm{e}^{\mathrm{i}\xi x}\,\mathrm{d}\,\mathbb{P}_X(x).$$

Remarque. Lorsque $\xi = 0$, on retrouve les moments (à une constante près).

Corollaire. Avec les mêmes hypothèses et notations, on a le développement de Taylor suivant :

$$\left|\varphi(\xi + t) - \varphi(\xi) - \sum_{k=1}^{n-1} \frac{t^k}{k!}\varphi^{(k)}(\xi)\right| \leq M_n \frac{|t|^n}{n!}.$$

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1+\ldots+X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3 - 3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

1.3 Développement en série entière

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n x^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $]-R, R[$ et diverge grossièrement en dehors de l'intervalle fermé $[-R, R]$. On appelle R le *rayon de convergence de la série entière.*

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha, \alpha[$. La série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha, \alpha[$.

Exemple. On obtient plein de développements usuels.

Proposition. Si X admet des moments à tout ordre et si $\limsup \frac{\mathbb{E}(|X|^n)^{1/n}}{n} = \frac{1}{R} < \infty$, alors φ est développable en série entière au voisinage de tout réel avec un rayon de convergence supérieur à $\frac{R}{\mathrm{e}}$. En particulier, elle est caractérisée par ses moments.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

2 Suites récurrentes

2.1 Résolution approchée d'équations

Théorème. (Méthode de Newton)
Soit $f : [c, d] \to \mathbb{R}$ une fonction $\mathcal{C}^2$. On suppose $f(c) < 0 < f(d)$ et f' strictement positive sur $[c, d]$. On considère la suite récurrente

$$x_{n+1} = F(x_n) \text{ où } F(x) = x - \frac{f(x)}{f'(x)}.$$

Soit a l'unique zéro de f dans $[c, d]$. Alors il existe $C > 0$ tel que $|F(x) - a| \leq C|x - a|^2$ pour tout $x \in [c, d]$ et $\alpha > 0$ tel que $[a - \alpha, a + \alpha]$ est stable par F. De plus, pour tout $x_0 \in I$, $(x_n)_n$ converge vers a avec $C|x_n - a| \leq (C\alpha)^{2^n}$.

2.2 Asymptotique de suites récurrentes

Remarque. Si on a une équation du type $u_{n+1} = f(u_n)$ qui converge vers ℓ, un développement asymptotique de f au voisinage de ℓ peut permettre d'obtenir un développement asymptotique de u_n.

Exemple. Pour x_0 au voisinage de 0 et $f = \arctan$, on a un équivalent en $\sqrt{\frac{3}{2n}}$.

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

3 En dimension supérieure

3.1 Formules de Taylor en dimension supérieure

Théorème. (formule de Taylor-Young)
Soit $f : \Omega \to F$ une application n-fois différentiable en a. Alors :

$$f(a + h) = f(a) + \sum_{k=1}^{n} \frac{1}{k!}\mathrm{d}^k f(a)(h, \ldots, h) + \mathrm{o}_0(\|h\|_H^n).$$

Théorème. Soit $\varphi = U \to V$ un $\mathcal{C}^1$-difféomorphisme entre deux ouverts de $\mathbb{R}^d$. Alors pour toute fonction mesurable $f : V \to \mathbb{C}$, la composée $f \circ \varphi$ est aussi mesurable et si f est intégrable, $f \circ \varphi$ aussi, avec la relation :

$$\int_V f(y)\,\mathrm{d}y = \int_U f(\varphi(x))|\mathrm{Jac}(\varphi)(x)|\,\mathrm{d}x.$$

Exemple. (Coordonnées polaires)
Si on considère $\varphi : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$, alors on définit un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}_+^* \times]-\pi, \pi[$ dans $\mathbb{R}^2 \backslash \mathbb{R}_-$, et on a, pour f Lebesgue intégrable sur $\mathbb{R}^2$:

$$\int_{\mathbb{R}^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \int_0^{2\pi} f(r\cos(\theta), r\sin(\theta))\, r\mathrm{d}r\,\mathrm{d}\theta.$$

Application. L'intégrale $\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x$ vaut $\sqrt{\pi}$.

Théorème. (formule de Taylor avec reste intégral)
Soit $f \in \mathcal{C}^k(\Omega, F)$ avec $[a, a+h] \subset \Omega$ alors

$$f(a+h) = f(a) + \sum_{j=1}^{k-1} \frac{1}{j!} \mathrm{d}^k f(a)(h, \ldots, h) + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \mathrm{d}^k f(a+th)(h, \ldots, h)\, \mathrm{d}t.$$

3.2 Comportement d'une fonction au voisinage d'un extremum

Définition. Un *point critique* de f est un $a \in \omega$ tel que $\mathrm{d}f(a) = 0$.

Théorème.
- Si f est différentiable en a et admet un extremum local en a, alors a est un point critique de f.
- Si f est $\mathcal{C}^2$ admet un minimum local en a alors $\mathrm{H}f(a) \geq 0$.
- Si f est $\mathcal{C}^2$ admet un point critique en a de $\mathrm{H}f(a) > 0$ alors a est un minimum local de f.
- Si f est $\mathcal{C}^2$ admet un point critique en a et $\mathrm{H}f(a)$ est positive en un vecteur et négative en un autre, alors f n'admet pas d'extremum en a.

Exemple. La fonction $f(x, y, z) \mapsto x^2 + y^2 + z^3 \mathrm{e}^x$ possède un point critique $(0, 0, 0)$ mais ce n'est pas un extremum.

Proposition. (Changement de base $\mathcal{C}^1$)

Soit A_0 une matrice symétrique réelle inversible. Alors il existe un voisinage V de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ telle que

$$A = \rho(A)^\top A_0 \rho(A) \qquad \forall A \in V.$$

Théorème. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

Application. (méthode de Laplace)
Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\| \to \infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x) \mathrm{e}^{-t\varphi(x)}\, \mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}} \left(\frac{2\pi}{t}\right)^{d/2} \mathrm{e}^{-t\varphi(x_0)}.$$

Chapitre 46

219. Extremums : existence, caractérisation, recherche. Exemples et applications.

Sommaire

1 Extremums, cadre général 322
1.1 Minimum local, global 322
1.2 Compacité 322
1.3 Convexité 322
2 Calcul différentiel 323
2.1 Propriétés au premier ordre 323
2.2 Deuxième ordre et hessienne 323
2.3 Optimisation sous contraintes 324
2.4 Résolution de systèmes linéaires 324
3 Espaces de Hilbert et méthodes hilbertiennes 324
3.1 Projection sur un convexe fermé 324
3.2 Résolution de problèmes aux limites linéaires 325
3.3 Résolution de problèmes non linéaires 326

Présentation du plan

On considère E un $\mathbb{R}$-espace vectoriel normé.

1 Extremums, cadre général

1.1 Minimum local, global

Définition. Soit $f : E \to \mathbb{R}$. On dit que a est un *minimum* (global) de f si pour tout $x \in E$, $f(x) \geq f(a)$. C'est un minimum local si c'est un minimum global de la restriction de f à un voisinage de a. On définit de même un maximum. Un *extremum* est un point qui est un minimum ou un maximum.

Exemple. La fonction carré admet un unique minimum global en 0. La fonction $(x-1)^2(x+1)^2$ admet un maximum local en 0, qui n'est pas global.

Exemple. Soit A une partie de E. Pour $x \in E$, on définit la *distance* de A à x comme

$$d(x, A) = \inf_{a \in A} \|x - a\|.$$

1.2 Compacité

Proposition. Une fonction continue sur un compact est bornée est atteint ses bornes.

Corollaire. (théorème de Rolle)
Soit une fonction continue sur $[a, b]$, dérivable à l'intérieur, alors il existe $x \in]a, b[$ tel que $f'(x) = 0$.

Théorème. Soit K compact et $f : K \to K$ continue et strictement contractante. Alors f admet un unique point fixe dans K.

Définition. Une fonction $f : E \mapsto \mathbb{R}$ est dite *coercitive* lorsque f tend vers l'infini quand la norme de x tend vers l'infini.

Proposition. Si E est de dimension finie, alors une fonction coercitive admet un minimum global.

Remarque. L'hypothèse de compacité est facile à obtenir en dimension finie, mais le théorème de F.Riesz la rend trop forte en dimension infinie ; on cherchera dans la suite des hypothèses plus faibles pouvant quand même amener à l'existence d'extremums.

1.3 Convexité

Proposition. Un minimum local d'une fonction convexe est un fait un minimum global.

Proposition. Une fonction strictement convexe admet au plus un minimum.

Exemple. La réciproque est fausse, la fonction exponentielle sur $\mathbb{R}$ est strictement convexe mais n'admet pas de minimum pour autant.

Théorème. (point fixe de Kakutani)
Soit G un sous-groupe compact de $\mathrm{GL}(E)$, et K une partie convexe et compacte de E stable par G. Alors il existe u un point fixe sous G dans K.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

2 Calcul différentiel

2.1 Propriétés au premier ordre

Définition. Un *point critique* de f est un $a \in \Omega$ tel que $\mathrm{d}f(a) = 0$.

Proposition. Si f est différentiable en a et admet un extremum local en a, alors a est un point critique de f.

Exemple. La fonction $x \mapsto x^3$ admet un point critique en 0 mais ce n'est pas un extremum. La réciproque est cependant vraie dans le cas d'une fonction convexe.

Exemple. La fonction $f(x, y, z) \mapsto x^2 + y^2 + z^3 \mathrm{e}^x$ possède un point critique $(0, 0, 0)$ mais ce n'est pas un extremum.

2.2 Deuxième ordre et hessienne

On se place ici en dimension finie.

Théorème. — Si f est $\mathcal{C}^2$ admet un minimum local en a alors $\mathrm{H}f(a) \geq 0$.

— Si f est $\mathcal{C}^2$ admet un point critique en a de $\mathrm{H}f(a) > 0$ alors a est un minimum local de f.

— Si f est $\mathcal{C}^2$ admet un point critique en a et $\mathrm{H}f(a)$ est positive en un vecteur et négative en un autre, alors f n'admet pas d'extremum en a.

Théorème. (Lemme de Morse)
Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$ tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

Application. Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\| \to \infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x) \mathrm{e}^{-t\varphi(x)} \, \mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}} \left(\frac{2\pi}{t}\right)^{d/2} \mathrm{e}^{-t\varphi(x_0)}.$$

2.3 Optimisation sous contraintes

Théorème. (des extremums liés)
Soient $f, g_1, \ldots, g_p$ des fonctions réelles de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$, et X l'ensemble défini par les équations $g_1(x) = \ldots = g_p(x) = 0, x \in U$. Si la restriction de f à X admet un extremum local en a et si les différentielles $\mathrm{d}g_1(a), \ldots, \mathrm{d}g_p(a)$ sont des formes linéaires indépendantes sur $\mathbb{R}^n$ alors il existe des constantes $(\lambda_1, \ldots, \lambda_p)$ appelées *multiplicateurs de Lagrange* tels que $\mathrm{d}f(a) = \sum_{i=1}^p \lambda_i \mathrm{d}g_i(a)$.

Application. Le périmètre maximal d'un polygone inscrit dans un cercle est atteint pour un polygone régulier.

2.4 Résolution de systèmes linéaires

Proposition. Soit $A \in \mathrm{GL}_n(\mathbb{R})$ et $b \in \mathbb{R}^n$. La solution du système $Ax = b$ est l'unique minimum de la fonction $x \mapsto \|Ax - b\|_2^2$.

Remarque. Cette formulation présente un intérêt numérique, pour les méthodes de gradient.

Proposition. Dans le cas où A est rectangulaire, l'unique minimum de la fonction $x \mapsto \|Ax-b\|_2^2$ est $A^\dagger b$, avec $A^\dagger$ la *pseudo-inverse* de A, obtenue par décomposition en valeurs singulières.

3 Espaces de Hilbert et méthodes hilbertiennes

3.1 Projection sur un convexe fermé

Théorème. (projection sur un convexe fermé)
Soit H un espace de Hilbert et C un convexe fermé non vide de H.

- Pour tout $x \in H$, il existe un unique point $p_C(x) \in C$ qui réalise la distance à C
- Il est caractérisé par la propriété suivante :

$$p_C(x) \in C \text{ et } \mathrm{Re}(\langle x - p_C(x), y - p_C(x)\rangle) \forall y \in C$$

- p_C est une application 1-lipschitzienne
- Si C est un sous-espace vectoriel fermé de H, alors $p_C(x)$ est caractérisé par

$$p_C(x) \in C \text{ et } (x - p_C(x)) \perp F.$$

Remarque. Seule la complétude de C est nécessaire dans la preuve.

Proposition. L'espérance conditionnelle de $X \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ sachant la tribu $\mathcal{G}$ est définie comme la projection de X sur le convexe fermé $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ et vérifie alors, pour tout Y $\mathcal{G}$-mesurable :

$$\mathbb{E}(XY) = \mathbb{E}(Y\,\mathbb{E}(X|\mathcal{G})).$$

Théorème. (supplémentaire orthogonal)
Soit F un sous-espace vectoriel fermé de H un Hilbert. Alors H est somme directe orthogonale de F et $F^\perp$.

Corollaire. Un sous-espace vectoriel F est dense dans H si et seulement si $F^\perp = 0$.

Théorème. (représentation de Riesz)
Pour tout $\varphi \in H'$, où H est un espace de Hilbert, il existe un unique a dans H tel que $\varphi = \langle \cdot, a\rangle$.

3.2 Résolution de problèmes aux limites linéaires

Proposition. (Problème de Dirichlet)
Soit $f \in \mathrm{L}^2(0,1)$ ($I =]0,1[$). On cherche à résoudre $-u'' + u = f$ avec les conditions aux limites

$u(0) = u(1) = 0$. On dit que $u \in \mathrm{H}_0^1$ est une *solution faible* de ce problème si pour toute fonction test φ :

$$\int_I u'\varphi' + \int_I u\varphi = \int_I f\varphi.$$

Ce problème admet une unique solution faible.

Théorème. (Problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

3.3 Résolution de problèmes non linéaires

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 47

220. Illustrer par des exemples la théorie des équations différentielles ordinaires.

Sommaire

1 Théorie de Cauchy-Lipschitz 327
1.1 Existence et unicité de solutions 327
1.2 Théorème des bouts, continuité des solutions 329
2 Résolution explicite 329
2.1 Cas linéaire 329
2.2 Cas où l'on se ramène au cas linéaire 330
2.3 Décomposition de fonctions 330
3 Étude qualitative 331
3.1 Champ de vecteurs, portraits de phase 331
3.2 Cas linéaire 331
3.3 Équilibres, stabilité 331

Présentation du plan

1 Théorie de Cauchy-Lipschitz

1.1 Existence et unicité de solutions

Proposition. (lemme de Grönwall)
Soit I un intervalle de $\mathbb{R}$, $t_0 \in I$, $a \in \mathbb{R}$, u et v des fonctions réelles continues sur I avec v positive. Si pour tout $t \geq t_0$

$$u(t) \leq a + \int_{t_0}^{t} v(s)u(s)\,\mathrm{d}s$$

alors pour tout $t \geq t_0$:

$$u(t) \leq a\mathrm{e}^{\int_{t_0}^{t} v(s)\,\mathrm{d}s}.$$

Soit Ω un ouvert de $\mathbb{R} \times \mathbb{R}^N$, $f : \Omega \to \mathbb{R}^N$, on s'intéresse au problème $y' = f(t, y)$ avec $t, y \in \Omega$.

Remarque. On n'étudie pas les systèmes à l'ordre supérieur car on peut les ramener à l'ordre 1 quitte à augmenter la dimension.

Définition. La fonction f est dite *localement lipschitzienne* (par rapport à la seconde variable) lorsque pour tout $(\widetilde{t}, \widetilde{y}) \in \Omega$, il existe $T, r > 0$ et $k \geq 0$ tels que pour tout $t \in [\widetilde{t} \pm T]$, pour tous $y_1, y_2 \in \overline{B}(\widetilde{y}, r) \subset \Omega$:

$$\|f(t, y_1) - f(t, y_2)\| \leq k\|y_1 - y_2\|.$$

Proposition. Si f est $\mathcal{C}^1$, alors elle est localement lipschitzienne.

Définition. Dans le cas où $\Omega = I \times X$, on dit que f est globalement lipschitzienne (en la seconde variable) si pour tout J compact, il existe $k > 0$ tel que pour $(t, y_1), (t, y_2) \in J \times X$:

$$\|f(t, y_1) - f(t, y_2)\| \leq k\|y_1 - y_2\|.$$

Théorème. Soit $t_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^N$, $\alpha, \beta \in \mathbb{R}_+$ (qui ne sont pas tous les deux nuls), $r_0 \in \mathbb{R}_+^*$. Soit $I = [t_0 - \alpha, t_0 + \beta]$ et $f : I \times \overline{B}(y_0, r_0) \to \mathbb{R}^N$ continue et globalement lipschitzienne en y. On considère l'opérateur

$$\Phi(y)(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\,\mathrm{d}s.$$

Si $\Phi(y) \in \mathcal{C}(I, \overline{B}(y_0, r_0))$ pour un $y \in \mathcal{C}(I, \overline{B}(y_0, r_0))$ alors il existe une unique solution globale de notre problème de Cauchy.

Théorème. (de Cauchy-Lipschitz global)
Soit I un intervalle de $\mathbb{R}$, $f : I \times \mathbb{R}^N \to \mathbb{R}^N$ continue et globalement lipschitzienne en la seconde variable. Il existe une unique solution globale à notre problème de Cauchy, c'est-à-dire définie sur I tout entier.

Corollaire. (théorème de Cauchy-Lipschitz linéaire)
Soit I un intervalle de $\mathbb{R}$, $A : I \to \mathcal{M}_N(\mathbb{R})$ continue et $b : I \to \mathbb{R}^N$ continue. Pour une condition initiale donnée, le problème $y' = A(t)y + b(t)$ admet une unique solution globale.

Théorème. (de Cauchy-Lipschitz local)
Si f est localement lipschitzienne alors on a existence et unicité d'une solution maximale définie sur un intervalle ouvert.

Exemple. Pour $y' = -\sqrt{y}$ on voit que l'hypothèse Lipschitz est nécessaire. Par ailleurs, si f n'est pas globalement lipschitzienne, la solution n'est pas forcément globale, comme en témoigne $x' = x^2$.

1.2 Théorème des bouts, continuité des solutions

Théorème. (de sortie de tout compact)
Soit $f : \Omega \to \mathbb{R}^N$ continue, localement lipschitzienne. Soit $y :]c, d[\to \mathbb{R}^N$ une solution maximale. Alors $(t, y(t))$ sort de tout compact de Ω quand $t \to d$.

Corollaire. (théorème des bouts)
Soit $\Omega =]a, b[\times \mathbb{R}^n$, $f : \Omega \to \mathbb{R}^N$ continue et localement lipschitzienne. Soit y une solution maximale définie sur $]c, d[$. Si $d < b$ alors forcément $\|y(t)\|$ tend vers l'infini quand t tend vers d.

Remarque. On se sert de cela pour dire que si la solution est bornée, alors elle est globale.

Dans la suite, on rajoute un espace de paramètres $\Omega \subset \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^p$. On ajuste alors naturellement la notion de fonction localement lipschitzienne.

Théorème. Soit f continue localement lipschitzienne comme ci-dessus. Soit $(t_0, y_0, \lambda_0) \in \Omega'$. Il existe $C = [t_0 - \alpha, t_0 + \alpha] \times \overline{B}(y_0, r) \times \overline{B}(\lambda_0, \beta) \subset \Omega'$ tel que pour tout t^*, y^*, λ, on a une unique solution définie sur un voisinage de t_0 telle que $y_{t^*,y^*,\lambda}(t^*) = y^*$. De plus l'application $(t, t^*, y^*, \lambda) \mapsto y_{t^*,y^*,\lambda}(t)$ est continue.

Théorème. (de Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$ une fonction propre telle que pour tout x, $\mathrm{d}f(x)$ soit inversible. Alors f est un $\mathcal{C}^1$ difféomorphisme global.

2 Résolution explicite

2.1 Cas linéaire

On s'intéresse ici au problème linéaire présenté plus haut. L'équation homogène est celle pour $b \equiv 0$. On note $\mathcal{S}$ l'ensemble des solutions de l'équation et $\mathcal{S}_H$ l'ensemble des solutions de l'équation homogène.

Proposition. Pour $t_0 \in I$ fixé,

$$\Phi_{t_0} : \mathcal{S}_H \to \mathbb{R}^N, y \mapsto y(t_0)$$

est un isomorphisme d'espaces vectoriels. L'espace des solutions est donc un espace affine de dimension N dirigé par l'espace des solutions homogènes.

Définition. On appelle *système fondamental* une base de solutions de l'équation homogène. La matrice associée est appelée *matrice fondamentale* ou *résolvante*. Son déterminant est appelé *wronskien*.

Proposition. Une matrice est une résolvante si et seulement si elle est dérivable, inversible pour tout t et si $\Phi' = A\Phi$.

Proposition. Dans le cas où A est constante, on a $\Phi(t) = \mathrm{e}^{tA}$.

Exemple. La solution du système homogène telle que $y(t_0) = y_0$ est donnée par $y(t) = \Phi(t)\Phi(t_0)^{-1}y_0$. Exemple pratique du Berthelin.

Proposition. Si W est le wronskien associé à notre équation, alors $W' = \mathrm{Tr}(A)W$.

Théorème. Si on connaît une résolvante de notre système homogène Φ, alors la solution du système général telle que $y(t_0) = y_0$ est donnée par

$$y(t) = \Phi(t)\Phi(t_0)^{-1}y_0 + \int_{t_0}^{t} \Phi(t)\Phi(s)^{-1}b(s)\,\mathrm{d}s.$$

Exemple. Résolution de $y'' - 2y' + y = \mathrm{e}^t$ avec $y(-1) = 0$ et $y'(-1) = 1$.

2.2 Cas où l'on se ramène au cas linéaire

Exemple. Exemples d'équation de Bernoulli et de Ricatti du Berthelin.

2.3 Décomposition de fonctions

Remarque. Pour trouver des solutions, on peut les chercher développables en série entière ou en série de Fourier.

Théorème. Soit E l'ensemble des fonctions continues de (t, x) sur $\mathbb{R}_+ \times \mathbb{R}$ et 2π-périodiques en la seconde variable. Soit f une fonction $\mathcal{C}^1$ et 2π-périodique. Il existe une unique solution au problème suivant :

— $u \in E \cap \mathcal{C}^2(\mathbb{R}_+^* \times \mathbb{R})$

— $\forall t > 0, \forall x \in \mathbb{R}$, $\partial_t(u)(t, x) = \partial_x^2 u(t, x)$

— $u(0, .) = f$.

Proposition. Résolution de $xy'' + 2y' - xy$ sur $\mathbb{R}$.

3 Étude qualitative

On s'intéresse ici à des systèmes autonomes, c'est-à-dire où f ne dépend pas du temps t.

3.1 Champ de vecteurs, portraits de phase

Remarque. Tracer le portrait de phase d'un système permet d'obtenir des informations qualitatives sur son comportement.

Exemple. Plein d'exemples en annexe.

3.2 Cas linéaire

Proposition. Classification des systèmes constants d'ordre deux de Poincaré.

Exemple. Exemple du pendule linéarisé.

3.3 Équilibres, stabilité

Définition. Soit $y' = f(y)$ un système différentiel. Un *point d'équilibre* y_0 est un zéro de f. Il est dit *stable* si pour tout $\varepsilon > 0$, il existe un voisinage V de y_0 tel que toute solution du problème qui part d'un point de V reste proche de y_0 à ε près. Il est dit *asymptotiquement stable* s'il est stable et s'il existe un voisinage de y_0 tel que toutes les solutions qui partent de ce voisinage convergent vers y_0.

Proposition. (Stabilité des systèmes linéaires homogènes)
On considère le système différentiel $y' = Ay$, pour $A \in \mathcal{M}_n(\mathbb{R})$. Le point d'équilibre 0 est asymptotiquement stable si et seulement si les valeurs propres complexes de A ont toutes une partie réelle strictement négative, et il est stable si et seulement si toutes les valeurs propres ont une partie réelle négative, et que les valeurs propres de parties réelles nulles sont non défectives.

Proposition. (Linéarisation en première approche)
On considère le système différentiel $y' = f(y)$ sur $\mathbb{R}^n$, avec $f \in \mathcal{C}^1$. Soit y_0 un point d'équilibre de ce système et $A = \mathrm{Jac}(f)(y_0)$. Si toutes les valeurs propres de A ont une partie réelle strictement négative, alors y_0 est asymptotiquement stable.

Exemple. Cas de $x' = \pm x^3$.

Application. Application au pendule.

Chapitre 48

221. Équations différentielles linéaires. Systèmes d'équations différentielles linéaires. Exemples et applications.

Sommaire

1 Théorie générale 333
1.1 Problème de Cauchy linéaire 333
1.2 Espace des solutions, système fondamental, wronskien 334
1.3 Exemples 335
2 Étude qualitative 335
2.1 Une équation particulière 335
2.2 Coefficients périodiques 335
2.3 Équilibres et stabilité 335
3 Méthodes hilbertiennes et problèmes aux limites 336
3.1 Espace de Sobolev en dimension un 336
3.2 Problème de Sturm-Liouville 336

Présentation du plan

1 Théorie générale

1.1 Problème de Cauchy linéaire

Proposition. (lemme de Grönwall)
Soit I un intervalle de $\mathbb{R}$, $t_0 \in I$, $a \in \mathbb{R}$, u et v des fonctions réelles continues sur I avec v positive. Si pour tout $t \geq t_0$

$$u(t) \leq a + \int_{t_0}^{t} v(s)u(s)\,\mathrm{d}s$$

alors pour tout $t \geq t_0$:

$$u(t) \leq a\mathrm{e}^{\int_{t_0}^{t} v(s)\,\mathrm{d}s}.$$

Théorème. (théorème de Cauchy-Lipschitz linéaire)
Soit I un intervalle de $\mathbb{R}$, $A : I \to \mathcal{M}_N(\mathbb{R})$ continue et $b : I \to \mathbb{R}^N$ continue. Pour une condition initiale donnée, le problème $y' = A(t)y + b(t)$ admet une unique solution globale.

Remarque. L'existence et d'une solution globale n'est plus assurée dans le cas non linéaire.

Remarque. On peut se ramener à l'ordre un en ajoutant des dimensions pour un système réel d'ordre n.

1.2 Espace des solutions, système fondamental, wronskien

On s'intéresse ici au problème linéaire présenté plus haut. L'équation homogène est celle pour $b \equiv 0$. On note $\mathcal{S}$ l'ensemble des solutions de l'équation et $\mathcal{S}_H$ l'ensemble des solutions de l'équation homogène.

Proposition. Pour $t_0 \in I$ fixé,

$$\Phi_{t_0} : \mathcal{S}_H \to \mathbb{R}^N, y \mapsto y(t_0)$$

est un isomorphisme d'espaces vectoriels. L'espace des solutions est donc un espace affine de dimension N dirigé par l'espace des solutions homogènes.

Définition. On appelle *système fondamental* une base de solutions de l'équation homogène. La matrice associée est appelée *matrice fondamentale* ou *résolvante*. Son déterminant est appelé *wronskien*.

Proposition. Une matrice est une résolvante si et seulement si elle est dérivable, inversible pour tout t et si $\Phi' = A\Phi$.

Proposition. Dans le cas où A est constante, on a $\Phi(t) = \mathrm{e}^{tA}$.

Exemple. La solution du système homogène telle que $y(t_0) = y_0$ est donnée par $y(t) = \Phi(t)\Phi(t_0)^{-1}y_0$. Exemple pratique du Berthelin.

Proposition. Si W est le wronskien associé à notre équation, alors $W' = \mathrm{Tr}(A)W$.

1.3 Exemples

Proposition. Résolutions des équations d'ordre 1.

Exemple. Un exemple.

Proposition. Variation des constantes à l'ordre 2.

Exemple. Un exemple.

2 Étude qualitative

2.1 Une équation particulière

Proposition. On considère l'équation $y'' + qy = 0$ sur $\mathbb{R}$. Soit q une fonction continue sur $\mathbb{R}$.

- Si q est strictement négative, alors la seule solution bornée sur $\mathbb{R}$ est la fonction nulle.
- Si q est $\mathcal{C}^1$ sur $\mathbb{R}_+$ et strictement positive et strictement croissante, alors les solutions sont bornées.
- Toujours sur $\mathbb{R}_+$, si q est continue et intégrable, alors il existe une solution non bornée.

2.2 Coefficients périodiques

Exemple. Dans le cas de coefficients périodiques, il peut tout se passer.

Théorème. (théorème de Floquet)
Soit A à valeurs matricielles continue sur $\mathbb{R}$ et T périodique. Alors il existe $u \in \mathcal{C}(\mathbb{R}, \mathrm{GL}(\mathbb{C}))$ T-périodique et B une matrice complexe telles que $R(t,0) = u(t)\mathrm{e}^{tB}$ (R désigne la résolvante).

2.3 Équilibres et stabilité

Définition. Soit $y' = f(y)$ un système différentiel. Un *point d'équilibre* y_0 est un zéro de f. Il est dit *stable* si pour tout $\varepsilon > 0$, il existe un voisinage V de y_0 tel que toute solution du problème qui part d'un point de V reste proche de y_0 à ε près. Il est dit *asymptotiquement stable* s'il est stable et s'il existe un voisinage de y_0 tel que toutes les solutions qui partent de ce voisinage convergent vers y_0.

Proposition. (Stabilité des systèmes linéaires homogènes)
On considère le système différentiel $y' = Ay$, pour $A \in \mathcal{M}_n(\mathbb{R})$. Le point d'équilibre 0 est asymptotiquement stable si et seulement si les valeurs propres complexes de A ont toutes une partie réelle strictement négative, et il est stable si et seulement si toutes les valeurs propres ont une partie réelle négative, et que les valeurs propres de parties réelles nulles sont non défectives.

Proposition. (Linéarisation en première approche)
On considère le système différentiel $y' = f(y)$ sur $\mathbb{R}^n$, avec $f \in \mathcal{C}^1$. Soit y_0 un point d'équilibre de ce système et $A = \mathrm{Jac}(f)(y_0)$. Si toutes les valeurs propres de A ont une partie réelle strictement négative, alors y_0 est asymptotiquement stable.

Exemple. Cas de $x' = \pm x^3$.

Application. Application au pendule.

3 Méthodes hilbertiennes et problèmes aux limites

3.1 Espace de Sobolev en dimension un

Définition. On définit l'*espace de Sobolev* $\mathrm{H}^1(I)$ par

$$\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I); \exists g \in \mathrm{L}^p(I), \int_I u\varphi' = -\int_I g\varphi, \forall \varphi \in \mathcal{C}^1_{\mathrm{c}}(I) \right\}.$$

Pour $u \in \mathrm{H}^1(I)$, le g de la définition est unique et on le note u', car il coïncide avec la dérivée usuelle lorsque u est $\mathcal{C}^1$.

Théorème. Muni de la norme $\sqrt{\| \cdot \|_2^2 + \| \cdot' \|_2^2}$, $\mathrm{H}^1(I)$ est un espace de Hilbert séparable. De plus, si $u \in \mathrm{H}^1(I)$, alors u admet un représentant continu.

Définition. On définit $\mathrm{H}_0^1(I)$ l'adhérence des fonctions test dans $\mathrm{H}^1(I)$. C'est un espace de Hilbert séparable muni de la norme induite.

Proposition. Si $I =]a, b[$, alors $u \in \mathrm{H}_0^1(I)$ si et seulement si $u(a) = u(b) = 0$.

Théorème. (Inégalité de Poincaré)
Supposons ici que I est borné. Il existe une constante $C > 0$ telle que pour $u \in \mathrm{H}_0^1(I)$, $\|u\|_{\mathrm{H}^1} \leq C\|u'\|_2$.

3.2 Problème de Sturm-Liouville

Proposition. (Problème de Dirichlet)
Soit $f \in \mathrm{L}^2(0,1)$ ($I =]0,1[$). On cherche à résoudre $-u'' + u = f$ avec les conditions aux limites $u(0) = u(1) = 0$. On dit que $u \in \mathrm{H}_0^1$ est une *solution faible* de ce problème si pour toute fonction test φ :

$$\int_I u'\varphi' + \int_I u\varphi = \int_I f\varphi.$$

Ce problème admet une unique solution faible.

Théorème. (Problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

Chapitre 49

223. Suites réelles et complexes. Convergence, valeurs d'adhérence. Exemples et applications.

Sommaire

1 Convergence de suites numériques 340
1.1 Limite et construction de la droite réelle 340
1.2 Théorème sur les suites réelles 340
1.3 Valeurs d'adhérence 340
1.4 Asymptotique 341
2 Séries numériques 341
2.1 Série convergente, divergente, absolument convergente 341
2.2 Séries semi-convergentes 342
2.3 Étude asymptotique 342
2.4 Séries entières 343
3 Suites récurrentes 344
3.1 Convergence et asymptotique 344
3.2 Points périodiques 345
3.3 Processus de Galton-Watson 345

Présentation du plan

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1 Convergence de suites numériques

1.1 Limite et construction de la droite réelle

Définition. Soit $(u_n) \in \mathbb{K}^{\mathbb{N}}$. On dit que u_n tend (ou converge) vers $\ell \in \mathbb{K}$ si pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $|u_n - \ell| < \varepsilon$. Le scalaire ℓ est appelé *limite* de la suite (u_n) et est unique.

Proposition. La limite est linéaire, et stable par application d'une fonction continue en la limite.

Proposition. Une suite qui converge est bornée.

Définition. On dit que (u_n) est de Cauchy lorsque pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que $\forall k \in \mathbb{N}$, $|u_{N+k} - u_N| < \varepsilon$.

Proposition. Le corps $\mathbb{R}$ est défini comme le plus petit corps complet contenant $\mathbb{Q}$, en quotientant les suites de Cauchy sur $\mathbb{Q}$ par la limite de leur différence. Par construction, les corps $\mathbb{R}$ et $\mathbb{C}$ sont complets, *i.e.* toute suite de Cauchy y converge.

1.2 Théorème sur les suites réelles

Proposition. (théorème de la limite monotone)
Toute suite croissante majorée converge. Toute suite décroissante minorée converge.

Exemple. Cela permet par exemple d'établir la convergence de la suite des intégrales de Wallis.

Proposition. (théorème des gendarmes)
Si une suite est encadrée par deux suites qui convergent vers ℓ, alors cette même suite converge vers ℓ.

Proposition. (théorème des suites adjacentes)
Soient (u_n) et (v_n) deux suites réelles. On suppose que (u_n) est croissante, que (v_n) est décroissante et que $(v_n - u_n)$ tend vers 0. Alors (u_n) et (v_n) convergent vers la même limite.

1.3 Valeurs d'adhérence

Définition. Le scalaire ℓ est *valeur d'adhérence* de la suite (u_n) s'il existe une extractrice φ telle que $(u_{\varphi(n)})$.

Proposition. Une suite qui converge admet une unique valeur d'adhérence, égale à sa limite.

Théorème. (de Bolzano-Weierstrass)
Toute suite réelle ou complexe bornée admet une valeur d'adhérence.

Remarque. L'ensemble des valeurs d'adhérence d'une suite est un fermé.

Exemple. L'ensemble des valeurs d'adhérence de $((-1)^n)$ est $\{-1, 1\}$; l'ensemble des valeurs d'adhérence de $(\sin(n))_n$ est $[-1, 1]$.

Définition. Dans le cas d'une suite réelle, on définit sa *limite inférieure*, notée $\liminf$ comme sa plus petite valeur d'adhérence. On définit de même sa *limite supérieure*, notée $\limsup$ comme sa plus grande valeur d'adhérence.

Remarque. Contrairement à la limite, ces deux quantités existent toujours, et leur introduction permet parfois d'alléger les écritures.

1.4 Asymptotique

Définition. Soient (u_n) et (v_n) deux suites réelles. On dit que

- u_n est *négligeable* devant v_n, noté $u_n = \mathrm{o}(v_n)$ s'il existe $(\varepsilon_n)_n$ qui tend vers 0 telle que $u_n = \varepsilon_n v_n$
- u_n est *dominée* par v_n, noté $u_n = \mathrm{O}(v_n)$ s'il existe $(\varepsilon_n)_n$ bornée telle que $u_n = \varepsilon_n v_n$
- u_n est équivalente à v_n si $u_n - v_n = \mathrm{o}(v_n)$, on note $u_n \sim v_n$.

Remarque. Dans le cas où v_n ne s'annule pas, on se ramène à des propriétés sur le quotient.

Proposition. Les relation o et O sont transitives, et la relation d'équivalence est une relation d'équivalence.

Proposition. On peut multiplier et mettre les équivalents à une puissance constante, mais pas sommer, composer ni élever à une puissance non constante.

Exemple. La suite $(u_n := (1 + \frac{1}{n})^n)$ tend vers e quand n tend vers l'infini.

Exemple. Les intégrales de Wallis sont équivalentes à $\sqrt{\frac{\pi}{2n}}$, de quoi on peut déduire la formule de Stirling.

2 Séries numériques

2.1 Série convergente, divergente, absolument convergente

Définition. On dit que la série $\sum_n u_n$ *converge* si la suite des sommes partielles correspondante converge. Dans le cas contraire on dit qu'elle *diverge*. On dit qu'une série converge absolument si $\sum_n |u_n|$ converge. La somme d'une série est alors la limite en question.

Proposition. Si une série converge, son terme général tend vers 0.

Remarque. Une suite (u_n) converge si et seulement si la série $\sum_n (u_{n+1} - u_n)$ converge.

Exemple. La série $\sum_n \frac{1}{n(n+1)}$ converge et sa somme vaut 1.

Proposition. Une série absolument convergente converge. La réciproque est fausse.

Proposition. Soient (u_n) et (v_n) deux suites de réels positifs. Si $u_n \leq v_n$ pour tout n alors la convergence de la série des v_n entraîne celle des u_n. On en déduit le même genre de résultats pour les notations de Landau.

Théorème. (critère de Riemann)
La série $\sum n^\alpha$ converge si et seulement si $\alpha < -1$.

Proposition. La série $\sum_{n \geq 2} n^{-\alpha} (\ln(n))^{-\beta}$ converge si et seulement si $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$).

2.2 Séries semi-convergentes

Définition. Une série est dite *semi-convergente* lorsqu'elle est convergente mais pas absolument convergente.

Exemple. La série $\sum_n \frac{(-1)^n}{n}$ est semi-convergente.

Théorème. (critère spécial pour les séries alternée)
Soit (a_n) une suite positive décroissante qui tend vers 0. Alors la série $\sum_n (-1)^n a_n$ converge, et $|\sum_{k=n+1}^{\infty}(-1)^n a_n| \leq a_{n+1}$.

Exemple. La série $\sum_n \frac{(-1)^n}{n-\sqrt{n}}$ converge.

Proposition. (transformation d'Abel)
Soient (a_n) et (b_n) deux suites de complexes. On pose $A_n = \sum_{k=0}^{n} a_k$. Alors

$$\sum_{k=0}^{n} a_k b_k = a_0 b_0 - A_0 b_1 + A_n b_n + \sum_{k=1}^{n-1} A_k (b_k - b_{k+1}).$$

Application. La série $\sum_n \frac{\sin(n)}{n}$ converge.

2.3 Étude asymptotique

Proposition. (lemme de Cesàro)
Soit (u_n) une suite qui converge vers ℓ ; alors $\frac{1}{n}\sum_{k=1}^{n} u_n$ converge vers ℓ également.

Théorème. Soient (u_n) et (v_n) deux suites de réels positifs. Si les séries associées sont divergentes, on peut sommer les relations de comparaison pour les sommes partielles. Si elles sont convergentes, on peut le faire pour les restes.

Application. Soit $u_0 \in]0,1[$ et $u_{n+1} = \sin(u_n)$. Alors $u_n \sim \sqrt{\frac{3}{n}}$

Proposition. (Lemme de Kronecker)
Soient (x_n) le terme général d'une série convergente et (b_n) une suite croissante de réels qui tend vers l'infini. On a alors :

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{j=1}^{n} b_j x_j = 0.$$

On s'intéresse à des séries de la forme $\sum_n X_n$, où les (X_i) sont des variables aléatoires indépendantes et L^2. On définit par (S_n) la suite des sommes partielles, et on note $\sigma_i^2 = \mathrm{Var}(X_i)$.

Théorème. (inégalité maximale de Kolmogorov) : On suppose les (X_i) centrées. Alors pour $\varepsilon > 0$ et pour tout n :

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \sigma_i^2.$$

Proposition. Si les (X_i) sont centrées, et si la série des variances converge, alors la série $\sum_n X_n$ converge presque sûrement (et dans L^2).

Théorème. (loi des grands nombres avec un moment d'ordre deux)
Si $\mathbb{E}(X_n)$ tend vers m et si $\sum_{j=1}^{+\infty} \frac{\sigma_j^2}{j^2} < \infty$ alors $\overline{X_n}$ converge presque sûrement et dans L^2 vers m.

2.4 Séries entières

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(0,R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0,R)$. On appelle R le *rayon de convergence de la série entière* et $D(0,R)$ le *disque de convergence*.

Exemple. Soit $\mu = (\mu_n)$ une mesure de probabilité sur $\mathbb{N}$. On définit sa *série génératrice* G_μ par $G_\mu = \sum_n \mu_n z^n$. Alors le rayon de convergence de cette série est au moins 1.

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence. Elle est donc $\mathcal{C}^\infty$ sur l'intervalle de convergence. Une série entière est holomorphe sur son disque de convergence.

Exemple. La série exponentielle $\sum_n \frac{z^n}{n!}$ converge sur tout $\mathbb{C}$ vers une fonction notée exp. La fonction exp est alors holomorphe sur $\mathbb{C}$ avec $\exp' = \exp$.
La série $\sum_n z^n$ converge sur $D(0,1)$, mais pas sur tout le disque fermé. De plus, la convergence n'est pas uniforme sur le disque.

Théorème. Soit $\sum_n a_n z n^n$ une série entière. On suppose que la suite (a_n) ne s'annule pas à partir d'un certain rang, et que $|\frac{a_{n+1}}{a_n}|$ converge vers une limite ℓ. Alors le rayon de convergence R de la série entière vaut $\frac{1}{\ell}$. Plus généralement, on a la formule suivante :

$$\frac{1}{R} = \limsup_{n\to+\infty} |a_n|^{\frac{1}{n}}.$$

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha,\alpha[$. La série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha,\alpha[$.

Exemple. On obtient plein de développements usuels.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

Remarque. Sur le disque fermé de convergence, il peut tout se passer.

Proposition. Soit $\sum_n a_n z^n$ une série entière de rayon de convergence $R > 0$. On suppose que la somme $\sum_{n=0}^{+\infty} a_n R^n$ est finie. Alors la série converge uniformément sur $[0,R]$.

Exemple. Convergence de la série logarithme sur $[0,1]$, et de l'arctangente.

Théorème. (de Tauber fort) Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1. On suppose que la somme admet une limite en 1 par valeurs inférieures. Si $a_n = \mathrm{O}(1/n)$, alors cette limite est la somme des (a_n).

3 Suites récurrentes

3.1 Convergence et asymptotique

Proposition. Si $f \in \mathcal{C}^1(I)$, et $u_{n+1} = f(u_n)$. Si (u_n) converge vers ℓ, alors la convergence est lente si $|f'(\ell)| = 1$, géométrique si $|f'(\ell)| \in]0,1[$, rapide si $|f'(\ell)| = 0$.

Remarque. Puisque la fonction f est continue, si la suite (u_n) converge, c'est forcément vers un point fixe de f.

Exemple. Pour tout $x_0 > 0$, la suite $x_{n+1} = \arctan(x_n)$ tend vers 0, et la convergence est lente.

Remarque. Lorsqu'on connaît un développement limité de la fonction au voisinage de la limite de notre suite, on peut l'utiliser pour obtenir un équivalent de la suite.

Exemple. Avec la notation précédente, on a $x_n \sim \sqrt{\frac{3}{2n}}$.

3.2 Points périodiques

Définition. Un point $x \in D$ est de période p ou p-périodique lorsque :

$$f^p(x) = x \text{ et } f^k(x) \neq x \text{ pour tout } k \in \{1, \dots, p-1\}.$$

Pour $p = 1$, cela correspond à un point fixe.

Proposition. Soit g une fonction continue définie d'un segment $[a, b]$ dans lui-même. Alors g admet un point fixe.

Théorème. (de Charkovski)
Si f possède un point de période 3, alors elle possède un point à toute période.

Exemple. La fonction $x \mapsto (x + \frac{1}{2})\mathbf{1}_{[0,\frac{1}{2}]} + (4 - 2x)\mathbf{1}_{[\frac{1}{2},1]}$, aussi simple soit-elle, admet des points périodiques à toutes les périodes.

Par contre, la fonction $x \mapsto (1 - x)$ n'a que des points de période 2, et on voit que l'existence d'un point de période trois est importante pour avoir le résultat du théorème.

3.3 Processus de Galton-Watson

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

Chapitre 50

224. Exemples de développements asymptotiques de suites et de fonctions.

Sommaire

1 Développements asymptotiques, développements limités 348
1.1 Définitions 348
1.2 Formule de Taylor Young 349
2 Séries numériques 349
2.1 Critères de convergence 349
2.2 Sommation des relations de comparaison 349
2.3 Séries entières 350
3 Étude de suites récurrentes 351
3.1 Méthode générale 351
3.2 Processus de Galton-Watson 351
4 Asymptotique d'intégrales à paramètres 352
4.1 Méthode de Laplace à une dimension 352
4.2 En dimension finie quelconque 352
4.3 Théorème limite central 353

Présentation du plan

1 Développements asymptotiques, développements limités

1.1 Définitions

On donne les définitions suivantes pour des fonctions, mais elles se généralisent pour les suites à l'infini.

Définition. Soient f et g deux fonctions définies sur $A \subset \mathbb{R}$ et soit a un point d'accumulation de $\overline{A}$. On dit que

— f est *négligeable* devant g au voisinage de a et on note $f(x) = \mathrm{o}_a(g(x))$ s'il existe une fonction h définie sur un voisinage V de a qui tend vers 0 quand x tend vers a et telle que pour tout $x \in V \cap A$, $f(x) = h(x)g(x)$.

— f est *dominée* par g et on note $f(x) = \mathrm{O}_a g(x)$ si la fonction h ci-dessus et bornée.

— f et g sont équivalentes au voisinage de a si $f(x) - g(x) = \mathrm{o}(g(x))$ et on note alors $f \sim_a g$.

Remarque. Ces deux premières relations définissent des relations d'ordre partielles et la dernière une relation d'équivalence.

Définition. Une échelle de comparaison est une famille de fonctions ordonnée pour la relation o_a.

Exemple. La famille $((x-a)^k)_{k\in\mathbb{N}}$ est une échelle de comparaison au voisinage de a. La famille $(x^\lambda)_{\lambda\in\mathbb{R}}$ est une échelle de comparaison au voisinage de l'infini, tout comme $(x^\alpha \ln(x)^\beta)_{\alpha,\beta>0}$.

Définition. Un *développement asymptotique* de f définie au voisinage de $a \in \mathbb{R}$ à l'ordre $k \in \mathbb{N}$ est une expression de la forme

$$f(x) = f(a) + \lambda_1 g_1(x) + \cdots + \lambda_k g_k(x) + \mathrm{o}(g_k(x)) \qquad \lambda_i \in \mathbb{R},$$

où les (g_i) sont dans une échelle de comparaison et où $g_{i+1}(x) = \mathrm{o}_a g_i(x)$. Dans le cas où $g_i(x) = (x-a)^i$, on parle de *développement limité*.
Un développement asymptotique est unique.

Remarque. L'intérêt pratique de ces expressions et de pouvoir étudier une fonction au voisinage d'un point sans avoir besoin de connaître toute la fonction. De plus, si notre échelle de comparaison est stable par produit, on peut multiplier les développements asymptotiques des facteurs pour avoir un développement asymptotique du produit. On peut aussi composer les développements asymptotiques, comme on compose les limites.

Exemple. La formule de Taylor pour les polynômes donne un développement limite de tout polynôme au voisinage de tout point. Dans la suite on va généraliser cela à une fonction quelconque (avec des hypothèses suffisantes).

1.2 Formule de Taylor Young

Théorème. (formule de Taylor-Young)
Soit $f \in \mathcal{C}^n(I, E)$, E espace vectoriel normé de dimension finie, $(n+1)$ fois dérivable en a. Alors :

$$f(a+h) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a) h^k}{k!} + \mathrm{o}(h^{n+1}).$$

Remarque. Cette formule garantit l'existence de développements limités à tout ordre de fonctions usuelles. On peut de plus dériver et intégrer les développements limités.

Exemple. Développements limités de l'exponentielle, du logarithme, de l'arctangente.

Proposition. Une fonction est dérivable en a si et seulement si elle admet un développement limité à l'ordre 1 en a. Ce résultat tombe en défaut dès que l'ordre dépasse 2.

2 Séries numériques

Les critères de convergence et de sommation présentés ici peuvent être généralisés à l'intégrale.

2.1 Critères de convergence

Proposition. Soient (u_n) et (v_n) deux suites de réels positifs.

— Si $u_n = o(v_n)$ alors si $\sum v_n$ converge $\sum u_n$ aussi et si $\sum u_n$ diverge alors $\sum v_n$ aussi.

— de même pour les O

— Si $u_n \sim v_n$, alors $\sum u_n$ et $\sum v_n$ ont la même nature.

Remarque. Couplé au critère de Riemann, ce résultat permet de déterminer la nature de nombreuses séries à termes positifs.

Exemple. La série de terme général $(\frac{1}{1+\sin(\frac{1}{n^3})} - 1)$ converge.

2.2 Sommation des relations de comparaison

Proposition. (lemme de Cesàro)
Soit (u_n) une suite qui converge vers ℓ ; alors $\frac{1}{n}\sum_{k=1}^{n} u_n$ converge vers ℓ également.

Théorème. Soient (u_n) et (v_n) deux suites de réels positifs. Si les séries associées sont divergentes, on peut sommer les relations de comparaison pour les sommes partielles. Si elles sont convergentes, on peut le faire pour les restes.

Proposition. (Lemme de Kronecker)
Soient (x_n) le terme général d'une série convergente et (b_n) une suite croissante de réels qui tend vers l'infini. On a alors :

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{j=1}^{n} b_j x_j = 0.$$

2.3 Séries entières

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(0,R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0,R)$. On appelle R le *rayon de convergence de la série entière* et $D(0,R)$ le *disque de convergence.*

Exemple. Soit $\mu = (\mu_n)$ une mesure de probabilité sur $\mathbb{N}$. On définit sa *série génératrice* G_μ par $G_\mu = \sum_n \mu_n z^n$. Alors le rayon de convergence de cette série est au moins 1.

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence. Elle est donc $\mathcal{C}^\infty$ sur l'intervalle de convergence.

Proposition. Une série entière est holomorphe sur son disque de convergence.

Exemple. La série $\sum_n z^n$ converge sur $D(0,1)$, mais pas sur tout le disque fermé. De plus, la convergence n'est pas uniforme sur le disque.

Exemple. La série exponentielle $\sum_n \frac{z^n}{n!}$ converge sur tout $\mathbb{C}$ vers une fonction notée exp. La fonction exp est alors holomorphe sur $\mathbb{C}$ avec $\exp' = \exp$.

Théorème. Soit $\sum_n a_n z n^n$ une série entière. On suppose que la suite (a_n) ne s'annule pas à partir d'un certain rang, et que $|\frac{a_{n+1}}{a_n}|$ converge vers une limite ℓ. Alors le rayon de convergence R de la série entière vaut $\frac{1}{\ell}$. Plus généralement, on a la formule suivante :

$$\frac{1}{R} = \limsup_{n\to+\infty} |a_n|^{\frac{1}{n}}.$$

Corollaire. Le rayon de convergence d'une série entière et de sa dérivée formelle est le même.

Proposition. Si $a_n = \mathrm{o}(b_n)$, alors le rayon de convergence associé aux (b_n) est inférieur à celui des (a_n).

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha, \alpha[$. La série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha, \alpha[$.

Exemple. On obtient plein de développements usuels.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

3 Étude de suites récurrentes

3.1 Méthode générale

Remarque. Lorsqu'on cherche un développement asymptotique d'une suite de la forme $u_{n+1} - f(u_n)$, on peut s'appuyer sur l'étude de la fonction au voisinage de sa limite.

Exemple. Cas du sinus, de l'arctangente et du logarithme.

3.2 Processus de Galton-Watson

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson*.

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

4 Asymptotique d'intégrales à paramètres

4.1 Méthode de Laplace à une dimension

La méthode de Laplace est une façon de calculer des équivalents d'une intégrale à paramètre. Plus précisément, on cherche à calculer un équivalent en l'infini de fonctions de la forme :

$$F(s) = \int_a^b \mathrm{e}^{-s\varphi(t)} f(t)\,\mathrm{d}t,$$

avec φ une fonction régulière, et f continue en a qui ne s'annule pas en a.

Proposition. (Méthode de Laplace au premier ordre)
Si φ' est strictement positive sur $[a, b[$ et si f est intégrable, alors :

$$F(s) \sim \frac{1}{\varphi'(a)} \frac{\mathrm{e}^{-s\varphi(a)} f(a)}{s}.$$

Proposition. (Méthode de Laplace au second ordre)
Dans le cas où $\varphi' > 0$ sur $]a, b[$, $\varphi'(a) = 0$ et $\varphi''(a) > 0$, on a l'équivalent :

$$F(s) \sim \sqrt{\frac{\pi}{2\varphi''(a)}} \frac{\mathrm{e}^{-s\varphi(a) f(a)}}{\sqrt{s}}.$$

4.2 En dimension finie quelconque

Théorème. Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\|\to\infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x)\mathrm{e}^{-t\varphi(x)}\,\mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}}\left(\frac{2\pi}{t}\right)^{d/2}\mathrm{e}^{-t\varphi(x_0)}.$$

4.3 Théorème limite central

Lemme. Si X admet un moment d'ordre n alors sa fonction caractéristique φ est de classe $\mathcal{C}^n$, et

$$\varphi^{(n)}(\xi) = \mathrm{i}^n \int_{\mathbb{R}} x^n \mathrm{e}^{\mathrm{i}\xi x}\,\mathrm{d}\,\mathbb{P}_X(x).$$

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1+\ldots+X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3-3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

Chapitre 51

226. Suites vectorielles et réelles définies par une relation de récurrence $u_{n+1} = f(u_n)$. Exemples. Applications à la résolution approchée d'équations.

Sommaire

1 Cas d'une fonction de la variable réelle 355
1.1 Comportement des suites récurrentes 356
1.2 Résolution approchée d'équations 356
1.3 Points fixes et points périodiques 357
1.4 Processus de Galton-Watson 357
2 Cas des fonctions sur un espace de Banach 358
2.1 Résolution de systèmes linéaires 358
2.2 Cas des matrices à coefficients strictement positifs 358
2.3 Suites récurrentes linéaires 359
2.4 Théorème de point fixe 359

Présentation du plan

Dans tous le plan, on considère $f : E \mapsto E$ une fonction continue sur un espace métrique.

1 Cas d'une fonction de la variable réelle

Ici, on considère $E = I$ un intervalle de $\mathbb{R}$.

1.1 Comportement des suites récurrentes

Théorème. (de la limite monotone)
Toute suite croissante majorée converge.

Définition. Soit (u_n) une suite qui converge vers une limite ℓ et qui ne vaut jamais ℓ. On pose $v_n = \frac{|u_{n+1}-\ell|}{|u_n-\ell|}$. On suppose que la suite (v_n) tend vers un réel $a \in [0,1]$.

— Si $a = 1$, on parle de convergence *lente*.

— Si $a \in]0,1[$, on parle de convergence *géométrique*.

— Si $a = 0$, on parle de convergence *rapide*.

Exemple. Les suites $(\frac{1}{n+1})$, (2^{-n}) et $(1/n!)$ convergent respectivement lentement, géométriquement et rapidement.

Proposition. Si $f \in \mathcal{C}^1(I)$, et $u_{n+1} = f(u_n)$. Si (u_n) converge vers ℓ, alors la convergence est lente si $|f'(\ell)| = 1$, géométrique si $|f'(\ell)| \in]0,1[$, rapide si $|f'(\ell)| = 0$.

Remarque. Puisque la fonction f est continue, si la suite (u_n) converge, c'est forcément vers un point fixe de f.

Exemple. Pour tout $x_0 > 0$, la suite $x_{n+1} = \arctan(x_n)$ tend vers 0, et la convergence est lente.

Remarque. Lorsqu'on connaît un développement limité de la fonction au voisinage de la limite de notre suite, on peut l'utiliser pour obtenir un équivalent de la suite.

Exemple. Avec la notation précédente, on a $x_n \sim \sqrt{\frac{3}{2n}}$.

Proposition. (Comportement général de la suite au voisinage d'un point fixe)
On suppose ici encore $f \in \mathcal{C}^1(I)$. Soit a un point fixe de f. Si $|f'(a)| > 1$, on parle de point *répulsif*, la suite des itérées ne peut converger que si elle stationne en a. Si $|f'(a)| < 1$, on parle de point *attractif*, et dans ce cas la suite des itérées converge vers a pourvu qu'on prenne le premier terme suffisamment proche de a.

Remarque. Lorsque $|f'(a)| = 1$, il peut tout se passer, comme le montre l'exemple de exp en 0.

1.2 Résolution approchée d'équations

Théorème. (Méthode de Newton)
Soit $f : [c,d] \to \mathbb{R}$ une fonction $\mathcal{C}^2$. On suppose $f(c) < 0 < f(d)$ et f' strictement positive sur $[c,d]$. On considère la suite récurrente

$$x_{n+1} = F(x_n) \text{ où } F(x) = x - \frac{f(x)}{f'(x)}.$$

Soit a l'unique zéro de f dans $[c,d]$. Alors il existe $C > 0$ tel que $|F(x) - a| \leq C|x-a|^2$ pour tout $x \in [c,d]$ et $\alpha > 0$ tel que $[a-\alpha, a+\alpha]$ est stable par F. De plus, pour tout $x_0 \in I$, $(x_n)_n$ converge vers a avec $C|x_n - a| \leq (C\alpha)^{2^n}$.

1.3 Points fixes et points périodiques

Définition. Un point $x \in D$ est de période p ou p-périodique lorsque :

$$f^p(x) = x \text{ et } f^k(x) \neq x \text{ pour tout } k \in \{1, \ldots, p-1\}.$$

Pour $p = 1$, cela correspond à un point fixe.

Proposition. Soit g une fonction continue définie d'un segment $[a, b]$ dans lui-même. Alors g admet un point fixe.

Théorème. (de Charkovski)
Si f possède un point de période 3, alors elle possède un point à toute période.

Exemple. La fonction $x \mapsto (x+\frac{1}{2})\mathbf{1}_{[0,\frac{1}{2}]} + (4-2x)\mathbf{1}_{[\frac{1}{2},1]}$, aussi simple soit-elle, admet des points périodiques à toutes les périodes.

Par contre, la fonction $x \mapsto (1-x)$ n'a que des points de période 2, et on voit que l'existence d'un point de période trois est importante pour avoir le résultat du théorème.

1.4 Processus de Galton-Watson

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

2 Cas des fonctions sur un espace de Banach

On considère ici $(E, \|\cdot\|)$ un espace de Banach.

2.1 Résolution de systèmes linéaires

On souhaite résoudre le système linéaire $Ax = b$, $b \in \mathbb{R}^n$, A inversible.

Définition. On écrit alors $A = P - N$, P inversible, et on considère la suite récurrente

$$x_{k+1} = Bx_k + c$$

avec $x_0 \in \mathbb{R}^n$, $B = I_n - P^{-1}A$, $c = P^{-1}b$. On appelle ce genre de procédé une *méthode itérative*.

Lemme. La suite des puissances d'une matrices tend vers 0 si et seulement si son rayon spectral est strictement inférieur à 1.

Théorème. La suite (x_k) converge pour tout x_0 vers la solution du système si et seulement si $\rho(B) < 1$.

Exemple. (Méthode de Richarsdon)
Il s'agit de prendre $P = \frac{1}{\alpha} I_n$ avec $\alpha \neq 0$. Dans ce cas, on peut essayer d'ajuster le paramètre α pour avoir $\rho(B) < 1$.

2.2 Cas des matrices à coefficients strictement positifs

Définition. Soit $A \in \mathcal{M}_n(\mathbb{R})$. A est dite *positive* (resp. *strictement positive*) et on note $A \geq 0$ (resp. $A > 0$) lorsque tous ses coefficients sont positifs (resp. strictement positifs). On définit de même un vecteur positif ou strictement positif.

Proposition. Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors A est strictement positive si et seulement si :

$$\forall x \geq 0 \text{ non nul }, Ax > 0.$$

Cette propriété donne une interprétation géométrique des matrices strictement positives : ce sont celles qui préservent le cône $\{x > 0\}$ invariant. Le théorème suivant assure que dans ce cône, il y a une direction qui va être laissée invariante par A.

Théorème. (de Perron-Frobenius)
Soit A une matrice réelle strictement positive. Alors le rayon spectral de A est l'unique valeur propre de A de module $\rho(A)$. De plus, le sous espace propre associé est de dimension 1, engendré par un vecteur strictement positif.

Application. Soit $A > 0$ et $X_0 \geq 0$. On s'intéresse au système dynamique $X_{t+1} = AX_t$ pour $t \in \mathbb{N}^*$. Lorsque t tend vers l'infini, on a l'asymptotique suivante :

$$X_t \sim \rho(A)^t \langle v, X_0 \rangle w$$

où $w = \frac{x}{\|x\|_1}$ et v vecteur propre à gauche associé à $\rho(A)$ choisi de telle sorte que $\langle v, w \rangle = 1$.

2.3 Suites récurrentes linéaires

Définition. On appelle suite récurrente d'ordre p une suite de $\mathbb{C}^{\mathbb{N}}$ définie comme suit :

$$u_{n+p} = a_0 u_n + \ldots + a_{p-1} u_{n+p-1}.$$

Proposition. Si on note (U_n) le vecteur $(u_n, \ldots, u_{n+p-1})$, on a alors $U_{n+1} = CU_n$, avec C la matrice compagnon du polynôme $P := X^p + a_{p-1}X^{-1} + \ldots + a_0$.

Proposition. Si on note $((r_1, \alpha_1), \ldots, (r_q, \alpha_q))$ les racines de P et leur multiplicité, alors on a pour tout n :

$$u_n = P_1(n)r_1^n + \ldots + P_q(n)r_q^n,$$

avec $\deg(P_i) < \alpha_i$.

Exemple. Soit (a_n) définie par $a_{n+3} = 3a_{n+2} - 4a_n$, avec $a_0 = 0, a_1 = -1, a_2 = 5$. Alors $a_n = (-1)^n + (n-1)2^n$.

2.4 Théorème de point fixe

On reprend le cas E Banach.

Théorème. (de point fixe de Picard)
Soit E un espace vectoriel normé, et A une partie complète de E. Soit $f : A \to A$ une application contractante. Alors

- Il existe un unique $a \in A$ tel que $f(a) = a$,
- Pour tout x_0, la suite des itérées de x_0 par f converge géométriquement vers a.

Corollaire. (théorème de Cauchy-Lipschitz)
Soit $f : I \times \mathbb{R}^m \to \mathbb{R}^m$ une application continue globalement lipschitzienne en la seconde variable. Alors le système différentiel

$$y' = f(t, y), y(t_0) = x$$

avec $t_0 \in I$ et $x \in \mathbb{R}^m$ donnés admet une unique solution définie sur I tout entier.

Théorème. (d'inversion locale)
Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction $\mathcal{C}^1$ et $a \in \Omega$ tel que le jacobien de f en a est non nul. Alors il existe un voisinage ouvert de $f(a)$ et un voisinage ouvert de a tel que f soit un $\mathcal{C}^1$-difféomorphisme entre ces deux voisinages.

Exemple. La fonction $x \mapsto x^3$ sur $\mathbb{R}$ fournit un contre-exemple quand la différentielle est non inversible.

Exemple. La fonction $x \mapsto x + x^2 \sin(\frac{\pi}{x})$ (qui s'écrase à MORT) n'est pas $\mathcal{C}^1$ sur $\mathbb{R}$ et ne réalise pas de difféomorphisme au voisinage de zéro, même si sa dérivée y est non nulle.

Chapitre 52

228. Continuité, dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications.

Sommaire

1 Continuité d'une fonction de la variable réelle 362

1.1 Résultats généraux 362

1.2 Fonctions continues sur un segment 363

1.3 Théorème des valeurs intermédiaires 363

2 Dérivabilité 364

2.1 Définitions 364

2.2 Théorème de Rolle 364

2.3 Formules de Taylor 365

2.4 Théorème limite central 366

3 Monotonie et convexité 367

3.1 Monotonie 367

3.2 Convexité 368

3.3 Processus de Galton-Watson 368

4 Suites et séries de fonctions 369

4.1 Modes de convergence 369

4.2 Régularité de la limite 370

4.3 Séries entières 370

Présentation du plan

On désignera par I un intervalle ouvert de $\mathbb{R}$.

1 Continuité d'une fonction de la variable réelle

1.1 Résultats généraux

Définition. Soit $f : I \to \mathbb{R}$ une fonction. On dit qu'elle est *continue* en $x_0 \in I$ lorsque pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que si $|x_0 - y| \leq \delta$ alors $|f(x_0) - f(y)| \leq \varepsilon$, *i.e.* $\lim_{y \to x_0} f(x) = f(x_0)$. Si on ne regarde que la limite à droite (resp. à gauche) en x_0, on parle de continuité *à droite* (resp. à gauche). Si $I = [a, b]$, on définit la continuité en a par la continuité à droite, idem en b. Si f est continue en tout point de I, on dit qu'elle est continue sur I.

Exemple. (Quelques exemples)

- La fonction $x \mapsto 2x^2$ est continue sur $\mathbb{R}$.
- La fonction de répartition d'une variable aléatoire réelle est continue à droite sur $\mathbb{R}$.
- La fonction $\mathbf{1}_{\mathbb{Q}}$ n'est continue en aucun point de $\mathbb{R}$.
- Soit $T : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = 0$ si $x \notin \mathbb{Q}$ et $f(p/q) = 1/q$ si p/q est une fraction irréductible non nulle, $(q > 0)$. La fonction T est discontinue sur $\mathbb{Q}$ mais continue sur $\mathbb{R} \backslash \mathbb{Q}$.

Proposition. Une fonction est continue en x_0 si et seulement si pour toute suite (y_n) qui converge vers x_0, $f(x_n)$ tend vers $f(x_0)$.

Exemple. La fonction $f : x \mapsto \sin(1/x)$ sur $\mathbb{R}_+^*$ et $f(0) = 0$ n'est pas continue en 0.

Corollaire. Toute somme, tout produit, toute composée (quand c'est bien définie) de fonctions continues est continue.

Définition. Soit $f : I \mapsto \mathbb{R}$.

- f est *uniformément continue* si $\forall \varepsilon > 0, \exists \delta, \forall x \in I$, si $|y - x| \leq \delta$ alors $|f(y) - f(x)| \leq \varepsilon$.
- f est *K-lipschitzienne*, $K > 0$, si pour tout $(x, y) \in I^2$, $|f(x) - f(y)| \leq K|x - y|$.

Proposition. Une fonction uniformément continue est continue, et une fonction lipschitzienne (K-lipschitzienne pour un certain K) est uniformément continue.

Exemple. Les réciproques ne sont pas vraies en général. La fonction $x \mapsto \sqrt{x}$ est uniformément continue sur $\mathbb{R}_+$ mais pas lipschitzienne. La fonction $x \mapsto x^2$ est continue sur $\mathbb{R}_+$ mais pas uniformément.

1.2 Fonctions continues sur un segment

Théorème. Une fonction continue sur un segment est bornée et atteint ses bornes.

Théorème. (de Heine)
Une fonction uniformément continue sur un segment y est uniformément continue.

Remarque. Le théorème de Heine est utile pour définir l'intégrale de Riemann.

Théorème. (Weierstrass)
Toute fonction continue sur un segment est limite uniforme de fonctions polynomiales.

1.3 Théorème des valeurs intermédiaires

Théorème. (des valeurs intermédiaires)
L'image d'un intervalle par une fonction continue est un intervalle. En particulier, si I est un segment, alors $f(I)$ aussi.

Remarque. La réciproque est fausse en général ($\sin(1/x)$).

Proposition. Soit g une fonction continue définie d'un segment $[a, b]$ dans lui-même. Alors g admet un point fixe.

Définition. Soit $f : D \mapsto \mathbb{R}$ une fonction continue. Un point $x \in D$ est de période p ou p-périodique lorsque :

$$f^p(x) = x \text{ et } f^k(x) \neq x \text{ pour tout } k \in \{1, \dots, p-1\}.$$

Pour $p = 1$, cela correspond à un point fixe.

Théorème. (de Charkovski)
Si f possède un point de période 3, alors elle possède un point à toute période.

Remarque. Ce qui coince dans le cas où f a un point périodique de période 4 survient dans le cas où l'orbite n'est pas croissante. Dans le cas $x < f(x) < f^2(x) < f^3(x)$, on peut faire la même preuve et avoir un point périodique à toute période.

Exemple. La fonction $x \mapsto (x+\frac{1}{2})\mathbf{1}_{[0,\frac{1}{2}]} + (4-2x)\mathbf{1}_{[\frac{1}{2},1]}$, aussi simple soit-elle, admet des points périodiques à toutes les périodes.

Par contre, la fonction $x \mapsto (1-x)$ n'a que des points de période 2, et on voit que l'existence d'un point de période trois est importante.

2 Dérivabilité

2.1 Définitions

Définition. Soit $f : I \to \mathbb{R}$. On dit que f est *dérivable* en $x_0 \in I$ s'il existe a tel que $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = a$. On note alors $f'(x_0) = a$. On parle de même de dérivabilité à droite et à gauche, et de dérivabilité sur un intervalle. On peut alors voir f' comme une fonction, appelée *fonction dérivée* de f.

Proposition. Soit $f : I \to \mathbb{R}$.

- Si f est dérivable en x_0, elle est continue en x_0.
- La dérivabilité équivaut à un développement limité à l'ordre un.
- Si x est un extremum local de f, alors $f'(x) = 0$.
- Si f est dérivable et croissante sa dérivée est positive.

Proposition. Soient f et g deux fonctions dérivables sur leur intervalle de définition.

— $\lambda f + g$ est dérivable, et $(\lambda f + g)' = \lambda f' + g'$.

— fg est dérivable, et $(fg)' = f'g + fg'$.

— Si c'est bien défini, $f \circ g$ est dérivable, avec $(f \circ g)' = g'(f' \circ g)$.

— Si g ne s'annule pas, $\frac{f}{g}$ est dérivable, avec $(\frac{f}{g})' = \frac{f'g - fg'}{g^2}$.

Exemple. La fonction valeur absolue est continue sur $\mathbb{R}$, mais n'est pas dérivable en 0. Plus surprenant, on verra que l'on peut construire des fonctions continues partout et dérivables nulle part.

Définition. Une fonction $f : i \to \mathbb{R}$ est *de classe* $\mathcal{C}^1$, noté $f \in \mathcal{C}^1(I, \mathbb{R})$ si f est dérivable sur I et que sa dérivée y est continue. De même, on peut définir les classes $\mathcal{C}^n$ et $\mathcal{C}^\infty$, qui sont toutes des espaces vectoriels. On note $f^{(n)}(a)$ la dérivée d'ordre n en a.

Exemple. La fonction définie par $x \mapsto x^2 \sin(1/x)$ et prolongée en 0 par continuité est dérivable (car on l'a écrasée à MORT) sur $\mathbb{R}$ mais n'est pas $\mathcal{C}^1$.

Proposition. (formule de Leibniz)
Soient f et g deux fonctions dérivables n fois en un a. Alors (fg) est dérivable n fois en a avec

$$(fg)^{(n)}(a) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(a) f^{(n-k)}(a).$$

2.2 Théorème de Rolle

Théorème. (de Rolle)
Soient $a < b$ dans $\mathbb{R}$ et $f : [a, b] \to \mathbb{R}$ continue sur l'intervalle fermé, dérivable sur l'intervalle ouvert, avec $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.

Remarque. Cela n'est plus vrai en général quand f est à valeurs complexes.

Théorème. (Égalité des accroissements finis)
Soit $f \in \mathcal{C}^0([a, b], \mathbb{R})$ dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$.

Application. (théorème de Darboux)
Soit $f : I \to \mathbb{R}$ dérivable sur I. Alors f' vérifie le théorème des valeurs intermédiaires.

Proposition. Soit $f :]a, b[\to \mathbb{R}$ dérivable. La fonction f est croissante si et seulement si $f' \geq 0$. Si $f' > 0$ alors f est strictement croissante.

Exemple. Il est possible que la dérivée s'annule mais que la fonction reste strictement croissante, comme le montre $x \mapsto x^3$.

Théorème. (formule de Taylor-Lagrange)
Soit $f \in \mathcal{C}^n([a, b], \mathbb{R})$ une fonction $(n + 1)$ fois dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b - a)^k + \frac{(b - a)^{n+1}}{(n + 1)!} f^{(n+1)}(c).$$

2.3 Formules de Taylor

Théorème. Soit f et φ deux fonctions continues sur $[a,b]$ dérivables sur $]a,b[$. On suppose que pour tout $t \in]a,b[$, $|f'(t)| \leq \varphi'(t)$. Alors $|f(b)-f(a)| \leq \varphi(b)-\varphi(a)$.

Corollaire. Soit f une fonction continue et $\mathcal{C}^1$ par morceaux sur $[a,b]$. Alors :

$$\int_a^b f' = f(b) - f(a).$$

Théorème. (formule de Taylor-Young)
Soit $f \in \mathcal{C}^n(I,\mathbb{R})$ $(n+1)$ fois dérivable en a. Alors :

$$f(a+h) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a)h^k}{k!} + \mathrm{o}(h^{n+1}).$$

Définition. Soit $f : I \to \mathbb{R}$. On dit que f admet un *développement limité* à l'ordre n en $x_0 \in I$ si on a une expression de la forme :

$$f(x_0+h) = P(h) + \mathrm{o}_0(h^n),$$

où P est un polynôme réel de degré au plus n.

Remarque. La formule de Taylor-Young montre qu'une fonction $\mathcal{C}^n$ admet un développement limité à l'ordre n (et donc d'avoir tous les dl usuels), mais la réciproque est fausse. En effet, $x \mapsto \mathrm{e}^{-1/x}\sin(\mathrm{e}^{1/x})$ prolongée en 0 par continuité est un $\mathrm{o}(x^n)$ en 0 pour tout n mais n'est pas $\mathcal{C}^1$. Par contre, la dérivabilité équivaut à l'existence d'un développement limité à l'ordre un, ce qui montre que le nombre dérivée est la pente de la tangente qui approxime le mieux la fonction au voisinage du point en question.

Théorème. (Formule de Taylor avec reste intégral)
Soit $f \in \mathcal{C}^{n+1}([a,b],\mathbb{R})$. Alors

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)(b-a)^k}{k!} + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t.$$

Exemple. Soit $f \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ telle que $f(0)=0$. Alors la fonction définie par $h(x) = \frac{f(x)}{x}$ et prolongée par $f'(0)$ en 0 définit une fonction $\mathcal{C}^\infty$ sur $\mathbb{R}$.

Proposition. Soit f continue sur $[a,b]$ et dérivable sur $]a,b[$. Si f' admet une limite au bord, alors on peut la prolonger avec cette valeur pour en faire une fonction dérivable jusqu'au bord.

2.4 Théorème limite central

Lemme. Pour $n \in \mathbb{N}$ et $t > 0$:

$$\left|\mathrm{e}^{\mathrm{i}t} - 1 - \mathrm{i}t - \cdots - \frac{(\mathrm{i}t)^{n-1}}{(n-1)!}\right| \leq \frac{t^n}{n!}.$$

Lemme. Si X admet un moment d'ordre n alors sa fonction caractéristique φ est de classe $\mathcal{C}^n$, et

$$\varphi^{(n)}(\xi) = \mathrm{i}^n \int_{\mathbb{R}} x^n \mathrm{e}^{\mathrm{i}\xi x} \,\mathrm{d}\,\mathbb{P}_X(x).$$

Remarque. Lorsque $\xi = 0$, on retrouve les moments (à une constante près).

Corollaire. Avec les mêmes hypothèses et notations, on a le développement de Taylor suivant :

$$\left|\varphi(\xi+t) - \varphi(\xi) - \sum_{k=1}^{n-1} \frac{t^k}{k!}\varphi^{(k)}(\xi)\right| \leq M_n \frac{|t|^n}{n!}.$$

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1+\ldots+X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3-3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

3 Monotonie et convexité

3.1 Monotonie

Proposition. L'ensemble des points de discontinuité d'une fonction monotone est dénombrable.

Théorème. (de la bijection continue)
Soit $f : I \to f(I)$ continue et monotone, alors f est un homéomorphisme (c'est-à-dire une bijection bicontinue).

Remarque. Si de plus f est $\mathcal{C}^1$ sur I et que sa dérivée ne s'annule pas, alors la réciproque de f est aussi $\mathcal{C}^1$, on parle de *difféomorphisme*. C'est ce qui permet de faire des changements de variable.

Proposition. (Méthode de Laplace au second ordre)
On cherche à calculer un équivalent en l'infini de fonctions de la forme :

$$F(s) = \int_a^b \mathrm{e}^{-s\varphi(t)} f(t)\,\mathrm{d}t,$$

avec φ une fonction régulière, et f continue en a qui ne s'annule pas en a.

Dans le cas où $\varphi' > 0$ sur $]a, b[$, $\varphi'(a) = 0$ et $\varphi''(a) > 0$, on a l'équivalent :

$$F(s) \sim \sqrt{\frac{\pi}{2\varphi''(a)}} \frac{\mathrm{e}^{-s\varphi(a)f(a)}}{\sqrt{s}}.$$

Remarque. Il est toujours possible de se ramener au cas où $t \mapsto \mathrm{e}^{-s_0 t} f(t)$ est intégrable pour un réel s_0 en écrivant :

$$\mathrm{e}^{-st} f(t) = \mathrm{e}^{-(s-s_0)t} \mathrm{e}^{-s_0 t} f(t),$$

et F est alors bien définie pour $s \geq s_0$.

L'application la plus fameuse de ce résultat est la formule de Stirling.

Application. (formule de Striling) Quand x tend vers l'infini :

$$\Gamma(x+1) = \int_0^\infty t^x \mathrm{e}^{-t} \, \mathrm{d}t \sim \frac{x^x}{e^x} \sqrt{2\pi x}.$$

3.2 Convexité

Proposition. Soit $f : I \to \mathbb{R}$ une fonction convexe.

- f est lipschitzienne sur tout segment.
- f admet des dérivées à droite et à gauche en tout point.
- Les dérivées à gauche et à droite sont des fonctions croissantes.
- L'ensemble des points en lesquels f n'est pas dérivable est dénombrable et f est continue en dehors de ces points.

Proposition. Soit $f :]a, b[\to \mathbb{R}$. Si f est dérivable sur $]a, b[$, alors f est convexe si et seulement si f' est croissante sur $]a, b[$. Si f est deux fois dérivable, alors la convexité de f équivaut à la positivité de f''.

Exemple. Sur $[0, \pi/2]$, $\sin(x) \geq \frac{2}{\pi} x$.
Inégalité arithmético-géométrique.

Proposition. Soit $f :]a, b[\to \mathbb{R}$. Si f est convexe et dérivable sur $]a, b[$, alors on a équivalence, pour $c \in]a, b[$, entre le fait que c soit un minimum de f et un point critique de f. En particulier, si f est strictement convexe elle admet au plus un minimum.

3.3 Processus de Galton-Watson

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k \in \mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

4 Suites et séries de fonctions

4.1 Modes de convergence

Définition. Une suite de fonctions $(f_n)_n$ converge *simplement* vers une fonction $f : I \to \mathbb{R}$ lorsque pour tout $x \in I$, la suite $(f_n(x)) \in \mathbb{R}^{\mathbb{N}}$ converge dans $\mathbb{R}$. Une série de fonctions converge simplement lorsque la suite de ses sommes partielles converge simplement.

Définition. Une suite de fonctions $(f_n)_n$ converge *uniformément* (sur E) vers une fonction $f : E \to \mathbb{C}$ lorsque la quantité :

$$\|f - f_n\|_\infty := \sup_{x\in I} |f(x) - f_n(x)|$$

tend vers 0 quand n tend vers l'infini. Une série de fonctions converge uniformément lorsque la suite de ses sommes partielles converge uniformément.

Remarque. La fonction $\|\cdot\|_\infty$ définit une norme sur l'ensemble des fonctions de E dans $\mathbb{C}$.

Remarque. La convergence uniforme entraîne la convergence simple.

Exemple. La suite de fonctions $(x \mapsto \frac{x}{n})_n \geq 1$ converge simplement vers 0 sur $\mathbb{R}$, mais la convergence n'y est pas uniforme.

Définition. Une série de fonctions $(f_n)_n$ converge *normalement* lorsque la série $\sum_{n\in\mathbb{N}} \|f_n\|_\infty$ converge.

Proposition. La convergence normale entraîne la convergence uniforme.

Exemple. La série de fonctions $\sum_{n\in\mathbb{N}} (s \mapsto \frac{1}{n^s})$, $s > 1$, converge normalement sur tout intervalle $[\alpha, +\infty[$, avec $\alpha > 1$. On peut donc définir sa somme $\zeta :]1, +\infty[\to \mathbb{R}$.

Exemple. La série de fonctions $\sum_{n\in\mathbb{Z}} (x \mapsto \mathrm{e}^{-\pi n^2 x})$, $x > 0$, converge normalement sur tout intervalle $[\alpha, +\infty[$, avec $\alpha > 0$. On note sa somme $\vartheta : \mathbb{R}_+^* \to \mathbb{R}$.

4.2 Régularité de la limite

Proposition. Une limite uniforme de fonctions continue est continue.

Proposition. Si $(f_n = [a,b] \to E)_n$, E Banach, converge uniformément vers f, alors la suite des primitives converge uniformément sur la primitive de f correspondante. En particulier, on peut intervertir limite et intégrale.

Proposition. Soit I un intervalle de $\mathbb{R}$. Si une suite de fonctions $(f_n) \in \mathcal{C}^k$ converge uniformément vers une fonction f et que pour tout $j \in \{1,\dots,k\}$, la suite $(f_n^{(j)})_n$ converge uniformément vers une fonction f_j, alors f est $\mathcal{C}^k$ et pour tout $j \in \{1,\dots,k\}$, $f^{(j)} = f_j$.

Remarque. La convergence simple en un point suffirait.

Exemple. Les fonctions ζ et ϑ sont $\mathcal{C}^\infty$ sur leurs domaines respectifs de définition.

Exemple. La suite de fonctions $(x \mapsto \sqrt{x^2 + \frac{1}{n^2}})_n$ est une suite de fonctions dérivables qui converge uniformément vers la fonction valeur absolue, qui n'est pas dérivable.

Proposition. On considère g la fonction 4-périodique affine par morceaux définie sur $\mathbb{R}$ par $f(-2) = -1, f(-1) = 0, f(0) = 1, f(1) = 0$. On considère alors :

$$f = \sum_{n=0}^{\infty} \frac{g(2^{2^n}\cdot)}{2^n}.$$

La fonction f définit une fonction continue sur $\mathbb{R}$ mais nulle part dérivable.

4.3 Séries entières

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n x^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $]-R,R[$ et diverge grossièrement en dehors de l'intervalle fermé $[-R,R]$. On appelle R le *rayon de convergence de la série entière.*

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha,\alpha[$. La série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha,\alpha[$.

Exemple. On obtient plein de développements usuels.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

Chapitre 53

229. Fonctions monotones, fonctions convexes. Exemples et applications.

Sommaire

1 Fonctions monotones 373
1.1 Premières propriétés de régularité 373
1.2 Suites récurrentes, séries 374
1.3 Fonctions de répartition 374
2 Fonctions convexes de la variable réelle 375
2.1 Régularité 375
2.2 Inégalités de convexité 375
2.3 Processus de Galton-Watson 376
3 Cas général 377
3.1 Définition, lien avec les différentielles 377
3.2 Optimisation 377

Présentation du plan

1 Fonctions monotones

1.1 Premières propriétés de régularité

Définition. Une fonction f définie sur une partie de $\mathbb{R}$ est *croissante* si $x \leq y$ entraîne $f(x) \leq f(y)$, et *décroissante* si $f(x) \geq f(y)$. On dit qu'elle est monotone si elle est l'un ou l'autre.

Exemple. La fonction $x \mapsto x^2$ est croissante sur $\mathbb{R}_+$, décroissante sur $\mathbb{R}_-$ et n'est donc pas monotone sur $\mathbb{R}$.

Proposition. Une fonction monotone a des limites à gauche et à droite définies en tout point. En particulier, une fonction monotone n'a qu'un nombre dénombrable de discontinuités, et ce ne sont que des discontinuités de première espèce.

Proposition. Si f est dérivable sur un ouvert de $\mathbb{R}$, alors sa monotonie équivaut à la constance du signe de sa dérivée.

Théorème. Soit I un intervalle de $\mathbb{R}$ et f une fonction continue de I dans $\mathbb{R}$. On suppose que f est continue et strictement monotone. Alors f réalise une bijection de I sur $f(I)$, la réciproque est continue de même stricte monotonie que f, et si f est $\mathcal{C}^1$ et que sa dérivée ne s'annule pas, on a un $\mathcal{C}^1$ difféomorphisme.

Théorème. (théorème de Dini)
La convergence simple d'une suite (f_n) de fonctions réelles d'une variable réelle définies et croissantes sur un segment $[a, b]$ vers une fonction f continue entraîne sa convergence uniforme.

1.2 Suites récurrentes, séries

Remarque. La monotonie de f peut permettre de d'obtenir des informations sur une suite récurrente de la forme $u_{n+1} = f(u_n)$.

Exemple. Soit $x_0 \in [0, 1]$, la suite $u_{n+1} = \sin(u_n)$ tend vers 0.

Remarque. (comparaison à une intégrale)
De la même manière, la monotonie de f peut permettre l'étude asymptotique de $\sum_{k=1}^{n} f(n)$.

Proposition. (critère de Riemann)
La série de terme général (n^α) converge si et seulement si $\alpha < -1$.

1.3 Fonctions de répartition

Proposition. Soit X une variable aléatoire réelle. Sa fonction de répartition F vérifie les propriétés suivantes :

- F est entre 0 et 1
- Sa limite en $-\infty$ est 0 et sa limite en $+\infty$ est 1
- F est continue à droite avec des limites à gauche et est croissante

Définition. (fonction quantile)
Pour $u \in [0, 1]$, et F une fonction de répartition, on définit

$$F^{-1}(u) = \inf\{x, F(x) \geq u\},$$

on définit alors une fonction croissante de $[0, 1]$ dans $\mathbb{R}$.

Théorème. Si F vérifie les propriétés de la propositions précédentes, alors F est la fonction de répartition d'une variable aléatoire.

Théorème. (Glivenko-Cantelli)
La fonction de répartition empirique associée à un n-échantillon $F_n := \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{[X_k, \infty[}$ converge presque sûrement uniformément vers la fonction de répartition F commune aux X_i.

Remarque. En particulier, on montre dans ce théorème que la statistique $D_n := \sqrt{n}\|F_n - F\|$ est libre, et donc on peut construire un test d'adéquation non asymptotique à partir de celle-ci en estimant les quantiles par Monte-Carlo.

2 Fonctions convexes de la variable réelle

2.1 Régularité

Définition. Soit I un intervalle de $\mathbb{R}$. On dit que $f : I \to \mathbb{R}$ est *convexe* si pour tout $t \in]0,1[$ et tout couple $(x,y) \in I$, on a $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Proposition. Soit $f : I \to \mathbb{R}$ une fonction convexe.

— f est lipschitzienne sur tout segment.

— f admet des dérivées à droite et à gauche en tout point.

— Les dérivées à gauche et à droite sont des fonctions croissantes.

— L'ensemble des points en lesquels f n'est pas dérivable est dénombrable et f est continue en dehors de ces points.

Proposition. Soit $f :]a,b[\to \mathbb{R}$. Si f est dérivable sur $]a,b[$, alors f est convexe si et seulement si f' est croissante sur $]a,b[$. Si f est deux fois dérivable, alors la convexité de f équivaut à la positivité de f''.

Exemple. La fonction logarithme est concave, le sinus aussi (sur $[0, \frac{\pi}{2}]$).

Proposition. Soit $f :]a,b[\to \mathbb{R}$. Si f est convexe et dérivable sur $]a,b[$, alors on a équivalence, pour $c \in]a,b[$, entre le fait que c soit un minimum de f et un point critique de f. En particulier, si f est strictement convexe elle admet au plus un minimum.

Exemple. La fonction exp est strictement convexe mais n'admet pas de minimum.

2.2 Inégalités de convexité

Proposition. Soit $x \in [0, \pi/2]$, alors $\frac{2}{\pi}x \leq \sin(x) \leq x$.

Application. La fonction sinus cardinal n'est pas intégrable sur $\mathbb{R}$. Son intégrale est donc semi-convergente.

Proposition. Soient $x_1, \ldots, x_n$ des réels positifs, alors $\sqrt[n]{x_1 \cdots x_n} \leq \frac{1}{n}\sum_{i=1}^{n} x_i$.

Proposition. (inégalité de Jensen)
Soit (Ω, μ) un espace de probabilité, et g une fonction intégrable à valeurs dans un intervalle I de $\mathbb{R}$. Soit φ une fonction convexe de I dans $\mathbb{R}$. Alors

$$\varphi\left(\int_\Omega g \, \mathrm{d}\mu\right) \leq \int_\Omega \varphi \circ g \, \mathrm{d}\mu.$$

Proposition. (inégalité de Hölder)
Si $f \in \mathrm{L}^p$ et $g \in \mathrm{L}^q$ ($\frac{1}{p} + \frac{1}{q} = 1$), alors $fg \in \mathrm{L}^1$ avec $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

2.3 Processus de Galton-Watson

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

3 Cas général

3.1 Définition, lien avec les différentielles

Soit Ω un (ouvert quand c'est nécessaire) convexe de E un espace vectoriel normé.

Définition. On dit que $f : \Omega \to \mathbb{R}$ est convexe si pour tous $x, y \in E$, $t \in]0, 1[$, $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$. On dit qu'elle est strictement convexe si l'inégalité est stricte dès que $x \neq y$.

Proposition. Si f est différentiable sur Ω, alors f est convexe sur Ω si et seulement si $f(y) - f(x) \geq \mathrm{d}f(x)(y - x)$ pour tout $x, y \in \Omega$. Si f est deux fois différentiable, on peut remplacer la condition par $\mathrm{d}^2 f(x) \geq 0$ pour tout $x \in \Omega$. Si f est strictement convexe, on peut remplacer les inégalités larges par des inégalités strictes.

3.2 Optimisation

Proposition. Si f est convexe sur Ω différentiable et si $c \in \Omega$ alors

- c est un minimum de f si et seulement si c est un point critique de f
- Si f est strictement convexe, alors elle admet au plus un minimum sur Ω.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 54

230. Séries de nombres réelles et complexes. Comportement des restes ou des sommes partielles des séries numériques. Exemples.

Sommaire

1 Séries numériques 379
1.1 Série convergente, divergente 379
1.2 Critères de convergence 380
1.3 Comportement de sommes partielles et de restes 381
2 Séries de fonctions définies à partir d'une série 381
2.1 Séries entières 381
2.2 Séries de Fourier 382
3 Probabilités : séries génératrices, séries de variables aléatoires . . . 384
3.1 Séries de variables aléatoires indépendantes 384
3.2 Série génératrice 385

Présentation du plan

1 Séries numériques

1.1 Série convergente, divergente

Définition. On dit que la série $\sum_n u_n$ *converge* si la suite des sommes partielles correspondante converge. Dans le cas contraire on dit qu'elle *diverge*. On dit qu'une série converge absolument si $\sum_n |u_n|$ converge. La somme d'une série est alors la limite en question.

Exemple. La série $\sum_n q^n$ converge si et seulement si $|q| < 1$, et sa somme vaut alors $\frac{1}{1-q}$.

Proposition. Si une série converge, son terme général tend vers 0.

Remarque. Une suite (u_n) converge si et seulement si la série $\sum_n (u_{n+1} - u_n)$ converge.

Exemple. La série $\sum_n \frac{1}{n(n+1)}$ converge et sa somme vaut 1.

Proposition. Une série absolument convergente converge. La réciproque est fausse.

Proposition. Soient (u_n) et (v_n) deux suites de réels positifs. Si $u_n \leq v_n$ pour tout n alors la convergence de la série des v_n entraîne celle des u_n. On en déduit le même genre de résultats pour les notations de Landau.

Théorème. (critère de Riemann)
La série $\sum n^\alpha$ converge si et seulement si $\alpha < -1$.

Proposition. La série $\sum_{n\geq 2} n^{-\alpha}(\ln(n))^{-\beta}$ converge si et seulement si $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$).

1.2 Critères de convergence

Définition. Une série est dite *semi-convergente* lorsqu'elle est convergente mais pas absolument convergente.

Exemple. La série $\sum_n \frac{(-1)^n}{n}$ est semi-convergente.

Théorème. (critère spécial pour les séries alternée)
Soit (a_n) une suite positive décroissante qui tend vers 0. Alors la série $\sum_n (-1)^n a_n$ converge, et $|\sum_{k=n+1}^{\infty} (-1)^n a_n| \leq a_{n+1}$.

Exemple. La série $\sum_n \frac{(-1)^n}{n-\sqrt{n}}$ converge.

Proposition. (transformation d'Abel)
Soient (a_n) et (b_n) deux suites de complexes. On pose $A_n = \sum_{k=0}^{n} a_k$. Alors

$$\sum_{k=0}^{n} a_k b_k = a_0 b_0 - A_0 b_1 + A_n b_n + \sum_{k=1}^{n-1} A_k (b_k - b_{k+1}).$$

Application. La série $\sum_n \frac{\sin(n)}{n}$ converge.

1.3 Comportement de sommes partielles et de restes

Proposition. (lemme de Cesàro)
Soit (u_n) une suite qui converge vers ℓ ; alors $\frac{1}{n}\sum_{k=1}^{n} u_n$ converge vers ℓ également.

Théorème. Soient (u_n) et (v_n) deux suites de réels positifs. Si les séries associées sont divergentes, on peut sommer les relations de comparaison pour les sommes partielles. Si elles sont convergentes, on peut le faire pour les restes.

Application. Soit $u_0 \in]0, 1[$ et $u_{n+1} = \sin(u_n)$. Alors $u_n \sim \sqrt{\frac{3}{n}}$

Proposition. (Lemme de Kronecker)
Soient (x_n) le terme général d'une série convergente et (b_n) une suite croissante de réels qui tend vers l'infini. On a alors :

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{j=1}^{n} b_j x_j = 0.$$

2 Séries de fonctions définies à partir d'une série

2.1 Séries entières

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(0,R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0,R)$. On appelle R le *rayon de convergence de la série entière* et $D(0,R)$ le *disque de convergence.*

Exemple. Soit $\mu = (\mu_n)$ une mesure de probabilité sur $\mathbb{N}$. On définit sa *série génératrice* G_μ par $G_\mu = \sum_n \mu_n z^n$. Alors le rayon de convergence de cette série est au moins 1.

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence. Elle est donc $\mathcal{C}^\infty$ sur l'intervalle de convergence.

Proposition. Une série entière est holomorphe sur son disque de convergence.

Exemple. La série $\sum_n z^n$ converge sur $D(0,1)$, mais pas sur tout le disque fermé. De plus, la convergence n'est pas uniforme sur le disque.

Exemple. La série exponentielle $\sum_n \frac{z^n}{n!}$ converge sur tout $\mathbb{C}$ vers une fonction notée exp. La fonction exp est alors holomorphe sur $\mathbb{C}$ avec $\exp' = \exp$.

Théorème. Soit $\sum_n a_n z n^n$ une série entière. On suppose que la suite (a_n) ne s'annule pas à partir d'un certain rang, et que $|\frac{a_{n+1}}{a_n}|$ converge vers une limite ℓ. Alors le rayon de convergence R de la série entière vaut $\frac{1}{\ell}$. Plus généralement, on a la formule suivante :

$$\frac{1}{R} = \limsup_{n\to+\infty} |a_n|^{\frac{1}{n}}.$$

Corollaire. Le rayon de convergence d'une série entière et de sa dérivée formelle est le même.

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha,\alpha[$. La série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha,\alpha[$.

Exemple. On obtient plein de développements usuels.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

Remarque. Sur le disque fermé de convergence, il peut tout se passer.

Proposition. Soit $\sum_n a_n z^n$ une série entière de rayon de convergence $R > 0$. On suppose que la somme $\sum_{n=0}^{+\infty} a_n R^n$ est finie. Alors la série converge uniformément sur $[0, R]$.

Exemple. Convergence de la série logarithme sur $[0, 1]$, et de l'arctangente.

Théorème. (de Tauber fort) Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1. On suppose que la somme admet une limite en 1 par valeurs inférieures. Si $a_n = \mathrm{O}(1/n)$, alors cette limite est la somme des (a_n).

2.2 Séries de Fourier

Définition. On note $e_k : \mathbb{R} \to \mathbb{C}$ définie par $e_k(t) = \mathrm{e}^{\mathrm{i}kt}$.
Pour une fonction $f \in \mathrm{L}^1_{2\pi}$ et $n \in \mathbb{Z}$, on définit le n-ième coefficient de Fourier par :

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e_{-n}(t) \, \mathrm{d}t \in \mathbb{C}.$$

Remarque. Pour $f \in \mathrm{L}^2_{2\pi} \subset \mathrm{L}^1_{2\pi}$, on a $c_n(f) = \langle f, e_n \rangle$ (en prenant par convention l'antilinéarité à droite).

Exemple. Coefficient de Fourier d'une porte et $c_n(e_k) = \delta_{k,n}$. Coefficient de Fourier de $\cos(\alpha t)$.

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1_{2\pi}$. Alors la suite de ses coefficients de Fourier tend vers 0 en $\pm\infty$.

Théorème. On peut définir $F : f \mapsto (c_n(f))_n$ de $\mathrm{L}^1_{2\pi}$ dans $c_0(\mathbb{Z}, \mathbb{C})$. C'est une application linéaire de norme d'opérateur 1.

Définition. On définit la *somme partielle de Fourier* par $S_N(f) \sum_{n=-N}^{N} c_n(f) e_n$ et la série de Fourier (convergente ou non, c'est une série de fonctions) $S(f) = \sum_{n \in \mathbb{Z}} c_n(f) e_n$.

Définition. Pour $N \in \mathbb{N}^*$, on définit le *noyau de Dirichlet* :

$$D_N(t) = \sum_{n=-N}^{N} \mathrm{e}^{\mathrm{i}nt} = \frac{\sin((N + 1/2)t)}{\sin(t/2)}$$

ainsi que le *noyau de Fejér* :

$$K_N(t) = \frac{1}{N} \sum_{n=0}^{N-1} D_n(t) = \frac{\sin(Nt/2)^2}{N \sin(t/2)^2}.$$

Ainsi, les sommes partielles de Fourier s'écrivent : $S_N(f) = (D_N * f)$, et leur moyenne de Cesàro s'écrit $(K_N * f)$. On montre facilement que K_N est une approximation de l'unité sur $[0, 2\pi]$ (en fait, elle converge au sens des distributions vers $\delta_{\mathbb{Z}}$).

Théorème. (de Fejér)
Soit f une fonction continue et 2π-périodique, alors $(K_N * f)_N$ converge uniformément vers f.

Remarque. Si la série de Fourier d'une fonction continue converge, c'est forcément vers la fonction.

Théorème. (de Dirichlet)
Soit $f \in \mathrm{L}^1_{2\pi}$ et $x_0 \in \mathbb{R}$. On suppose que f admet des limites à gauche et à droite en x_0, et que ses taux d'accroissements à gauche et à droite en x_0 sont intégrables au voisinage de x_0. Alors la série de Fourier de f converge en x_0 avec pour limite $\frac{f(x_0^+)+f(x_0^-)}{2}$.

Corollaire. Si f est continue et $\mathcal{C}^1$ par morceaux, f est somme de sa série de Fourier.

Application. $\zeta(2) = \frac{\pi^2}{6}$.

Remarque. La continuité de f ne suffit pas. Il existe des fonctions continues dont la série de Fourier diverge en 0.

Remarque. Si la série des coefficients de Fourier de f converge, alors la série de Fourier converge normalement. Grâce au théorème taubérien fort, on peut montrer que si les coefficients de Fourier d'une fonction continue décroissent comme $\frac{1}{|n|}$, f est somme de sa série de Fourier.

Proposition. Si f est continue et $\mathcal{C}^1$ par morceaux, alors sa série de Fourier converge normalement vers elle.

Application. Pour tout $x \in \mathbb{R}\backslash\mathbb{Z}$:

$$\cot(\pi x) = \frac{1}{\pi x} + \sum_{n=1}^{+\infty} \frac{2x}{\pi(x^2 - n^2)}.$$

Théorème. (factorisation du sinus)
Pour tout $t \in]-1, 1[$:

$$\frac{\sin(\pi t)}{\pi t} = \prod_{k=1}^{\infty} \left(1 - \frac{t^2}{k^2}\right).$$

3 Probabilités : séries génératrices, séries de variables aléatoires

3.1 Séries de variables aléatoires indépendantes

On s'intéresse à des séries de la forme $\sum_n X_n$, où les (X_i) sont des variables aléatoires indépendantes et L^2. On définit par (S_n) la suite des sommes partielles, et on note $\sigma_i^2 = \mathrm{Var}(X_i)$.

Théorème. (inégalité maximale de Kolmogorov) : On suppose les (X_i) centrées. Alors pour $\varepsilon > 0$ et pour tout n :

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \sigma_i^2.$$

Proposition. Si les (X_i) sont centrées, et si la série des variances converge, alors la série $\sum_n X_n$ converge presque sûrement (et dans L^2).

Théorème. (loi des grands nombres avec un moment d'ordre deux)
Si $\mathbb{E}(X_n)$ tend vers m et si $\sum_{j=1}^{+\infty} \frac{\sigma_j^2}{j^2} < \infty$ alors $\overline{X_n}$ converge presque sûrement et dans L^2 vers m.

3.2 Série génératrice

Définition. Si X est une variables aléatoire à valeurs dans $\mathbb{N}$, sa *série génératrice* est la fonction définie sur $[0,1]$ par $t \mapsto \mathbb{E}(t^X)$. Elle caractérise la loi.

Proposition. (Indécomposabilité de la loi de Poisson)
Soient X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ dont la somme suit une loi de Poisson. Alors X et Y suivent une loi de Poisson.

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0,1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

Chapitre 55

234. Fonctions et espaces de fonctions Lebesgue-intégrables.

Sommaire

1 Définition de l'intégrale de Lebesgue 387
1.1 Intégrale d'une fonction positive 387
1.2 Notion de fonction intégrable 388
1.3 Théorèmes fondamentaux du calcul intégral 388
2 Espaces de Lebesgue 389
2.1 Construction et premières propriétés 389
2.2 Convolution, régularisation 389
2.3 Transformée de Fourier 390
2.4 Compacité dans les espace de Lebesgue 391
3 Ajout de dérivée, les espaces de Sobolev 391
3.1 Dérivée au sens faible, définition 391
3.2 Conditions aux limites 391
3.3 Résolution de problèmes aux limites 392

Présentation du plan

On se place par défaut sur un espace mesuré $(\Omega, \mathcal{F}, \mu)$.

1 Définition de l'intégrale de Lebesgue

1.1 Intégrale d'une fonction positive

Définition. On définit l'intégrale d'une indicatrice comme la mesure de l'ensemble correspondant, puis on prolonge cette définition aux fonctions étagées. On définit alors l'intégrale d'une fonction mesurable positive comme la borne supérieure des intégrales des fonctions étagées inférieures. On a alors la linéarité, la croissance et autres propriétés classiques de l'intégrale de Riemann.

Théorème. (de convergence monotone)
Si $(f_n)_n$ est une suite croissante de fonctions mesurables positives, alors $\sup f_n$ est mesurable positive avec

$$\int_\Omega f_n \to \, \mathrm{d}\mu \int_\Omega \sup f_n \, \mathrm{d}\mu.$$

Théorème. (lemme de Fatou)
Si $(f_n)_n$ est une suite de fonctions mesurables positives, alors

$$\int \liminf f_n \leq \liminf \int f_n.$$

1.2 Notion de fonction intégrable

Définition. On dit qu'une fonction f mesurable est *intégrable* si $|f|$ est d'intégrale finie, et on définit alors son intégrale comme $\int f^+ - \int f_-$. Elle vérifie la linéarité, l'inégalité triangulaire et autres bonnes propriétés.

Remarque. Lorsque les deux sont définies, les intégrales de fonctions positives de Riemann et de Lebesgue coïncident. En particulier, sous l'hypothèse d'intégrabilité, les intégrales ont même valeur.

Cependant, l'intégrale de Lebesgue, contrairement à l'intégrale de Riemann, ne traite pas les intégrales semi-convergentes. On se ramène alors dans ce cas à l'intégrale de Riemann.

1.3 Théorèmes fondamentaux du calcul intégral

Théorème. (de convergence dominée)
Soit (f_n) une suite de fonctions mesurables qui converge simplement vers f, avec une domination intégrable g uniforme en n, alors $\int f_n$ converge vers $\int f$.

Remarque. Une condition nécessaire à l'interversion de la limite et de l'intégrale est l'intégrabilité uniforme de la famille de fonctions (f_n).

Théorème. (Fubini-Tonelli)
Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. Si f est une fonction positive mesurable sur l'espace produit par rapport à la mesure produit $\mu \otimes \nu$, alors $x \mapsto \int_Y f(x, \cdot)$ et $y \mapsto_X f(y, \cdot)$ sont respectivement $\mathcal{A}$ et $\mathcal{B}$ mesurables, avec

$$\int_{X\times Y} f(x,y) \, \mathrm{d}(\mu \otimes \nu)(x,y) = \int_X \int_Y f(x,y) \, \mathrm{d}\nu(y) \, \mathrm{d}\mu(x) = \int_Y \int_X f(x,y) \, \mathrm{d}\mu(x) \, \mathrm{d}\nu(y).$$

Théorème. (Fubini)
Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. Si f est une fonction intégrable sur l'espace produit par rapport à la mesure produit $\mu \otimes \nu$, alors les intégrales par rapport à chacune des variables sont intégrables et on a la même formule que précédemment.

2 Espaces de Lebesgue

2.1 Construction et premières propriétés

Définition. Pour $(\Omega, \mathcal{F}, \mu)$ un espace mesuré, on définit $\mathrm{L}^p(\Omega, \mathcal{F}, \mu)$ comme

— l'ensemble des fonctions p-intégrables sur Ω définies μ-presque partout si $p \in [1, \infty[$

— l'ensemble des fonctions essentiellement bornées sur Ω définies μ-presque partout si $p = \infty$.

On garde ces notations dans la suite.

Proposition. (inégalité de Hölder)
Soit $f \in \mathrm{L}^p(\Omega)$ et $g \in \mathrm{L}^q(\Omega)$ avec $\frac{1}{p} + \frac{1}{q} = 1$, alors $fg \in \mathrm{L}^1(\Omega)$ avec

$$\int_\Omega |fg| \,\mathrm{d}\mu \leq \left(\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{1/p} \left(\int_\Omega |g|^q \,\mathrm{d}\mu\right)^{1/q}.$$

Corollaire. Pour $p \in [1, \infty]$, si on définit $\|f\|_p = (\int_\Omega |f|^p)^{1/p}$ pour $p < \infty$ et $\|f\|_\infty = \mathrm{esssup}(|f|)$, alors on fait de $\mathrm{L}^p(\Omega)$ un espace vectoriel normé.

Théorème. (Riesz-Fischer)
Pour tout p, $\mathrm{L}^p(\Omega)$ est un espace complet.

Corollaire. Si $(f_n) \in (\mathrm{L}^p(\Omega))^{\mathbb{N}}$ converge vers f dans $\mathrm{L}^p(\Omega)$, alors elle converge presque partout simplement vers f à extraction près.

Proposition. Les translations sont continues dans $\mathrm{L}^p(\mathbb{R}^d)$ pour $p \neq \infty$.

2.2 Convolution, régularisation

Définition. Soient f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{C}$. On dit qu'elles sont *convoluables* lorsque $t \mapsto f(x-t)g(t)$ est intégrable pour presque tout x. On définit alors le *produit de convolution* de f et g par la formule :

$$f * g(x) = \int_{\mathbb{R}} f(x-t)g(t) \,\mathrm{d}t.$$

Théorème. L'espace $(\mathrm{L}^1(\mathbb{R}), +, *)$ est une algèbre de Banach commutative.

Proposition. On peut définir le produit de convolution d'une fonction intégrable f et d'une fonction test φ. Alors $f * \varphi$ est $\mathcal{C}^\infty$, avec $(f * \varphi)^{(n)} = f * \varphi^{(n)}$.

Définition. Soit $(\rho_n)_n$ une suite de fonctions positives de $\mathrm{L}^1(\mathbb{R})$. Cette suite est une *approximation de l'unité* si pour tout n, $\int_{\mathbb{R}} \rho_n = 1$ et pour tout $\delta > 0$, $\lim_{n\to\infty} \int_{|x|>\delta} \rho_n(x) \,\mathrm{d}x = 0$.

Exemple. Pour φ une fonction test, la suite $(n\varphi(n\cdot))_n$ définit une approximation de l'unité.

Théorème. Soit $p \in [1, +\infty[$, $k \in \mathbb{N}$ et (φ_n) une approximation de l'unité dans $(\mathcal{D}(\mathbb{R}))^{\mathbb{N}}$. Si $f \in \mathrm{L}^p(\mathbb{R})$, $\varphi_n * f \in \mathrm{L}^p(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R})$ et $\|\varphi_n * f - f\|_p \to f$.

Corollaire. Pour $p < \infty$, l'ensemble des fonctions test est dense dans $\mathrm{L}^p(\mathbb{R})$.

2.3 Transformée de Fourier

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Exemple. Si $g : x \mapsto \mathrm{e}^{-\alpha x}$, pour $\alpha > 0$, alors $\widehat{g} : \xi \mapsto \sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\xi^2}{4\alpha}}$.

Remarque. Il arrive qu'une fonction soit dans $\mathrm{L}^1(\mathbb{R})$ sans que sa transformée de Fourier le soit pour autant. C'est le cas par exemple de la fonction porte $x \mapsto \mathbf{1}_{[-1/2,1/2]}$, donc la transformée de Fourier est $\mathrm{sinc}(\pi\cdot) \notin \mathrm{L}^1$.

Remarque. Cette définition peut être étendue à $\mathrm{L}^1(\mathbb{R}^d)$ en considérant $\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\langle x,\xi\rangle} f(x)\,\mathrm{d}x$.

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}_0^0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}_0^0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un opérateur linéaire continu.

Proposition. La transformée de Fourier est un morphisme d'algèbres de Banach de $(\mathrm{L}^1(\mathbb{R}), +, *)$ dans $(\mathcal{C}_0^0(\mathbb{R}, +, \times))$.

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Proposition. Les fonctions test sont dans la classe de Schwartz, qui est donc dense dans L^p, $p < \infty$.

Proposition. La classe de Schwartz est stable par transformée de Fourier.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x} \widehat{f}(\xi)\,\mathrm{d}\xi.$$

Par dualité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$.

2.4 Compacité dans les espace de Lebesgue

Théorème. (Riesz-Fréchet-Kolmogorov)
Soit F une partie fermée de $\mathrm{L}^p(\mathbb{R})$. Alors F est compacte si et seulement si

— F est bornée

— Les translations sont continues uniformément en F

— Pour tout $\varepsilon > 0$, $\exists R > 0$, $\forall f \in F$,

$$\int_{|x|\geq R} |f(x)|^p \, \mathrm{d}x \leq \varepsilon.$$

3 Ajout de dérivée, les espaces de Sobolev

3.1 Dérivée au sens faible, définition

Définition. On définit l'*espace de Sobolev* $\mathrm{H}^1(I)$ par

$$\mathrm{H}^1(I) = \left\{ u \in \mathrm{L}^2(I); \exists g \in \mathrm{L}^p(I), \int_I u\varphi' = -\int_I g\varphi, \forall \varphi \in \mathcal{C}^1_{\mathrm{c}}(I) \right\}.$$

Pour $u \in \mathrm{H}^1(I)$, le g de la définition est unique et on le note u', car il coïncide avec la dérivée usuelle lorsque u est $\mathcal{C}^1$.

Théorème. Muni de la norme $\sqrt{\| \cdot \|_2^2 + \| \cdot' \|_2^2}$, $\mathrm{H}^1(I)$ est un espace de Hilbert séparable. De plus, si $u \in \mathrm{H}^1(I)$, alors u admet un représentant continu.

3.2 Conditions aux limites

Définition. On définit $\mathrm{H}^1_0(I)$ l'adhérence des fonctions test dans $\mathrm{H}^1(I)$. C'est un espace de Hilbert séparable muni de la norme induite.

Proposition. Si $I =]a, b[$, alors $u \in \mathrm{H}^1_0(I)$ si et seulement si $u(a) = u(b) = 0$.

Théorème. (Inégalité de Poincaré)
Supposons ici que I est borné. Il existe une constante $C > 0$ telle que pour $u \in \mathrm{H}^1_0(I)$, $\|u\|_{H^1} \leq C\|u'\|_2$.

3.3 Résolution de problèmes aux limites

Proposition. (Problème de Dirichlet) Soit $f \in \mathrm{L}^2(0,1)$ ($I =]0,1[$). On cherche à résoudre $-u'' + u = f$ avec les conditions aux limites $u(0) = u(1) = 0$. On dit que $u \in \mathrm{H}^1_0$ est une *solution faible* de ce problème si pour toute fonction test φ :

$$\int_I u'\varphi' + \int_I u\varphi = \int_I f\varphi.$$

Ce problème admet une unique solution faible.

Théorème. (Problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

Chapitre 56

235. Problèmes d'interversion de symboles en analyse.

Sommaire

1 Suites et séries de fonctions 393
1.1 Convergence uniforme 393
1.2 Continuité uniforme 395
1.3 Séries entières 395
2 Intégration 396
2.1 Théorèmes d'intégration 396
2.2 Intégrabilité uniforme 397
2.3 Un exemple, la fonction Gamma 398
3 Théorie de Fourier 398
3.1 Séries de Fourier 398
3.2 Transformée de Fourier 399
3.3 Formule de Poisson 400

Présentation du plan

1 Suites et séries de fonctions

1.1 Convergence uniforme

Définition. Une suite de fonctions $(f_n)_n$ converge *uniformément* (sur E) vers une fonction $f : E \to \mathbb{C}$ lorsque la quantité :

$$\|f - f_n\|_\infty := \sup_{x \in E} |f(x) - f_n(x)|$$

tend vers 0 quand n tend vers l'infini. Une série de fonctions converge uniformément lorsque la suite de ses sommes partielles converge uniformément.

Remarque. La fonction $\|\cdot\|_\infty$ définit une norme sur l'ensemble des fonctions de E dans $\mathbb{C}$.

Remarque. La convergence uniforme entraîne la convergence simple.

Exemple. La suite de fonctions $(x \mapsto \frac{x}{n})_n \geq 1$ converge simplement vers 0 sur $\mathbb{R}$, mais la convergence n'y est pas uniforme.

Définition. Une série de fonctions $(f_n)_n$ converge *normalement* lorsque la série $\sum_{n\in\mathbb{N}} \|f_n\|_\infty$ converge.

Proposition. La convergence normale entraîne la convergence uniforme.

Exemple. La série de fonctions $\sum_{n\in\mathbb{N}} (s \mapsto \frac{1}{n^s})$, $s > 1$, converge normalement sur tout intervalle $[\alpha, +\infty[$, avec $\alpha > 1$. On peut donc définir sa somme $\zeta :]1, +\infty[\to \mathbb{R}$.

Exemple. La série de fonctions $\sum_{n\in\mathbb{Z}} (x \mapsto \mathrm{e}^{-\pi n^2 x})$, $x > 0$, converge normalement sur tout intervalle $[\alpha, +\infty[$, avec $\alpha > 0$. On note sa somme $\vartheta : \mathbb{R}_+^* \to \mathbb{R}$.

Proposition. Si E est compact (ou bien que l'on ne regarde que les fonctions bornées), la norme $\|\cdot\|_\infty$ fait de l'ensemble des fonctions continues de E dans $\mathbb{C}$ un espace de Banach. En particulier, une limite uniforme de fonctions continue est continue.

Proposition. Si $(f_n = [a, b] \to E)_n$, E Banach, converge uniformément vers f, alors la suite des primitives converge uniformément sur la primitive de f correspondante. En particulier, on peut intervertir limite et intégrale.

Proposition. Soit I un intervalle de $\mathbb{R}$. Si une suite de fonctions $(f_n) \in \mathcal{C}^k$ converge uniformément vers une fonction f et que pour tout $j \in \{1, \ldots, k\}$, la suite $(f_n^{(j)})_n$ converge uniformément vers une fonction f_j, alors f est $\mathcal{C}^k$ et pour tout $j \in \{1, \ldots, k\}$, $f^{(j)} = f_j$.

Exemple. Les fonctions ζ et ϑ sont $\mathcal{C}^\infty$ sur leurs domaines respectifs de définition.

Exemple. La suite de fonctions $(x \mapsto \sqrt{x^2 + \frac{1}{n^2}})_n$ est une suite de fonctions dérivables qui converge uniformément vers la fonction valeur absolue, qui n'est pas dérivable.

Proposition. On considère g la fonction 4-périodique affine par morceaux définie sur $\mathbb{R}$ par $f(-2) = -1, f(-1) = 0, f(0) = 1, f(1) = 0$. On considère alors :

$$f = \sum_{n=0}^{\infty} \frac{g(2^{2^n}\cdot)}{2^n}.$$

La fonction f définit une fonction continue sur $\mathbb{R}$ mais nulle part dérivable.

Proposition. Soit $U \subset \mathbb{C}$ un ouvert et (f_n) une suite de fonctions holomorphes définies sur U. Si $(f_n)_n$ converge uniformément vers f sur tout compact de U, la limite f est holomorphe.

Exemple. La fonction ζ définit une fonction holomorphe sur $\{z \in \mathbb{C}, \mathrm{Re}(z) > 1\}$.

1.2 Continuité uniforme

Définition. Une fonction est uniformément continue si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que $|x - y| < \alpha$ entraîne $|f(x) - f(y)| < \varepsilon$.

Théorème. (de Heine)
Une fonction continue sur un compact est uniformément continue.

Théorème. (théorème de Dini)
La convergence simple d'une suite (f_n) de fonctions réelles d'une variable réelle définies et croissantes sur un segment $[a, b]$ vers une fonction f continue entraîne sa convergence uniforme.

Théorème. (Glivenko-Cantelli)
La fonction de répartition empirique associée à un n-échantillon $F_n := \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{[X_k,\infty[}$ converge presque sûrement uniformément vers la fonction de répartition F commune aux X_i.

Remarque. En particulier, on montre dans ce théorème que la statistique $D_n := \sqrt{n}\|F_n - F\|$ est libre, et donc on peut construire un test d'adéquation non asymptotique à partir de celle-ci en estimant les quantiles par Monte-Carlo.

1.3 Séries entières

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(0, R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0, R)$. On appelle R le *rayon de convergence de la série entière* et $D(0, R)$ le *disque de convergence*.

Exemple. Soit $\mu = (\mu_n)$ une mesure de probabilité sur $\mathbb{N}$. On définit sa *série génératrice* G_μ par $G_\mu = \sum_n \mu_n z^n$. Alors le rayon de convergence de cette série est au moins 1.

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence. Elle est donc $\mathcal{C}^\infty$ sur l'intervalle de convergence.

Proposition. Une série entière est holomorphe sur son disque de convergence.

Exemple. La série $\sum_n z^n$ converge sur $D(0, 1)$, mais pas sur tout le disque fermé. De plus, la convergence n'est pas uniforme sur le disque.

Exemple. La série exponentielle $\sum_n \frac{z^n}{n!}$ converge sur tout $\mathbb{C}$ vers une fonction notée exp. La fonction exp est alors holomorphe sur $\mathbb{C}$ avec $\exp' = \exp$.

Proposition. Soit $\sum_n a_n z^n$ une série entière de rayon de convergence $R > 0$. On suppose que la somme $\sum_{n=0}^{+\infty} a_n R^n$ est finie. Alors la série converge uniformément sur $[0, R]$.

Théorème. (de Tauber fort) Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1. On suppose que la somme admet une limite en 1 par valeurs inférieures. Si $a_n = \mathrm{O}(1/n)$, alors cette limite est la somme des (a_n).

2 Intégration

2.1 Théorèmes d'intégration

Théorème. (de convergence monotone)
Si $(f_n)_n$ est une suite croissante de fonctions mesurables positives, alors $\sup f_n$ est mesurable positive avec

$$\int_\Omega f_n \to \,\mathrm{d}\mu \int_\Omega \sup f_n \,\mathrm{d}\mu.$$

Théorème. (lemme de Fatou)
Si $(f_n)_n$ est une suite de fonctions mesurables positives, alors

$$\int \liminf f_n \leq \liminf \int f_n.$$

Théorème. (de convergence dominée)
Soit (f_n) une suite de fonctions mesurables qui converge simplement vers f, avec une domination intégrable g uniforme en n, alors $\int f_n$ converge vers $\int f$.

Théorème. En remplaçant dérivable par $\mathcal{C}^k$ et en dominant les dérivées à chaque ordre par des fonctions intégrables qui ne dépendent pas du paramètre, on obtient la régularité de F, et on peut dériver sous l'intégrale.

Exemple. Les fonctions Γ et ζ sont $\mathcal{C}^\infty$ sur leur intervalle de définition.

Exemple. L'intégrale $\int_\mathbb{R} \mathrm{e}^{-x^2}\,\mathrm{d}x$ vaut $\sqrt{\pi}$.

Théorème. (Fubini-Tonelli)
Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. Si f est une fonction positive mesurable sur l'espace produit par rapport à la mesure produit $\mu \otimes \nu$, alors $x \mapsto \int_Y f(x, \cdot)$ et $y \mapsto_X f(y, \cdot)$ sont respectivement $\mathcal{A}$ et $\mathcal{B}$ mesurables, avec

$$\int_{X\times Y} f(x,y)\,\mathrm{d}(\mu\otimes\nu)(x,y) = \int_X \int_Y f(x,y)\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x) = \int_Y \int_X f(x,y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y).$$

Théorème. (Fubini)
Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. Si f est une fonction intégrable sur l'espace produit par rapport à la mesure produit $\mu \otimes \nu$, alors les intégrales par rapport à chacune des variables sont intégrables et on a la même formule que précédemment.

Exemple. La fonction $(x, y) \mapsto \frac{x^2-y^2}{(x^2+y^2)^2}$ fournit un contre-exemple dans le cas où la fonction n'est pas intégrable en les deux variables.

Exemple.

$$\int_0^1 \frac{\ln(x)}{1+x}\,\mathrm{d}x = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2}.$$

2.2 Intégrabilité uniforme

On se place dans le cas d'une mesure finie ici, donc dans le formalisme des probabilités.

Définition. Une famille de variables aléatoires $(X_i)_{i\in I}$ est *uniformément intégrable* lorsque

$$\lim_{a\to\infty} \sup_{i\in I} \mathbb{E}(|X_i|\mathbf{1}_{|X_i|>a}) = 0.$$

Proposition. Si une famille est uniformément bornée par une variable intégrable positive, alors elle est uniformément intégrable.

Proposition. Si $(X_i)_i$ est bornée dans L^p pour $p > 1$ alors elle est uniformément intégrable.

Définition. Une famille de variables aléatoires (X_i) est dite *équicontinue* si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $\mathbb{P}(A) \leq \delta$ alors $\sup_i \mathbb{E}(|X_i|\mathbf{1}_A) \leq \varepsilon$.

Proposition. Une famille de variables aléatoires est uniformément intégrable si et seulement si elle est équicontinue et bornée dans L^1.

Théorème. Une suite de variables aléatoires converge dans L^1 si et seulement si elle converge en probabilité et qu'elle est uniformément intégrable.

Exemple. Pour $(X_n := \frac{1}{n^2}\delta_{n^2} + (1-\frac{1}{n})\delta_0)$, on a convergence en probabilité mais pas en norme 1.

2.3 Un exemple, la fonction Gamma

Définition. On définit la fonction gamma sur $\Omega_0 := \{z \in \mathbb{C}, \mathrm{Re}(z) > 0\}$ par :

$$\Gamma(z) = \int_0^{+\infty} t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t.$$

Proposition. La fonction gamma est bien définie sur Ω_0, et définit une fonction holomorphe sur tout ce domaine.

Remarque. En effectuant une intégration par parties, on peut remarquer que Γ satisfait l'équation fonctionnelle :

$$\Gamma(z+1) = z\Gamma(z),$$

et qu'en particulier pour $n \in \mathbb{N}$, $\Gamma(n+1) = n!$. En effet :

$$\Gamma(z+1) = \left[-t^z\mathrm{e}^{-t}\right]_0^{\infty} + z\int_0^{\infty} t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t = z\Gamma(z).$$

Théorème. La fonction gamma se prolonge en une fonction méromorphe sur $\Omega := \mathbb{C}\backslash -\mathbb{N}$, avec un pôle simple en $-n$ pour tout $n \in \mathbb{N}$, de résidu $\frac{(-1)^n}{n!}$.

Théorème. (Formule des compléments) Pour tout $z \in \mathbb{C}\backslash\mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Corollaire. La fonction Γ ne s'annule pas. De plus, la fonction $\frac{1}{\Gamma}$ est entière.

3 Théorie de Fourier

3.1 Séries de Fourier

Définition. On définit la *somme partielle de Fourier* par $S_N(f) = \sum_{n=-N}^{N} c_n(f)e_n$ et la série de Fourier (convergente ou non, c'est une série de fonctions) $S(f) = \sum_{n\in\mathbb{Z}} c_n(f)e_n$.

Définition. Pour $N \in \mathbb{N}^*$, on définit le *noyau de Dirichlet* :

$$D_N(t) = \sum_{n=-N}^{N} \mathrm{e}^{int} = \frac{\sin((N+1/2)t)}{\sin(t/2)}$$

ainsi que le *noyau de Fejér* :

$$K_N(t) = \frac{1}{N}\sum_{n=0}^{N-1} D_n(t) = \frac{\sin(Nt/2)^2}{N\sin(t/2)^2}.$$

Ainsi, les sommes partielles de Fourier s'écrivent : $S_N(f) = (D_N * f)$, et leur moyenne de Cesàro s'écrit $(K_N * f)$. On montre facilement que K_N est une approximation de l'unité sur $[0, 2\pi]$ (en fait, elle converge au sens des distributions vers $\delta_{\mathbb{Z}}$).

Théorème. (de Fejér)
Soit f une fonction continue et 2π-périodique, alors $(K_N * f)_N$ converge uniformément vers f.

Remarque. Si la série de Fourier d'une fonction continue converge, c'est forcément vers la fonction en question.

Définition. Soit $(\rho_n)_n$ une suite de fonctions de $\mathrm{L}^1_{2\pi} \cap \mathcal{C}^0([0, 2\pi])$. On dit que c'est une *approximation de l'unité* lorsque $\int \rho_n = 1$ pour tout n, et pour tout $\delta > 0$, $\lim_n \int_{|x|\geq\delta} \rho_n(x)\,\mathrm{d}x = 0$.

Théorème. Soit $(\rho_n)_n$ une approximation de l'unité, et $f \in \mathrm{L}^p_{2\pi}$. Alors, on a la convergence suivante dans $\mathrm{L}^p_{2\pi}$:

$$\rho_n * f \xrightarrow[n\to\infty]{} f.$$

Corollaire. (Fejér-Lebesgue)
Puisque le noyau de Fejér est une approximation de l'unité, les sommes de Fejér convergent dans L^p vers la fonction considérée. En particulier, les polynômes trigonométriques sont denses dans L^p.

Théorème. (de Dirichlet)
Soit $f \in \mathrm{L}^1_{2\pi}$ et $x_0 \in \mathbb{R}$. On suppose que f admet des limites à gauche et à droite en x_0, et que ses taux d'accroissements à gauche et à droite en x_0 sont intégrables au voisinage de x_0. Alors la série de Fourier de f converge en x_0 avec pour limite $\frac{f(x_0^+)+f(x_0^-)}{2}$.

Théorème. Soit E l'ensemble des fonctions continues de (t, x) sur $\mathbb{R}_+ \times \mathbb{R}$ et 2π-périodiques en la seconde variable. Soit f une fonction $\mathcal{C}^1$ et 2π-périodique. Il existe une unique solution au problème suivant :

— $u \in E \cap \mathcal{C}^2(\mathbb{R}_+^* \times \mathbb{R})$
— $\forall t > 0, \forall x \in \mathbb{R}$, $\partial_t(u)(t, x) = \partial_x^2 u(t, x)$
— $u(0, .) = f$.

3.2 Transformée de Fourier

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}_0^0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}_0^0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un opérateur linéaire continu.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

- $\mathcal{F}(\check{f}) = \mathcal{F}\check{(}f)$.
- $\mathcal{F}(\overline{f}) = \overline{\mathcal{F}\check{(}f)}$.
- Pour $\lambda > 0$, $\mathcal{F}(f(\lambda\cdot))\frac{1}{\lambda}\mathcal{F}(f)\left(\frac{\cdot}{\lambda}\right)$
- $\mathcal{F}(\tau_a f) = \mathrm{e}^{-\mathrm{i}a\cdot}\mathcal{F}(f)$ et $\mathcal{F}(\mathrm{e}^{-\mathrm{i}a\cdot}) = \tau_a(\mathcal{F}(f))$, $a \in \mathbb{R}$.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

- Si f est $\mathcal{C}^1$ sur $\mathbb{R}$, avec $f' \in \mathrm{L}^1$. Alors $\mathcal{F}(f') = \mathrm{i}\xi\mathcal{F}(f)$.
- Si xf est intégrable, alors $\mathcal{F}(f)$ est dérivable sur $\mathbb{R}$ avec $\mathcal{F}(f)' = -\mathrm{i}\mathcal{F}(xf)$.

Exemple. Cette propriété permet de recalculer la transformée de Fourier de la gaussienne.

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Proposition. Les fonctions test sont dans la classe de Schwartz, qui est donc dense dans L^p, $p < \infty$.

Proposition. La classe de Schwartz est stable par transformée de Fourier.

Proposition. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x}\widehat{f}(\xi)\,\mathrm{d}\xi.$$

Par dualité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$ (d'après la proposition suivante).

3.3 Formule de Poisson

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Remarque. On peut appliquer cette formule sous des hypothèses plus faibles. En effet, la seule chose dont on a vraiment besoin est le caractère $\mathcal{C}^1$ de la somme, c'est-à-dire la convergence normale de la série de fonctions, pour pouvoir appliquer le théorème de Dirichlet. On peut donc se restreindre aux hypothèses suivantes :

— $f \in \mathcal{C}^0 \cap \mathrm{L}^1$ et $|f| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1$

— $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty$ ou bien $|f'| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1.$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x} \widehat{f}(\xi)\, \mathrm{d}\xi.$$

Théorème. (Échantillonnage de Shannon)
Soit $f \in \mathcal{S}(\mathbb{R})$ telle que le support de $\widehat{f}$ soit inclus dans $[-F, F]$ avec $F > 0$ et $2F \leq 2\pi$, alors :

$$f = \sum_{k\in\mathbb{Z}} f(k)\, \mathrm{sinc}(\pi(\cdot - k)).$$

Application. (Fonction ϑ de Jacobi)
On pose $\vartheta : x \mapsto \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 x}$, pour $x > 0$. La fonction ϑ est solution de l'équation fonctionnelle suivante sur $\mathbb{R}_+^*$:

$$\vartheta(x) = \frac{1}{\sqrt{x}} \vartheta\left(\frac{1}{x}\right).$$

Chapitre 57

236. Illustrer par des exemples quelques méthodes de calcul d'intégrales de fonctions d'une ou plusieurs variables.

Sommaire

1 Méthodes générales de calcul d'intégrales 404
1.1 Le théorème fondamental 404
1.2 Les sommes de Riemann 404
1.3 L'intégration par parties 404
1.4 théorème de Fubini 405
1.5 Le changement de variable 405
2 Intégrales à paramètres 406
2.1 Utilisation de la convergence dominée : méthode de Feynman 406
2.2 Transformée de Fourier fonction intégrables 406
3 Intégration complexe 407
3.1 Intégration sur un contour 407
3.2 Théorie des résidus 408
4 Calcul approché d'intégrales 408
4.1 Méthode de Laplace 408
4.2 Méthode de Monte-Carlo 409

Présentation du plan

1 Méthodes générales de calcul d'intégrales

1.1 Le théorème fondamental

Théorème. Soit f une fonction continue sur un segment $[a, b]$ de $\mathbb{R}$. Alors la fonction $F : x \mapsto \int_a^x f$ est dérivable sur $]a, b[$ et sa dérivée est f. Ainsi, si G est une autre primitive de f, $G - F$ est constante.

Exemple. On déduit de ce théorème des primitives pour plein de fonctions usuelles.

Application. (Comparaison série-intégrale) La série $\sum_{n\in\mathbb{N}^*} \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

Exemple. Pour tout $z \in \mathbb{C}$ tel que $\mathrm{Re}(z) > 0$, $\int_0^\infty \mathrm{e}^{zx}\,\mathrm{d}x = \frac{1}{z}$.

Théorème. Si $f \in \mathrm{L}^1_{\mathrm{loc}}(I)$, et si F est une primitive de f (*i.e.* F est absolument continue et $F' = f$ presque partout), alors pour tout $a < b \in I$, $\int_a^b f = F(b) - F(a)$.

Exemple. Il existe une fonction f dont la dérivée est nulle presque partout mais qui vérifie $f(0) = 0$ et $f(1) = 1$.

1.2 Les sommes de Riemann

Proposition. Soit f une fonction continue sur $[a, b]$ à valeurs dans $\mathbb{R}$. Alors, par définition de l'intégrale de Riemann, si on pose $S_n = \frac{b-a}{n} \sum_{k=1}^n f\left(a + k\,\frac{b-a}{n}\right)$:

$$\int_a^b f = \lim_{n\to\infty} S_n.$$

Exemple. Quand n tend vers $+\infty$, $\frac{\pi}{4} - \sum_{k=1}^n \frac{n^2}{k^2+n^2} \sim \frac{1}{4n}$.

1.3 L'intégration par parties

Théorème. Soient deux fonctions $u, v \in \mathcal{C}^1([a, b])$. Alors :

$$\int_a^b uv' = u(b)v(b) - u(a)v(a) - \int_a^b u'v.$$

Exemple. Une primitive du logarithme est $x \mapsto x\ln(x) - x$ sur $\mathbb{R}_+^*$.

Exemple. On définit, pour $x > 0$, $\Gamma(x) = \int_0^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$. Alors pour tout $x > 0$, $x\Gamma(x) = \Gamma(x + 1)$.

Exemple. Les intégrales $\int_1^\infty \frac{\cos(t)}{t^\alpha}\,\mathrm{d}t$ et $\int_1^\infty \frac{\sin(t)}{t^\alpha}\,\mathrm{d}t$ convergent pour $\alpha > 0$.

1.4 théorème de Fubini

Théorème. (Fubini-Tonelli)
Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. Si f est une fonction positive mesurable sur l'espace produit par rapport à la mesure produit $\mu \otimes \nu$, alors $x \mapsto \int_Y f(x, \cdot)$ et $y \mapsto_X f(y, \cdot)$ sont respectivement $\mathcal{A}$ et $\mathcal{B}$ mesurables, avec

$$\int_{X\times Y} f(x, y)\,\mathrm{d}(\mu \otimes \nu)(x, y) = \int_X \int_Y f(x, y)\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x) = \int_Y \int_X f(x, y)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y).$$

Théorème. (Fubini)
Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces mesurés σ-finis. Si f est une fonction intégrable sur l'espace produit par rapport à la mesure produit $\mu \otimes \nu$, alors les intégrales par rapport à chacune des variables sont intégrables et on a la même formule que précédemment.

Remarque. Pour appliquer le théorème de Fubini, on vérifie que la fonction est intégrable sur le produit en utilisant Fubini-Tonelli à $|f|$. Si une des deux intégrales converge, alors les trois convergent et on a bien l'intégrabilité sur produit.

Remarque. Le théorème d'interversion série-intégrale est un cas particulier du théorème de Fubini.

Exemple.

$$\int_0^1 \frac{\ln(x)}{1+x}\,\mathrm{d}x = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2}.$$

1.5 Le changement de variable

Théorème. Soit φ un $\mathcal{C}^1$-difféomorphisme croissant de $[\alpha, \beta[$ vers $[a, b[$ et f une fonction continue sur $[a, b[$. Alors les intégrales

$$\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}t = \int_a^b f(x)\,\mathrm{d}t$$

sont de même nature et égales en cas de convergence.

Exemple. Soit f une fonction continue définie sur un segment $[a, b]$, alors :

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(a+b-y)\,\mathrm{d}y.$$

Ainsi, pour tout $\alpha > 0$, $\int_0^{\pi/2} \frac{\mathrm{d}t}{1+\tan(t)^\alpha} = \frac{\pi}{4}$ pour tout $\alpha > 0$.

Théorème. Soit $\varphi = U \to V$ un $\mathcal{C}^1$-difféomorphisme entre deux ouverts de $\mathbb{R}^d$. Alors pour toute fonction mesurable $f : V \to \mathbb{C}$, la composée $f \circ \varphi$ est aussi mesurable et si f est intégrable, $f \circ \varphi$ aussi, avec la relation :

$$\int_V f(y)\,\mathrm{d}y = \int_U f(\varphi(x))|\mathrm{Jac}(\varphi)(x)|\,\mathrm{d}x.$$

Exemple. (Coordonnées polaires)
Si on considère $\varphi : (r, \theta) \mapsto (r\cos(\theta), r\sin(\theta))$, alors on définit un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}_+^* \times]-\pi, \pi[$ dans $\mathbb{R}^2 \backslash \mathbb{R}_-$, et on a, pour f Lebesgue intégrable sur $\mathbb{R}^2$:

$$\int_{\mathbb{R}^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \int_0^{2\pi} f(r\cos(\theta), r\sin(\theta))\,r\mathrm{d}r\,\mathrm{d}\theta.$$

Application. L'intégrale $\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x$ vaut $\sqrt{\pi}$.

2 Intégrales à paramètres

2.1 Utilisation de la convergence dominée : méthode de Feynman

Le théorème de convergence dominée et ses applications constituent une grosse boîte à outils pour calculer des intégrales. En particulier, la *méthode de Feynman* constitue à faire apparaître un paramètre, et étudier la fonction définie par cette intégrale à paramètres pour calculer l'intégrale finale.

Exemple. On peut retrouver la valeur de l'intégrale précédente en étudiant l'intégrale à paramètre $x \mapsto \int_{\mathbb{R}} \frac{\mathrm{e}^{-x^2(1+t^2)}}{1+t^2}\,\mathrm{d}t$.

Exemple. (intégrale de Serret)

$$\int_0^1 \frac{\ln(1+x)}{1+x^2}\,\mathrm{d}x = \frac{\pi \ln(2)}{8}.$$

2.2 Transformée de Fourier fonction intégrables

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}_0^0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}_0^0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un morphisme d'algèbres de Banach de $(\mathrm{L}^1(\mathbb{R}), +, *)$ dans $(\mathcal{C}_0^0(\mathbb{R}, +, \times))$.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

— Si f est $\mathcal{C}^1$ sur $\mathbb{R}$, avec $f' \in \mathrm{L}^1$. Alors $\mathcal{F}(f') = \mathrm{i}\xi\mathcal{F}(f)$.

— Si xf est intégrable, alors $\mathcal{F}(f)$ est dérivable sur $\mathbb{R}$ avec $\mathcal{F}(f)' = -\mathrm{i}\mathcal{F}(xf)$.

Exemple. Si $g : x \mapsto \mathrm{e}^{-\alpha x}$, pour $\alpha > 0$, alors $\widehat{g} : \xi \mapsto \sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\xi^2}{4\alpha}}$.

Théorème. (Inversion de Fourier)
Soit f une fonction L^1 dont la transformée de Fourier est L^1. Alors presque partout :

$$f(x) = \int_{\mathbb{R}} \widehat{f}(\xi)\mathrm{e}^{\mathrm{i}\xi x}\,\mathrm{d}\xi.$$

3 Intégration complexe

3.1 Intégration sur un contour

Définition. Soit I un intervalle de $\mathbb{R}$ et $\gamma : I \to \mathbb{C}$ une fonction lisse (on dit que γ est un *chemin lisse*, si I est fermé et que γ prend la même valeur aux deux bords, on parle de chemin fermé). Pour f une fonction continue sur $\gamma(I)$, on pose :

$$\int_\gamma f = \int_\gamma f(\zeta)\,\mathrm{d}\zeta := \int_I f(\gamma(t))\gamma'(t)\,\mathrm{d}t.$$

Théorème. (de Cauchy)
Soit Ω un ouvert convexe et f holomorphe sur Ω. Alors pour tout chemin fermé tracé dans Ω :

$$\int_\gamma f = 0.$$

Théorème. (formule de Cauchy)
Soit $\gamma : I \to \mathbb{C}$ un chemin fermé dans un ouvert convexe Ω, et f holomorphe sur Ω. Pour $z \in \Omega \backslash \gamma(I)$:

$$f(z)\,\mathrm{Ind}_\gamma(z) = \frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{f(\zeta)}{\zeta - z}\,\mathrm{d}\zeta.$$

Exemple. On peut calculer $\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x$ avec la formule de Cauchy.

Corollaire. Pour tout ouvert du plan U, toute fonction holomorphe sur U y est développable en série entière (au voisinage de chaque point).

Corollaire. (Théorème de Liouville)
Toute fonction entière bornée est constante.

3.2 Théorie des résidus

Définition. Soit f une fonction holomorphe sur $\Omega\backslash\{a\}$, où Ω est un ouvert de $\mathbb{C}$. On dit que a est un *pôle d'ordre* $m \in \mathbb{N}$ pour f si on a un développement de la forme :

$$f(z) - \sum_{k=1}^{m} \frac{c_{-k}}{(z-a)^k} = g(z) \qquad \forall z \neq a,$$

où, $c_{-m} \neq 0$, g est holomorphe sur $\Omega\backslash\{a\}$ et admet une singularité effaçable en a. On note $\mathrm{Res}(f, a) = c_{-1}$ le *résidu* de f en a.

Théorème. (des résidus)
Soit Ω un ouvert convexe et f une fonction holomorphe sur $\Omega\backslash A$, où A est l'ensemble des pôles de f, supposé fini. Soit γ une chemin fermé dont l'image ne rencontre pas A. Alors :

$$\sum_{a\in A} \mathrm{Ind}(\gamma, a)\mathrm{Res}(f, a) = \frac{1}{2\mathrm{i}\pi}\int_\gamma f.$$

Application. (formule des compléments) Pour tout $z \in \mathbb{C}\backslash\mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)},$$

où Γ est un prolongement analytique de la fonction Γ définie précédemment.

4 Calcul approché d'intégrales

4.1 Méthode de Laplace

La méthode de Laplace est une façon de calculer des équivalents d'une intégrale à paramètre. Plus précisément, on cherche à calculer un équivalent en l'infini de fonctions de la forme :

$$F(s) = \int_a^b \mathrm{e}^{-s\varphi(t)} f(t)\,\mathrm{d}t,$$

avec φ une fonction régulière, et f continue en a qui ne s'annule pas en a.

Proposition. (Méthode de Laplace au premier ordre)
Si φ' est strictement positive sur $[a,b[$ et si f est intégrable, alors :

$$F(s) \sim \frac{1}{\varphi'(a)} \frac{\mathrm{e}^{-s\varphi(a)} f(a)}{s}.$$

Proposition. (Méthode de Laplace au second ordre)
Dans le cas où $\varphi' > 0$ sur $]a,b[$, $\varphi'(a) = 0$ et $\varphi''(a) > 0$, on a l'équivalent :

$$F(s) \sim \sqrt{\frac{\pi}{2\varphi''(a)}} \frac{\mathrm{e}^{-s\varphi(a) f(a)}}{\sqrt{s}}.$$

Remarque. Il est toujours possible de se ramener au cas où $t \mapsto \mathrm{e}^{-s_0 t} f(t)$ est intégrable pour un réel s_0 en écrivant :

$$\mathrm{e}^{-st} f(t) = \mathrm{e}^{-(s-s_0)t} \mathrm{e}^{-s_0 t} f(t),$$

et F est alors bien définie pour $s \geq s_0$.

Application. (formule de Stirling)
Quand x tend vers l'infini :

$$\Gamma(x+1) = \int_0^\infty t^x \mathrm{e}^{-t} \,\mathrm{d}t \sim \frac{x^x}{e^x} \sqrt{2\pi x}.$$

Théorème. Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\| \to \infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x) \mathrm{e}^{-t\varphi(x)} \,\mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}} \left(\frac{2\pi}{t}\right)^{d/2} \mathrm{e}^{-t\varphi(x_0)}.$$

4.2 Méthode de Monte-Carlo

Théorème. (loi des grands nombres)
Soit $(X_n)_n \in \mathrm{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ une suite de variables aléatoires réelles indépendantes i.i.d. de moyenne commune μ. Alors $\overline{X_n}$ converge presque sûrement vers μ.

Application. (Méthode de Monte-Carlo)
Avec les mêmes notations, si les (X_n) admettent pour densité commune f et si g est une fonction mesurable telle que $g(X_1)$ est intégrable, alors :

$$\frac{1}{n} \sum_{k=1}^n g(X_k) \xrightarrow[n\to\infty]{\text{ps}} \int_{\mathbb{R}} g(x) f(x) \,\mathrm{d}x.$$

Chapitre 58

239. Fonctions définies par un intégrale dépendant d'un paramètre. Exemples et applications.

Sommaire

1 Théorèmes de régularité 412
1.1 Continuité 412
1.2 Dérivabilité 412
1.3 Holomorphie 412
2 Convolution 413
2.1 Définition 413
2.2 Approximation par convolution 413
3 Transformée de Fourier 414
3.1 Pour les fonctions intégrables 414
3.2 Classe de Schwartz et inversion de Fourier 414
3.3 Transformée de Fourier de mesures 415
4 Asymptotique d'intégrale à paramètres 416
4.1 Méthode de Laplace 416
4.2 Théorème limite central pour les densités 417

Présentation du plan

On s'intéresse dans cette leçon à des fonctions de la forme :

$$F(x) = \int_\Omega f(x,\omega)\,\mathrm{d}\mu,$$

où $(\Omega, \mathcal{F}, \mu)$ est un espace mesuré, $f : E \times \Omega \to \mathbb{C}$, avec E un espace métrique et $f(x,\cdot)$ est mesurable pour tout x.

1 Théorèmes de régularité

Définition. Pour $s > 1$, on définit $\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$ la fonction *zêta de Riemann*.

Définition. Pour $x > 0$ de partie réelle positive, on définit $\Gamma(x) \int_{\mathbb{R}} t^{x-1} \mathrm{e}^{-t} \, \mathrm{d}t$ la fonction *Gamma*.

1.1 Continuité

Théorème. Soit $x_0 \in E$. On suppose que pour presque tout $\omega \in \Omega$, $x \mapsto f(x, \omega)$ est continue en x_0, et qu'il existe $g \in \mathrm{L}^1(\mu)$ positive telle que l'on ait presque partout $|f(x, \cdot)| \leq g$. Alors F est continue en x_0.

Exemple. La fonction ζ est continue sur $[1, +\infty[$.

Exemple. La fonction Γ est continue sur $\mathbb{R}_+^*$.

1.2 Dérivabilité

Théorème. On prend ici $E = I$ intervalle de $\mathbb{R}$. On suppose que pour presque tout $\omega \in \Omega$, $x \mapsto f(x, \omega)$ est dérivable sur I, et que pour tout K compact de I, il existe $g_K \in \mathrm{L}^1(\mu)$ positive telle que l'on ait presque partout $|\partial_x f(x, \cdot)| \leq g_K$. Alors F est dérivable sur I et sa dérivée est donnée en dérivant sous l'intégrale.

Théorème. En remplaçant dérivable par $\mathcal{C}^k$ et en dominant les dérivées à chaque ordre par des fonctions intégrables qui ne dépendent pas du paramètre, on obtient la régularité de F, et on peut dériver sous l'intégrale.

Exemple. Les fonctions Γ et ζ sont $\mathcal{C}^\infty$ sur leur intervalle de définition.

Exemple. L'intégrale $\int_{\mathbb{R}} \mathrm{e}^{-x^2} \, \mathrm{d}x$ vaut $\sqrt{\pi}$.

1.3 Holomorphie

Théorème. On prend ici $E = U$ un ouvert de $\mathbb{C}$. On suppose que pour presque tout $\omega \in \Omega$, $z \mapsto f(z, \omega)$ est holomorphe sur U, et qu'il existe, pour tout compact K de U, $g_K \in \mathrm{L}^1(\mu)$ positive telle que l'on ait presque partout (sur K) $|f(z, \cdot)| \leq g$. Alors F est holomorphe sur U.

Exemple. La fonction ζ se prolonge en une fonction holomorphe sur $\{z \in \mathbb{C}, \mathrm{Re}(z) > 1\}$, ainsi que Γ sur $\mathbb{C} \backslash \mathbb{Z}$.

Théorème. (Formule des compléments) Pour tout $z \in \mathbb{C} \backslash \mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Corollaire. La fonction Γ ne s'annule pas. De plus, la fonction $\frac{1}{\Gamma}$ est entière.

2 Convolution

2.1 Définition

Définition. Soient f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{C}$. On dit qu'elles sont *convoluables* lorsque $t \mapsto f(x-t)g(t)$ est intégrable pour presque tout x. On définit alors le *produit de convolution* de f et g par la formule :

$$f * g(x) = \int_{\mathbb{R}} f(x-t)g(t)\,\mathrm{d}t.$$

Théorème. L'espace $(\mathrm{L}^1(\mathbb{R}), +, *)$ est une algèbre de Banach.

Proposition. De la même manière, on peut définir le produit de convolution d'une fonction intégrable f et d'une fonction test φ. Alors $f * \varphi$ est $\mathcal{C}^\infty$, avec $(f * \varphi)^{(n)} = f * \varphi^{(n)}$.

Définition. Soit $(\rho_n)_n$ une suite de fonctions positives de $\mathrm{L}^1(\mathbb{R})$. Cette suite est une *approximation de l'unité* si pour tout n, $\int_{\mathbb{R}} \rho_n = 1$ et pour tout $\delta > 0$, $\lim_{n\to\infty} \int_{|x|>\delta} \rho_n(x)\,\mathrm{d}x = 0$.

Exemple. Pour φ une fonction test, la suite $(n\varphi(n\cdot))_n$ définit une approximation de l'unité.

2.2 Approximation par convolution

Théorème. Soit $p \in [1, +\infty[$, $k \in \mathbb{N}$ et (φ_n) une approximation de l'unité dans $(\mathcal{D}(\mathbb{R}))^{\mathbb{N}}$.

— Si $f \in \mathrm{L}^p(\mathbb{R})$, $\varphi_n * f \in \mathrm{L}^p(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R})$ et $\|\varphi_n * f - f\|_p \to f$.

— Si $f \in \mathcal{C}^k_{\mathrm{c}}(\mathbb{R})$, $\varphi_n * f \in \mathcal{D}(\mathbb{R})$ et $\|\varphi_n * f - f\|_\infty \to f$.

Corollaire. Pour $k \in \mathbb{N}$, $\mathcal{D}(\mathbb{R})$ est dense dans $\mathcal{C}^k_{\mathrm{c}}(\mathbb{R})$ pour la norme infinie et est dense dans $\mathrm{L}^p(\mathbb{R})$ pour la norme L^p.

3 Transformée de Fourier

3.1 Pour les fonctions intégrables

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Exemple. Si $g : x \mapsto \mathrm{e}^{-\alpha x}$, pour $\alpha > 0$, alors $\widehat{g} : \xi \mapsto \sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\xi^2}{4\alpha}}$.

Remarque. Il arrive qu'une fonction soit dans $\mathrm{L}^1(\mathbb{R})$ sans que sa transformée de Fourier le soit pour autant. C'est le cas par exemple de la fonction porte $x \mapsto \mathbf{1}_{[-1/2,1/2]}$, donc la transformée de Fourier est $\mathrm{sinc}(\pi\cdot) \notin \mathrm{L}^1$.

Remarque. Cette définition peut être étendue à $\mathrm{L}^1(\mathbb{R}^d)$ en considérant $\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\langle x,\xi\rangle} f(x)\,\mathrm{d}x$.

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}^0_0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}^0_0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un morphisme d'algèbres de Banach de $(\mathrm{L}^1(\mathbb{R}), +, *)$ dans $(\mathcal{C}^0_0(\mathbb{R}, +, \times))$.

3.2 Classe de Schwartz et inversion de Fourier

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Proposition. Les fonctions test sont denses dans la classe de Schwartz, qui est donc dense dans L^p, $p < \infty$.

Proposition. La classe de Schwartz est stable par transformée de Fourier.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x+2\pi n) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x}\widehat{f}(\xi)\,\mathrm{d}\xi.$$

Par dualité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$.

Théorème. (Échantillonnage de Shannon)
Soit $f \in \mathcal{S}(\mathbb{R})$ telle que le support de $\widehat{f}$ soit inclus dans $[-F, F]$ avec $F > 0$ et $2F \leq 2\pi$, alors :

$$f = \sum_{k\in\mathbb{Z}} f(k)\,\mathrm{sinc}(\pi(\,\cdot - k)).$$

3.3 Transformée de Fourier de mesures

Définition. Soit μ une mesure positive finie sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. On définit sa transformée de Fourier $\widehat{\mu} \in \mathcal{C}^0(\mathbb{R})$ par

$$\widehat{\mu}(t) = \int_R \mathrm{e}^{-\mathrm{i}\langle t,x\rangle}\,\mathrm{d}\mu(x).$$

Proposition. Si μ est à densité par rapport à la mesure de Lebesgue, *i.e.* $\mathrm{d}\mu(x) = f(x)\mathrm{d}x$, alors $\widehat{\mu} = \widehat{f}$.

Définition. Si μ et ν sont deux mesures positives finies sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, alors on définit leur *produit de convolution* par

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \mu * \nu(A) = \int_{(R^d)^2} \mathbf{1}_A(x+y)\,\mathrm{d}\mu(x)\mathrm{d}\nu(y).$$

C'est la mesure image de $(x, y) \mapsto x + y$ par $(\mu \otimes \nu)$.

Proposition. Avec les notations précédentes, on a la relation $\widehat{\mu * \nu} = \widehat{\mu}\widehat{\nu}$.

Proposition. (Inversion de Fourier pour les mesures positives)
Pour tous les réels $a < b$, et toute mesure μ positive finie, on a :

$$\frac{1}{2}\mu(\{a,b\}) + \mu(]a,b[) = \lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-ita} - \mathrm{e}^{-itb}}{\mathrm{i}t}\widehat{\mu}(-t)\,\mathrm{d}t,$$

où φ est la fonction caractéristique associée à la mesure μ.

Corollaire. Si, pour une variable aléatoire X, on note $\varphi_X : t \mapsto \mathbb{E}(\mathrm{e}^{\mathrm{i}tX}) = \widehat{\mathbb{P}_X}(-t)$, et que l'on appelle cette fonction la *fonction caractéristique de* X, on voit que la fonction caractéristique caractérise la loi de X.

Exemple. Si X suit une loi normale $\mathcal{N}(m, \sigma^2)$, alors $\varphi_X(t) = \mathrm{e}^{\mathrm{i}mt - t^2\sigma^2/2}$ pour tout $t \in \mathbb{R}$.

Théorème. (Cramér)
Si X et Y sont deux variables aléatoires indépendantes à densité par rapport à la mesure de Lebesgue dont la somme est une loi normale, alors les deux sont de loi normale.

Proposition. Soit X une variable aléatoire réelle et φ_X sa fonction caractéristique.

— Si X admet un moment d'ordre $n \in \mathbb{N}^*$, alors φ_X est $\mathcal{C}^n$, et pour $k \leq n$:

$$\varphi_X^{(k)}(t) = \mathrm{i}^k \int_{\mathbb{R}} x^k \mathrm{e}^{\mathrm{i}tx} \,\mathrm{d}\,\mathbb{P}_X(x).$$

En particulier, $\varphi_X^{(k)}(0) = \mathrm{i}^k\,\mathbb{E}(X)^k$.

— Si est k fois dérivable en 0, $k \geq 2$, X admet des moments jusqu'à l'ordre $2\lfloor \frac{k}{2} \rfloor$, donnés par la formule précédente.

Théorème. Soit X une variable aléatoire réelle et φ_X sa fonction caractéristique. On suppose que X admet des moments à tout ordre, notés $(m_n)_n$. Si de plus, $\limsup_n \frac{m_n^{1/n}}{n} < +\infty$, alors X est caractérisée par ses moments.

Théorème. (Paul Lévy)
Avec le mêmes notations, la suite (X_n) converge vers X si et seulement si la fonction caractéristique φ_{X_n} converge simplement vers $\varphi(X)$.

Théorème. (Théorème limite central)
Soit (X_n) une suite de variables L^2 i.i.d d'espérance commune μ, et de variance commune σ^2. Alors $Z_n := \sqrt{n}\frac{\overline{X_n} - \mu}{\sigma}$ converge en loi vers une loi $\mathcal{N}(0, 1)$.

4 Asymptotique d'intégrale à paramètres

4.1 Méthode de Laplace

La méthode de Laplace est une façon de calculer des équivalents d'une intégrale à paramètre. Plus précisément, on cherche à calculer un équivalent en l'infini de fonctions de la forme :

$$F(s) = \int_a^b \mathrm{e}^{-s\varphi(t)} f(t) \,\mathrm{d}t,$$

avec φ une fonction régulière, et f continue en a qui ne s'annule pas en a.

Proposition. (Méthode de Laplace au premier ordre)
Si φ' est strictement positive sur $[a, b[$ et si f est intégrable, alors :

$$F(s) \sim \frac{1}{\varphi'(a)} \frac{\mathrm{e}^{-s\varphi(a)} f(a)}{s}.$$

Proposition. (Méthode de Laplace au second ordre)
Dans le cas où $\varphi' > 0$ sur $]a, b[$, $\varphi'(a) = 0$ et $\varphi''(a) > 0$, on a l'équivalent :

$$F(s) \sim \sqrt{\frac{\pi}{2\varphi''(a)}} \frac{\mathrm{e}^{-s\varphi(a) f(a)}}{\sqrt{s}}.$$

Application. (Formule de Stirling)
Quand x tend vers l'infini :

$$\Gamma(x+1) = \int_0^\infty t^x \mathrm{e}^{-t} \,\mathrm{d}t \sim \frac{x^x}{e^x} \sqrt{2\pi x}.$$

Théorème. Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\| \to \infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x) \mathrm{e}^{-t\varphi(x)} \,\mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}} \left(\frac{2\pi}{t}\right)^{d/2} \mathrm{e}^{-t\varphi(x_0)}.$$

4.2 Théorème limite central pour les densités

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3 - 3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

Chapitre 59

241. Suites et séries de fonctions. Exemples et contre-exemples.

Sommaire

1 Convergence de suites et séries de fonction 419
1.1 Modes de convergence 419
1.2 Régularité de la limite 420
1.3 Séries de variables aléatoires indépendantes 421
2 Séries entières 421
2.1 Rayon de convergence 421
2.2 Convergence sur le disque fermé 423
2.3 Fonctions entières 423
3 Séries de Fourier 423
3.1 Définitions 423
3.2 Convergence de la série de Fourier 424
3.3 Formule sommatoire de Poisson 425

Présentation du plan

On considère dans tout le plan (E, d) un ensemble, éventuellement muni d'une distance lorsque l'on parle de continuité.

1 Convergence de suites et séries de fonction

1.1 Modes de convergence

Définition. Une suite de fonctions $(f_n)_n$ converge *simplement* vers une fonction $f : E \to \mathbb{C}$ lorsque pour tout $x \in \mathbb{E}$, la suite $(f_n(x)) \in \mathbb{C}^{\mathbb{N}}$ converge dans $\mathbb{C}$. Une série de fonctions converge simplement lorsque la suite de ses sommes partielles converge simplement.

Définition. Une suite de fonctions $(f_n)_n$ converge *uniformément* (sur E) vers une fonction

$f : E \to \mathbb{C}$ lorsque la quantité :

$$\|f - f_n\|_\infty := \sup_{x \in E} |f(x) - f_n(x)|$$

tend vers 0 quand n tend vers l'infini. Une série de fonctions converge uniformément lorsque la suite de ses sommes partielles converge uniformément.

Remarque. La fonction $\|\cdot\|_\infty$ définit une norme sur l'ensemble des fonctions de E dans $\mathbb{C}$.

Remarque. La convergence uniforme entraîne la convergence simple.

Exemple. La suite de fonctions $(x \mapsto \frac{x}{n})_n \geq 1$ converge simplement vers 0 sur $\mathbb{R}$, mais la convergence n'y est pas uniforme.

Proposition. Si E est indénombrable, alors la convergence simple n'est pas métrisable sur l'ensemble des fonctions de E dans $\mathbb{C}$.

Définition. Une série de fonctions $(f_n)_n$ converge *normalement* lorsque la série $\sum_{n \in \mathbb{N}} \|f_n\|_\infty$ converge.

Proposition. La convergence normale entraîne la convergence uniforme.

Exemple. La série de fonctions $\sum_{n \in \mathbb{N}} (s \mapsto \frac{1}{n^s})$, $s > 1$, converge normalement sur tout intervalle $[\alpha, +\infty[$, avec $\alpha > 1$. On peut donc définir sa somme $\zeta :]1, +\infty[\to \mathbb{R}$.

Exemple. La série de fonctions $\sum_{n \in \mathbb{Z}} (x \mapsto \mathrm{e}^{-\pi n^2 x})$, $x > 0$, converge normalement sur tout intervalle $[\alpha, +\infty[$, avec $\alpha > 0$. On note sa somme $\vartheta : \mathbb{R}_+^* \to \mathbb{R}$.

1.2 Régularité de la limite

Proposition. Si E est compact (ou bien que l'on ne regarde que les fonctions bornées), la norme $\|\cdot\|_\infty$ fait de l'ensemble des fonctions continues de E dans $\mathbb{C}$ un espace de Banach. En particulier, une limite uniforme de fonctions continue est continue.

Proposition. Si $(f_n : [a, b] \to E)_n$, E Banach, converge uniformément vers f, alors la suite des primitives converge uniformément sur la primitive de f correspondante. En particulier, on peut intervertir limite et intégrale.

Proposition. Soit I un intervalle de $\mathbb{R}$. Si une suite de fonctions $(f_n) \in \mathcal{C}^k$ converge uniformément vers une fonction f et que pour tout $j \in \{1, \ldots, k\}$, la suite $(f_n^{(j)})_n$ converge uniformément vers une fonction f_j, alors f est $\mathcal{C}^k$ et pour tout $j \in \{1, \ldots, k\}$, $f^{(j)} = f_j$.

Exemple. Les fonctions ζ et ϑ sont $\mathcal{C}^\infty$ sur leurs domaines respectifs de définition.

Exemple. La suite de fonctions $(x \mapsto \sqrt{x^2 + \frac{1}{n^2}})_n$ est une suite de fonctions dérivables qui converge uniformément vers la fonction valeur absolue, qui n'est pas dérivable.

Proposition. On considère g la fonction 4-périodique affine par morceaux définie sur $\mathbb{R}$ par $f(-2) = -1, f(-1) = 0, f(0) = 1, f(1) = 0$. On considère alors :

$$f = \sum_{n=0}^{\infty} \frac{g(2^{2^n} \cdot)}{2^n}.$$

La fonction f définit une fonction continue sur $\mathbb{R}$ mais nulle part dérivable.

Proposition. Soit $U \subset \mathbb{C}$ un ouvert et (f_n) une suite de fonctions holomorphes définies sur U. Si $(f_n)_n$ converge uniformément vers f sur tout compact de U, la limite f est holomorphe.

Exemple. La fonction ζ définit une fonction holomorphe sur $\{z \in \mathbb{C}, \mathrm{Re}(z) > 1\}$.

1.3 Séries de variables aléatoires indépendantes

On s'intéresse à des séries de la forme $\sum_n X_n$, où les (X_i) sont des variables aléatoires indépendantes et L^2. On définit par (S_n) la suite des sommes partielles, et on note $\sigma_i^2 = \mathrm{Var}(X_i)$.

Théorème. (inégalité maximale de Kolmogorov) : On suppose les (X_i) centrées. Alors pour $\varepsilon > 0$ et pour tout n :

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\sum_{i=1}^{n} \sigma_i^2.$$

Proposition. Si les (X_i) sont centrées, et si la série des variances converge, alors la série $\sum_n X_n$ converge presque sûrement (et dans L^2).

Théorème. (loi des grands nombres avec un moment d'ordre deux)
Si $\mathbb{E}(X_n)$ tend vers m et si $\sum_{j=1}^{+\infty} \frac{\sigma_j^2}{j^2} < \infty$ alors $\overline{X_n}$ converge presque sûrement et dans L^2 vers m.

2 Séries entières

2.1 Rayon de convergence

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(0,R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0,R)$. On appelle R le *rayon de convergence de la série entière* et $D(0,R)$ le *disque de convergence.*

Exemple. Soit $\mu = (\mu_n)$ une mesure de probabilité sur $\mathbb{N}$. On définit sa *série génératrice* G_μ par $G_\mu = \sum_n \mu_n z^n$. Alors le rayon de convergence de cette série est au moins 1.

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence. Elle est donc $\mathcal{C}^\infty$ sur l'intervalle de convergence.

Proposition. Une série entière est holomorphe sur son disque de convergence.

Exemple. La série $\sum_n z^n$ converge sur $D(0,1)$, mais pas sur tout le disque fermé. De plus, la convergence n'est pas uniforme sur le disque.

Exemple. La série exponentielle $\sum_n \frac{z^n}{n!}$ converge sur tout $\mathbb{C}$ vers une fonction notée exp. La fonction exp est alors holomorphe sur $\mathbb{C}$ avec $\exp' = \exp$.

Théorème. Soit $\sum_n a_n z n^n$ une série entière. On suppose que la suite (a_n) ne s'annule pas à partir d'un certain rang, et que $|\frac{a_{n+1}}{a_n}|$ converge vers une limite ℓ. Alors le rayon de convergence R de la série entière vaut $\frac{1}{\ell}$. Plus généralement, on a la formule suivante :

$$\frac{1}{R} = \limsup_{n\to+\infty} |a_n|^{\frac{1}{n}}.$$

Corollaire. Le rayon de convergence d'une série entière et de sa dérivée formelle est le même.

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha, \alpha[$. La série entière $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha, \alpha[$.

Exemple. On obtient plein de développements usuels.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

2.2 Convergence sur le disque fermé

Remarque. Sur le disque fermé de convergence, il peut tout se passer.

Proposition. Soit $\sum_n a_n z^n$ une série entière de rayon de convergence $R > 0$. On suppose que la somme $\sum_{n=0}^{+\infty} a_n R^n$ est finie. Alors la série converge uniformément sur $[0, R]$.

Exemple. Convergence de la série logarithme sur $[0, 1]$, et de l'arctangente.

Théorème. (de Tauber fort) Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1. On suppose que la somme admet une limite en 1 par valeurs inférieures. Si $a_n = \mathrm{O}(1/n)$, alors cette limite est la somme des (a_n).

Application. Si les coefficients de Fourier d'une fonction tendent vers 0 en $\mathrm{o}(1/|n|)$, alors f est somme de sa série de Fourier.

2.3 Fonctions entières

Définition. On appelle fonction entière toute fonction holomorphe définie sur $\mathbb{C}$ tout entier.

Théorème. Soit f une fonction entière. Alors f s'écrit comme une série entière au voisinage de chaque point, c'est à dire que pour tout $a \in \mathbb{C}$, il existe une suite de complexes $(a_n)_n$ telle que

$$\forall z \in \mathbb{C}, f(z) = \sum_{n=0}^{+\infty} a_n (z-a)^n.$$

On a de plus $a_n = \frac{f^{(n)}(a)}{n!}$.

Théorème. Toute fonction entière bornée est constante.

Corollaire. Le corps $\mathbb{C}$ est algébriquement clos.

Application. (Indécomposabilité de la loi de Poisson)
Soient X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ dont la somme suit une loi de Poisson. Alors X et Y suivent une loi de Poisson.

3 Séries de Fourier

3.1 Définitions

Définition. On note $e_k : \mathbb{R} \to \mathbb{C}$ définie par $e_k(t) = \mathrm{e}^{\mathrm{i}kt}$.
Pour une fonction $f \in \mathrm{L}^1_{2\pi}$ et $n \in \mathbb{Z}$, on définit le n-ième coefficient de Fourier par :

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e_{-n}(t)\, \mathrm{d}t \in \mathbb{C}.$$

Remarque. Pour $f \in \mathrm{L}^2_{2\pi} \subset \mathrm{L}^1_{2\pi}$, on a $c_n(f) = \langle f, e_n \rangle$ (en prenant par convention l'antilinéarité à droite).

Exemple. Coefficient de Fourier d'une porte et $c_n(e_k) = \delta_{k,n}$. Coefficient de Fourier de $\cos(\alpha t)$.

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1_{2\pi}$. Alors la suite de ses coefficients de Fourier tend vers 0 en $\pm\infty$.

Théorème. On peut définir $F : f \mapsto (c_n(f))_n$ de $\mathrm{L}^1_{2\pi}$ dans $c_0(\mathbb{Z}, \mathbb{C})$. C'est une application linéaire de norme d'opérateur 1.

Définition. On définit la *somme partielle de Fourier* par $S_N(f) \sum_{n=-N}^{N} c_n(f) e_n$ et la série de Fourier (convergente ou non, c'est une série de fonctions) $S(f) = \sum_{n \in \mathbb{Z}} c_n(f) e_n$.

3.2 Convergence de la série de Fourier

Définition. Pour $N \in \mathbb{N}^*$, on définit le *noyau de Dirichlet* :

$$D_N(t) = \sum_{n=-N}^{N} \mathrm{e}^{\mathrm{i}nt} = \frac{\sin((N + 1/2)t)}{\sin(t/2)}$$

ainsi que le *noyau de Fejér* :

$$K_N(t) = \frac{1}{N} \sum_{n=0}^{N-1} D_n(t) = \frac{\sin(Nt/2)^2}{N \sin(t/2)^2}.$$

Ainsi, les sommes partielles de Fourier s'écrivent : $S_N(f) = (D_N * f)$, et leur moyenne de Cesàro s'écrit $(K_N * f)$. On montre facilement que K_N est une approximation de l'unité sur $[0, 2\pi]$ (en fait, elle converge au sens des distributions vers $\delta_{\mathbb{Z}}$).

Théorème. (de Fejér)
Soit f une fonction continue et 2π-périodique, alors $(K_N * f)_N$ converge uniformément vers f.

Remarque. Si la série de Fourier d'une fonction continue converge, c'est forcément vers la fonction.

Théorème. (de Dirichlet)
Soit $f \in \mathrm{L}^1_{2\pi}$ et $x_0 \in \mathbb{R}$. On suppose que f admet des limites à gauche et à droite en x_0, et que ses taux d'accroissements à gauche et à droite en x_0 sont intégrables au voisinage de x_0. Alors la série de Fourier de f converge en x_0 avec pour limite $\frac{f(x_0^+)+f(x_0^-)}{2}$.

Corollaire. Si f est continue et $\mathcal{C}^1$ par morceaux, f est somme de sa série de Fourier.

Application. $\zeta(2) = \frac{\pi^2}{6}$.

Remarque. La continuité de f ne suffit pas. Il existe des fonctions continues dont la série de Fourier diverge en 0.

Remarque. Si la série des coefficients de Fourier de f converge, alors la série de Fourier converge normalement.

Proposition. Si f est continue et $\mathcal{C}^1$ par morceaux, alors sa série de Fourier converge normalement vers elle.

Application. Pour tout $x \in \mathbb{R}\backslash\mathbb{Z}$:

$$\cot(\pi x) = \frac{1}{\pi x} + \sum_{n=1}^{+\infty} \frac{2x}{\pi(x^2 - n^2)}.$$

Théorème. (factorisation du sinus)
Pour tout $t \in]-1, 1[$:

$$\frac{\sin(\pi t)}{\pi t} = \prod_{k=1}^{\infty} \left(1 - \frac{t^2}{k^2}\right).$$

3.3 Formule sommatoire de Poisson

On note $\mathcal{S}(\mathbb{R})$ la classe de Schwartz et on prend pour conventions :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) \mathrm{e}^{-\mathrm{i}\xi x} \, \mathrm{d}\xi$$

si f est dans L^1.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \widehat{f}(n) \mathrm{e}^{\mathrm{i}nx}.$$

Remarque. On peut appliquer cette formule sous des hypothèses plus faibles. En effet, la seule chose dont on a vraiment besoin est le caractère $\mathcal{C}^1$ de la somme, c'est-à-dire la convergence normale de la série de fonctions, pour pouvoir appliquer le théorème de Dirichlet. On peut donc se restreindre aux hypothèses suivantes :

— $f \in \mathcal{C}^0 \cap \mathrm{L}^1$ et $|f| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1$

— $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty$ ou bien $|f'| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1$.

Application. (Fonction ϑ de Jacobi)
On pose $\vartheta : x \mapsto \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 x}$, pour $x > 0$. La fonction ϑ est solution de l'équation fonctionnelle suivante sur $\mathbb{R}_+^*$:

$$\vartheta(x) = \frac{1}{\sqrt{x}}\vartheta\left(\frac{1}{x}\right).$$

Chapitre 60

243. Séries entières, propriétés de la somme. Exemples et applications.

Sommaire

1 Définition d'une série entière 427
1.1 Rayon de convergence 427
1.2 Régularité dans le disque de convergence 428
1.3 Au bord du disque 428
2 Développement en série entière 428
2.1 Fonction holomorphes 428
2.2 Fonctions de la variable réelle 429
3 Série génératrice 430
3.1 Nombres de Catalan 430
3.2 Cas des probabilités 430

Présentation du plan

1 Définition d'une série entière

1.1 Rayon de convergence

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n z^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(0,R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0,R)$. On appelle R le *rayon de convergence de la série entière* et $D(0,R)$ le *disque de convergence.*

Théorème. Soit $\sum_n a_n zn^n$ une série entière. On suppose que la suite (a_n) ne s'annule pas à partir d'un certain rang, et que $|\frac{a_{n+1}}{a_n}|$ converge vers une limite ℓ. Alors le rayon de convergence R de la série entière vaut $\frac{1}{\ell}$. Plus généralement, on a la formule suivante :

$$\frac{1}{R} = \limsup_{n\to+\infty} |a_n|^{\frac{1}{n}}.$$

Corollaire. Le rayon de convergence d'une série entière et de sa dérivée formelle est le même.

1.2 Régularité dans le disque de convergence

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence. Elle est donc $\mathcal{C}^\infty$ sur l'intervalle de convergence.

Proposition. Une série entière est holomorphe sur son disque de convergence.

Exemple. La série $\sum_n z^n$ converge sur $D(0,1)$, mais pas sur tout le disque fermé. De plus, la convergence n'est pas uniforme sur le disque.

Exemple. La série exponentielle $\sum_n \frac{z^n}{n!}$ converge sur tout $\mathbb{C}$ vers une fonction notée exp. La fonction exp est alors holomorphe sur $\mathbb{C}$ avec $\exp' = \exp$.

1.3 Au bord du disque

Remarque. Sur le disque fermé de convergence, il peut tout se passer.

Proposition. Soit $\sum_n a_n z^n$ une série entière de rayon de convergence $R > 0$. On suppose que la somme $\sum_{n=0}^{+\infty} a_n R^n$ est finie. Alors la série converge uniformément sur $[0, R]$.

Exemple. Convergence de la série logarithme sur $[0,1]$, et de l'arctangente.

Théorème. (de Tauber fort) Soit $\sum_n a_n z^n$ une série entière de rayon de convergence 1. On suppose que la somme admet une limite en 1 par valeurs inférieures. Si $a_n = \mathrm{O}(1/n)$, alors cette limite est la somme des (a_n).

2 Développement en série entière

2.1 Fonction holomorphes

Théorème. Soit f une fonction holomorphe définie sur un voisinage de $a \in \mathbb{C}$. Alors f est développable en série entière au voisinage de a : il existe un disque D centré en a tel que pour tout $z \in D$, $f(z) = \sum_n a_n (z-a)^n$. Si $a = 0$, on a alors pour r assez petit :

$$a_n r^n = \frac{1}{2\pi} \int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta}) \mathrm{e}^{-\mathrm{i}n\theta} \, \mathrm{d}\theta.$$

Théorème. (Principe des zéros isolés)
Soit $f \in \mathcal{H}(\Omega)$. L'ensemble des zéros de f est isolé dans Ω.

Corollaire. (Principe du prolongement analytique)
Si deux fonctions holomorphes sur Ω sont égales sur une partie qui possède un point d'accumulation dans Ω, alors elles sont égales sur Ω.

Exemple. Il existe une unique fonction holomorphe qui prolonge Γ à $\mathbb{C}\backslash\mathbb{Z}$.

Exemple. Une fonction de la classe de Schwartz non nulle ne peut pas être à la fois à support compact et de transformée de Fourier à support compact.

Définition. Une fonction est dite *entière* si elle est holomorphe sur $\mathbb{C}$ ou, de manière équivalente, si elle est développable en série entière sur $\mathbb{C}$ avec un rayon de convergence infini.

Théorème. (de Liouville)
Toute fonction entière bornée est constante. De plus, si une fonction entière est à croissance polynomiale de degré au plus k, elle est de degré au plus k.

Application. (Théorème de d'Alembert-Gauss)
Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine complexe.

Théorème. (de la partie réelle d'Hadamard)
Soit $f = u + \mathrm{i}v = \sum_{a_n} z^n$ une fonction entière et

$$A(r) := \sup_{|z|=r} u(z).$$

Alors pour $n \geq 1$ et $r > 0$, $|a_n| \leq 2\frac{A(r)-u(0)}{r^n}$. Si de plus $A(r) = O_{+\infty}(r^d)$ pour $d \in \mathbb{N}$ alors f est un polynôme de degré inférieur ou égal à d.

2.2 Fonctions de la variable réelle

Théorème. (développement en série entière)
Soit I un intervalle de $\mathbb{R}$ contenant un voisinage de 0, et une fonction $f : I \to \mathbb{R}$ de classe $\mathcal{C}^\infty$. S'il existe un $\alpha > 0$ tel que la suite de fonctions :

$$R_n : x \mapsto f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-\alpha, \alpha[$. La série entière$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$ a alors un rayon de convergence plus grand que α et f coïncide avec la somme sur $]-\alpha, \alpha[$.

Exemple. On obtient plein de développements usuels.

Exemple. La fonction $\mathrm{e}^{-1/x}$ prolongée par 0 sur $\mathbb{R}_-$ est $\mathcal{C}^\infty$ sur $\mathbb{R}$ mais développable en série entière sur aucun voisinage de 0. En effet, sinon elle serait nulle sur un voisinage de 0.

3 Série génératrice

3.1 Nombres de Catalan

Définition. Pour $(a_n)_n$ une suite de réels, on définit sa *série génératrice* comme la série entière $\sum a_n z_n$.

Remarque. Introduire la série génératrice permet de ramener l'étude à celle d'une fonction. On cherche à transférer les propriétés de la suite sur des propriétés plus simples à manipuler de la fonction. Une fois une formule explicite de la fonction trouvée, on récupère la suite en développant en série entière.

Définition. Un mot de Dyck est un mot écrit sur l'alphabet $\{(,)\}$ qui est bien parenthésé, c'est-à-dire tel qu'aucun de ses préfixes ne contienne plus de parenthèses fermantes que de parenthèses ouvrantes. On note C_n le nombre de mots de Dyck de taille $2n$.

Proposition. Le nombre C_n est égal à $\binom{2n}{n} - \binom{2n}{n-1}$.

3.2 Cas des probabilités

Définition. La série génératrice G_X d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est naturellement définie comme la série génératrice de sa loi. Elle est alors définie au moins sur le disque unité ouvert et vérifie $G_X(1) = 1$. Quand c'est défini, $G_X(z) = \mathbb{E}(z^X)$.

Proposition. Si deux variables sont indépendantes, la série génératrice de leur somme correspond au produit de leurs séries génératrices.

Exemple. La série génératrice d'une loi de Poisson $\mathcal{P}(\lambda)$ est définie sur $\mathbb{C}$ par $G_{\mathcal{P}(\lambda)}(z) = \mathrm{e}^{\lambda(z-1)}$. En particulier, une somme de lois de Poisson suit une loi de Poisson.

Théorème. Soient X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ dont la somme suit une loi de Poisson. Alors X et Y suivent une loi de Poisson.

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

Chapitre 61

245. Fonctions holomorphes et méromorphes sur un ouvert de $\mathbb{C}$. Exemples et applications.

Sommaire

1 Fonctions holomorphes 433
1.1 Définition et premier exemple 433
1.2 Intégration sur un contour 434
1.3 Théorie de Cauchy 435
2 Conséquences du théorème de Cauchy 435
2.1 Principe des zéros isolés 435
2.2 Fonctions entières 436
2.3 Principe du maximum 436
3 Fonctions méromorphes 437
3.1 Singularités 437
3.2 Fonctions méromorphes et calcul des résidus 437

Présentation du plan

1 Fonctions holomorphes

1.1 Définition et premier exemple

Définition. Soit f une fonction complexe définie sur Ω. Si $z_0 \in \Omega$ et si

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existe on note ce nombre $f'(z_0)$ et on l'appelle *dérivée* de f en z_0. Si $f'(z_0)$ existe sur tout Ω, on dit que f est holomorphe sur Ω et on note $f \in \mathcal{H}(\Omega)$.

Remarque. L'holomorphie et la différentiabilité des deux variables sont deux notions différentes. La conjugaison est différentiable mais n'est pas holomorphe.

Proposition. On peut sommer, composer, multiplier des fonctions holomorphes et le résultat sera aussi holomorphe.

Définition. On appelle *série entière* une série de fonctions de la forme $\sum_{n\in\mathbb{N}} a_n(z-z_0)^n$, où $(a_n) \in \mathbb{C}^{\mathbb{N}}$.

Proposition. Soit $\sum_{n\in\mathbb{N}} a_n(z-z_0)^n$ une série entière et $R = \sup\{\rho \geq 0, (a_n\rho^n)_n \text{ est bornée}\}$. Alors la série entière converge simplement sur $D(z_0, R)$ et diverge grossièrement en dehors du disque fermé $\overline{D}(0, R)$. On appelle R le *rayon de convergence de la série entière* et $D(z_0, R)$ le *disque de convergence*.

Exemple. Soit $\mu = (\mu_n)$ une mesure de probabilité sur $\mathbb{N}$. On définit sa *série génératrice* G_μ par $G_\mu = \sum_n \mu_n z^n$. Alors le rayon de convergence de cette série est au moins 1.

Proposition. Une série entière converge normalement sur tout compact de son disque de convergence.

Théorème. Une série entière est holomorphe sur son disque de convergence, et la dérivée coïncide avec la dérivée formelle de la série.

1.2 Intégration sur un contour

Définition. Soit I un intervalle de $\mathbb{R}$ et $\gamma : I \to \mathbb{C}$ une fonction lisse (on dit que γ est un *chemin lisse*, si I est fermé et que γ prend la même valeur aux deux bords, on parle de chemin fermé). Pour f une fonction continue sur $\gamma(I)$, on pose :

$$\int_\gamma f = \int_\gamma f(\zeta)\,\mathrm{d}\zeta := \int_I f(\gamma(t))\gamma'(t)\,\mathrm{d}t.$$

Théorème. Soit $\gamma : I \to \mathbb{C}$ un chemin fermé, Ω le complémentaire de $\gamma(I)$ dans $\mathbb{C}$. On définit l'*indice* de γ par rapport au point $z \in \mathbb{C}$:

$$\mathrm{Ind}(\gamma, c) = \frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{\mathrm{d}\zeta}{\zeta - z}.$$

La fonction $\mathrm{Ind}(\gamma, \cdot)$ est alors une fonction à valeurs entières sur Ω, constante sur chaque composante connexe, et nulle sur sa composante connexe non bornée.

Remarque. L'indice permet de compter le nombre de tours (algébriquement) effectués dans le sens direct autour du point z par γ.

1.3 Théorie de Cauchy

Proposition. Soit $F \in \mathcal{H}(\Omega)$ telle que F' est continue sur Ω. Alors pour tout chemin fermé tracé dans Ω :

$$\int_\gamma F'(\zeta)\,\mathrm{d}\zeta = 0.$$

Théorème. (lemme de Goursat)
Soit Δ un triangle fermé inclus dans un ouvert Ω du plan. Soit z_0 dans Ω et soit f continue sur Ω et holomorphe sur $\Omega\backslash\{z_0\}$.

Théorème. (de Cauchy)
Soit Ω un ouvert convexe et f holomorphe sur Ω. Alors pour tout chemin fermé tracé dans Ω :

$$\int_\gamma f = 0.$$

Théorème. (formule de Cauchy)
Soit $\gamma : I \to \mathbb{C}$ un chemin fermé dans un ouvert convexe Ω, et f holomorphe sur Ω. Pour $z \in \Omega\backslash\gamma(I)$:

$$f(z)\,\mathrm{Ind}_\gamma(z) = \frac{1}{2\mathrm{i}\pi}\int_\gamma \frac{f(\zeta)}{\zeta - z}\,\mathrm{d}\zeta.$$

Exemple. On peut calculer $\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x$ avec la formule de Cauchy.

Corollaire. Pour tout ouvert du plan Ω, toute fonction holomorphe sur Ω y est développable en série entière (au voisinage de chaque point).

Théorème. (de Morera)
Toute fonction qui vérifie le lemme de Goursat pour tout triangle est holomorphe.

Théorème. Soit Ω un ouvert de $\mathbb{C}$. On s'intéresse à $F(x) = \int_\Omega f(x,\omega)\,\mathrm{d}\mu$, avec $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. On suppose que pour presque tout $\omega \in \Omega$, $z \mapsto f(z,\omega)$ est holomorphe sur U, et qu'il existe, pour tout compact K de Ω, $g_K \in \mathrm{L}^1(\mu)$ positive telle que l'on ait presque partout (sur K) $|f(z,\cdot)| \leq g$. Alors F est holomorphe sur Ω.

2 Conséquences du théorème de Cauchy

2.1 Principe des zéros isolés

Théorème. (Principe des zéros isolés)
Soit $f \in \mathcal{H}(\Omega)$. L'ensemble des zéros de f est isolé dans Ω.

Corollaire. (Principe du prolongement analytique)
Si deux fonctions holomorphes sur Ω sont égales sur une partie qui possède un point d'accumulation dans Ω, alors elles sont égales sur Ω.

Exemple. Il existe une unique fonction holomorphe qui prolonge Γ à $\mathbb{C}\backslash\mathbb{Z}$.

Exemple. Une fonction de la classe de Schwartz non nulle ne peut pas être à la fois à support compact et de transformée de Fourier à support compact.

2.2 Fonctions entières

Définition. Une fonction est dite *entière* si elle est holomorphe sur $\mathbb{C}$ ou, de manière équivalente, si elle est développable en série entière sur $\mathbb{C}$ avec un rayon de convergence infini.

Théorème. (de Liouville)
Toute fonction entière bornée est constante. De plus, si une fonction entière est à croissance polynomiale de degré au plus k, elle est de degré au plus k.

Application. (Théorème de d'Alembert-Gauss)
Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine complexe.

Théorème. (de la partie réelle d'Hadamard)
Soit $f = u + \mathrm{i}v = \sum_{a_n} z^n$ une fonction entière et

$$A(r) := \sup_{|z|=r} u(z).$$

Alors pour $n \geq 1$ et $r > 0$, $|a_n| \leq 2\frac{A(r)-u(0)}{r^n}$. Si de plus $A(r) = O_{+\infty}(r^d)$ pour $d \in \mathbb{N}$ alors f est un polynôme de degré inférieur ou égal à d.

Théorème. (de Raïkov)
Si $\mathbb{P}_1$ et $\mathbb{P}_2$ sont deux mesures de probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ telles que $\mathbb{P}_1 * \mathbb{P}_2 = \mathcal{P}(\lambda)$ pour un $\lambda > 0$, alors $\mathbb{P}_1$ et $\mathbb{P}_2$ sont des lois de Poisson. Plus simplement, si X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes dont la somme suit une loi de Poisson, alors X et Y suivent une loi de Poisson.

2.3 Principe du maximum

Théorème. (principe du maximum)
Soit $f \in \mathcal{H}(\Omega)$. À moins que f ne soit constante, $|f|$ ne possède de maximum local en aucun point de Ω.

Exemple. Il est possible qu'une fonction holomorphe soit bornée sur la frontière d'un ouvert non borné, mais ne soit pas bornée à l'intérieur, comme le montre l'exemple de $\exp \circ \exp$ sur la bande de partie réelle entre $\pm\frac{\pi}{2}$.

Théorème. Soit $(f_j) \in \mathcal{H}(\Omega)^{\mathbb{N}}$ qui converge uniformément vers f sur les compacts de Ω. Alors f est holomorphe sur Ω et $(f_j^{(n)})$ converge uniformément sur les compacts vers $(f^{(n)})$.

3 Fonctions méromorphes

3.1 Singularités

Définition. Si $a \in \Omega$ ouvert et $f \in \mathcal{H}(\Omega\backslash\{a\})$. Si f peut être prolongée en une fonction holomorphe sur Ω, on dit que a est une *singularité artificielle*.

Proposition. Soit $f \in \mathcal{H}(\Omega\backslash\{a\})$, qui est de plus bornée sur un voisinage de a. Alors f a une singularité artificielle en a.

Théorème. Si $f \in \mathcal{H}(\Omega\backslash\{a\})$, l'un des trois cas doit se produire.

— f a une singularité artificielle en a.

— Il existe des complexes $c_1, \ldots, c_m$, avec $c_m \neq 0$, tels que

$$f(z) - \sum_{k=1}^{m} \frac{c_k}{(z-a)^k}$$

ait une singularité artificielle en a. On dit que m est un *pôle d'ordre* m en a pour f. La somme ci-dessus est alors appelée *partie principale* de f en a.

— Pour $r > 0$ et $D(a,r) \subset \Omega$, $f(D(a,r)\backslash\{a\})$ est dense $\mathbb{C}$. On dit que a est une singularité essentielle en a.

3.2 Fonctions méromorphes et calcul des résidus

Définition. Une fonction f est *méromorphe* sur un ouvert Ω s'il existe un ensemble $A \subset \Omega$ tel que

— A n'a pas de point d'accumulation dans Ω.

— f est holomorphe sur $\Omega\backslash A$

— chaque point de A est un pôle de f.

Définition. Soit f une fonction méromorphe et a un pôle de f. On appelle *résidu* de f en a le coefficient c_1 du développement précédent.

Théorème. (des résidus)
Soit Ω un ouvert convexe et f une fonction méromorphe sur Ω, où A est l'ensemble des pôles de f, supposé fini. Soit γ une chemin fermé dont l'image ne rencontre pas A. Alors :

$$\sum_{a \in A} \mathrm{Ind}(\gamma, a)\mathrm{Res}(f, a) = \frac{1}{2\mathrm{i}\pi} \int_\gamma f.$$

Exemple. On peut calculer l'intégrale :

$$\int_{\mathbb{R}} \frac{\cos(x)}{1+x^2}\,\mathrm{d}x = \frac{\pi}{\mathrm{e}}.$$

Application. (formule des compléments) Pour tout $z \in \mathbb{C}\backslash\mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Chapitre 62

246. Séries de Fourier. Exemples et applications.

Sommaire

1 Coefficients et séries de Fourier 439
1.1 Définitions 439
1.2 Opérateur série de Fourier 440
2 Convergence de la série de Fourier 440
2.1 Au sens de Cesàro 440
2.2 Convergence en norme hilbertienne 441
2.3 Convergence simple, convergence normale 441
3 Utilisation des séries de Fourier 442
3.1 Calculs remarquables 442
3.2 Équation de la chaleur 443
3.3 Formule sommatoire de Poisson 443

Présentation du plan

1 Coefficients et séries de Fourier

1.1 Définitions

Définition. On note $e_k : \mathbb{R} \to \mathbb{C}$ définie par $e_k(t) = \mathrm{e}^{\mathrm{i}kt}$.
Pour une fonction $f \in \mathrm{L}^1_{2\pi}$ et $n \in \mathbb{Z}$, on définit le n-ième *coefficient de Fourier* par :

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e_{-n}(t) \, \mathrm{d}t \in \mathbb{C}.$$

Remarque. Pour $f \in \mathrm{L}^2_{2\pi} \subset \mathrm{L}^1_{2\pi}$, on a $c_n(f) = \langle f, e_n \rangle$ (en prenant par convention l'antilinéarité à droite).

Exemple. Coefficient de Fourier d'une porte et $c_n(e_k) = \delta_{k,n}$. Coefficient de Fourier de $\cos(\alpha t)$.

Proposition. Liste de propriétés des coefficients de Fourier :

— $c_n(\tau_a f) = \mathrm{e}^{\mathrm{i}na} c_n(f)$
— $c_n(e_k f) = c_{n-k}(f)$
— Si f est $\mathcal{C}^{k-1}$ et $\mathcal{C}^k$ par morceaux, $c_n(f^{(k)}) = (\mathrm{i}n)^k c_n(f)$.

1.2 Opérateur série de Fourier

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1_{2\pi}$. Alors la suite de ses coefficients de Fourier tend vers 0 en $\pm\infty$.

Théorème. On peut définir $F : f \mapsto (c_n(f))_n$ de $\mathrm{L}^1_{2\pi}$ dans $\mathrm{c}_0(\mathbb{Z}, \mathbb{C})$. C'est une application linéaire de norme d'opérateur 1.

Définition. Pour $f \in \mathrm{L}^p_{2\pi}(0, 2\pi)$ et $g \in \mathcal{C}^0([0, 2\pi])$, on définit le *produit de convolution (périodique)* de f et g par :

$$f * g : x \mapsto \frac{1}{2\pi} \int_0^{2\pi} f(t) g(x - t) \, \mathrm{d}t,$$

qui définit une fonction de $\mathrm{L}^p_{2\pi}$.

Remarque. Par Cauchy-Schwarz, on peut définir le produit de convolution entre L^2 et L^2, qui est alors dans L^2, et par Fubini entre L^1 et L^1, le résultat est alors dans L^1.

Proposition. Pour $f, g \in \mathrm{L}^1_{2\pi}$, $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$.

Définition. On définit la *somme partielle de Fourier* par $S_N(f) = \sum_{n=-N}^{N} c_n(f) e_n$ et la série de Fourier (convergente ou non, c'est une série de fonctions) $S(f) = \sum_{n \in \mathbb{Z}} c_n(f) e_n$.

2 Convergence de la série de Fourier

2.1 Au sens de Cesàro

Définition. Pour $N \in \mathbb{N}^*$, on définit le *noyau de Dirichlet* :

$$D_N(t) = \sum_{n=-N}^{N} \mathrm{e}^{\mathrm{i}nt} = \frac{\sin((N + 1/2)t)}{\sin(t/2)}$$

ainsi que le *noyau de Fejér* :

$$K_N(t) = \frac{1}{N} \sum_{n=0}^{N-1} D_n(t) = \frac{\sin(Nt/2)^2}{N \sin(t/2)^2}.$$

Ainsi, les sommes partielles de Fourier s'écrivent : $S_N(f) = (D_N * f)$, et leur moyenne de Cesàro s'écrit $(K_N * f)$. On montre facilement que K_N est une approximation de l'unité sur $[0, 2\pi]$ (en fait, elle converge au sens des distributions vers $\delta_{\mathbb{Z}}$).

Théorème. (de Fejér)
Soit f une fonction continue et 2π-périodique, alors $(K_N * f)_N$ converge uniformément vers f.

Remarque. Si la série de Fourier d'une fonction continue converge, c'est forcément vers la fonction en question.

Définition. Soit $(\rho_n)_n$ une suite de fonctions de $\mathrm{L}^1_{2\pi} \cap \mathcal{C}^0([0,2\pi])$. On dit que c'est une *approximation de l'unité* lorsque $\int \rho_n = 1$ pour tout n, et pour tout $\delta > 0$, $\lim_n \int_{|x|\geq\delta} \rho_n(x)\,\mathrm{d}x = 0$.

Théorème. Soit $(\rho_n)_n$ une approximation de l'unité, et $f \in \mathrm{L}^p_{2\pi}$. Alors, on a la convergence suivante dans $\mathrm{L}^p_{2\pi}$:

$$\rho_n * f \xrightarrow[n\to\infty]{} f.$$

Corollaire. (Fejér-Lebesgue)
Puisque le noyau de Fejér est une approximation de l'unité, les sommes de Fejér convergent dans L^p vers la fonction considérée. En particulier, les polynômes trigonométriques sont denses dans L^p.

2.2 Convergence en norme hilbertienne

Théorème. La famille $(e_k)_k$ est une base hilbertienne de $\mathrm{L}^2_{2\pi}$.

Corollaire. On a donc l'égalité de Bessel et la formule de Parseval, la série de Fourier converge dans L^2 et on a une isométrie linéaire bijective.

Exemple. $\zeta(4) = \frac{\pi^4}{90}$.

2.3 Convergence simple, convergence normale

Théorème. (de Dirichlet)
Soit $f \in \mathrm{L}^1_{2\pi}$ et $x_0 \in \mathbb{R}$. On suppose que f admet des limites à gauche et à droite en x_0, et que ses taux d'accroissements à gauche et à droite en x_0 sont intégrables au voisinage de x_0. Alors la série de Fourier de f converge en x_0 avec pour limite $\frac{f(x_0^+)+f(x_0^-)}{2}$.

Corollaire. Si f est continue et $\mathcal{C}^1$ par morceaux, f est somme de sa série de Fourier.

Application. $\zeta(2) = \frac{\pi^2}{6}$.

Remarque. La continuité de f ne suffit pas. On peut montrer qu'il existe une fonction continue dont la série de Fourier diverge en 0.

Remarque. Si la série des coefficients de Fourier de f converge, alors la série de Fourier converge normalement.

Proposition. Si f est continue et $\mathcal{C}^1$ par morceaux, alors sa série de Fourier converge normalement vers elle. Plus généralement, la convergence est d'autant plus rapide que la fonction est régulière.

Proposition. Grâce au théorème taubérien fort, on a la convergence normale de la série de Fourier vers la fonction dès que les coefficients de Fourier de f sont en $\mathrm{O}(1/|n|)$.

Lemme. (lemme de Lebesgue)
Soit $(V, \|.\|)$ un espace vectoriel normé, U un sous-espace vectoriel de V et P un projecteur linéaire sur U. Alors pour tout $v \in V$, on a

$$\|v - P(v)\| \leq (1 + \|P\|) \inf_{u\in U} \|v - u\|.$$

Proposition. (qualité de l'approximation trigonométrique)
Soit u une fonction continue 2π-périodique. Alors

$$\|u - K_n * u\|_\infty \leq 2\omega\left(\frac{\pi}{\sqrt{n}}\right),$$

où ω est le module de continuité.

3 Utilisation des séries de Fourier

3.1 Calculs remarquables

Application. Pour tout $x \in \mathbb{R}\backslash\mathbb{Z}$:

$$\cot(\pi x) = \frac{1}{\pi x} + \sum_{n=1}^{+\infty} \frac{2x}{\pi(x^2 - n^2)}.$$

Théorème. (factorisation du sinus)
Pour tout $t \in]-1, 1[$:

$$\frac{\sin(\pi t)}{\pi t} = \prod_{k=1}^{\infty}\left(1 - \frac{t^2}{k^2}\right).$$

Théorème. (intégrale de Fresnel)
Les intégrales $\int_{\mathbb{R}} \cos(2\pi u^2)\,\mathrm{d}t$ et $\int_{\mathbb{R}} \sin(2\pi u^2)\,\mathrm{d}t$ sont semi-convergentes, et valent toutes deux $\frac{1}{2}$.

3.2 Équation de la chaleur

Théorème. Soit E l'ensemble des fonctions continues de (t, x) sur $\mathbb{R}_+ \times \mathbb{R}$ et 2π-périodiques en la seconde variable. Soit f une fonction $\mathcal{C}^1$ et 2π-périodique. Il existe une unique solution au problème suivant :

- $u \in E \cap \mathcal{C}^2(\mathbb{R}_+^* \times \mathbb{R})$
- $\forall t > 0, \forall x \in \mathbb{R}$, $\partial_t(u)(t, x) = \partial_x^2 u(t, x)$
- $u(0, .) = f$.

3.3 Formule sommatoire de Poisson

On note $\mathcal{S}(\mathbb{R})$ la classe de Schwartz et on prend pour conventions :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)\mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}\xi$$

si f est dans L^1.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Remarque. On peut appliquer cette formule sous des hypothèses plus faibles. En effet, la seule chose dont on a vraiment besoin est le caractère $\mathcal{C}^1$ de la somme, c'est-à-dire la convergence normale de la série de fonctions, pour pouvoir appliquer le théorème de Dirichlet. On peut donc se restreindre aux hypothèses suivantes :

— $f \in \mathcal{C}^0 \cap \mathrm{L}^1$ et $|f| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1$

— $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty$ ou bien $|f'| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1.$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x} \widehat{f}(\xi)\, \mathrm{d}\xi.$$

Théorème. (Échantillonnage de Shannon)
Soit $f \in \mathcal{S}(\mathbb{R})$ telle que le support de $\widehat{f}$ soit inclus dans $[-F, F]$ avec $F > 0$ et $2F \leq 2\pi$, alors :

$$f = \sum_{k\in\mathbb{Z}} f(k)\, \mathrm{sinc}(\pi(\cdot - k)).$$

Application. (Fonction ϑ de Jacobi)
On pose $\vartheta : x \mapsto \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 x}$, pour $x > 0$. La fonction ϑ est solution de l'équation fonctionnelle suivante sur $\mathbb{R}_+^*$:

$$\vartheta(x) = \frac{1}{\sqrt{x}} \vartheta\left(\frac{1}{x}\right).$$

Chapitre 63

250. Transformation de Fourier. Applications.

Sommaire

1 Transformée de Fourier d'une fonction intégrable 445
1.1 Définition 445
1.2 Opérateur transformée de Fourier 446
1.3 Convolution 446
2 Classe de Schwartz et prolongement de la transformée de Fourier . 447
2.1 Classe de Schwartz et inversion de Fourier 447
2.2 Prolongement de la transformée de Fourier pour une fonction de carré intégrable 447
2.3 Localisation et principe d'incertitude 448
3 Transformée de Fourier de mesures 448
3.1 Mesures positives et fonction caractéristique 448
3.2 Problème des moments 449
3.3 Convergence en loi 449

Présentation du plan

1 Transformée de Fourier d'une fonction intégrable

1.1 Définition

Définition. Soit $f \in \mathrm{L}^1(\mathbb{R})$. On définit sa *transformée de Fourier* par l'expression :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\xi x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}$$

Exemple. Si $g : x \mapsto \mathrm{e}^{-\alpha x}$, pour $\alpha > 0$, alors $\widehat{g} : \xi \mapsto \sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\xi^2}{4\alpha}}$.

Remarque. Il arrive qu'une fonction soit dans $\mathrm{L}^1(\mathbb{R})$ sans que sa transformée de Fourier le soit pour autant. C'est le cas par exemple de la fonction porte $x \mapsto \mathbf{1}_{[-1/2,1/2]}$, donc la transformée

de Fourier est $\mathrm{sinc}(\pi\cdot) \notin \mathrm{L}^1$.

Remarque. Cette définition peut être étendue à $\mathrm{L}^1(\mathbb{R}^d)$ en considérant $\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\langle x,\xi\rangle} f(x)\,\mathrm{d}x$.

1.2 Opérateur transformée de Fourier

Théorème. (Riemann-Lebesgue)
Soit $f \in \mathrm{L}^1$, alors $\lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0$ et $\widehat{f} \in \mathcal{C}_0^0(\mathbb{R})$.

Théorème. On définit l'opérateur $\mathcal{F} : \mathrm{L}^1(\mathbb{R}) \to \mathcal{C}_0^0(\mathbb{R})$ par $\mathcal{F}(f) = \widehat{f}$. C'est un opérateur linéaire continu.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

- $\mathcal{F}(\check{f}) = \mathcal{F}\check{(}f)$.
- $\mathcal{F}(\overline{f}) = \overline{\mathcal{F}\check{(}f)}$.
- Pour $\lambda > 0$, $\mathcal{F}(f(\lambda\cdot))\frac{1}{\lambda}\mathcal{F}(f)\left(\frac{\cdot}{\lambda}\right)$
- $\mathcal{F}(\tau_a f) = \mathrm{e}^{-\mathrm{i}a\cdot}\mathcal{F}(f)$ et $\mathcal{F}(\mathrm{e}^{-\mathrm{i}a\cdot}) = \tau_a(\mathcal{F}(f))$, $a \in \mathbb{R}$.

Proposition. Soit $f \in \mathrm{L}^1(\mathbb{R})$.

- Si f est $\mathcal{C}^1$ sur $\mathbb{R}$, avec $f' \in \mathrm{L}^1$. Alors $\mathcal{F}(f') = \mathrm{i}\xi\mathcal{F}(f)$.
- Si xf est intégrable, alors $\mathcal{F}(f)$ est dérivable sur $\mathbb{R}$ avec $\mathcal{F}(f)' = -\mathrm{i}\mathcal{F}(xf)$.

Exemple. Cette propriété permet de recalculer la transformée de Fourier de la gaussienne.

1.3 Convolution

Définition. Soient f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{C}$. On dit qu'elles sont *convoluables* lorsque $t \mapsto f(x-t)g(t)$ est intégrable pour presque tout x. On définit alors le *produit de convolution* de f et g par la formule :

$$f * g(x) = \int_{\mathbb{R}} f(x-t)g(t)\,\mathrm{d}t.$$

Théorème. L'espace $(\mathrm{L}^1(\mathbb{R}), +, *)$ est une algèbre de Banach.

Proposition. La transformée de Fourier est un morphisme d'algèbres de Banach de $(\mathrm{L}^1(\mathbb{R}), +, *)$ dans $(\mathcal{C}_0^0(\mathbb{R}, +, \times))$.

2 Classe de Schwartz et prolongement de la transformée de Fourier

2.1 Classe de Schwartz et inversion de Fourier

Définition. La *classe de Schwartz* $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}$ dont toutes les dérivées sont à décroissance rapide.

Proposition. Les fonctions test sont dans la classe de Schwartz, qui est donc dense dans L^p, $p < \infty$.

Proposition. La classe de Schwartz est stable par transformée de Fourier.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x+2\pi n) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x}\widehat{f}(\xi)\,\mathrm{d}\xi.$$

Par dualité, cette formule est valable pour $f \in \mathrm{L}^1$ dès que $\mathcal{F}(f) \in \mathrm{L}^1$ (d'après la proposition suivante).

Théorème. Pour tout $(f,g) \in \mathcal{S}(\mathbb{R})^2$,

$$\int_{\mathbb{R}} \widehat{f}\overline{\widehat{g}} = 2\pi\int_{\mathbb{R}} f\overline{g} \text{ et } \|\widehat{f}\|_2 = \sqrt{2\pi}\|f\|_2.$$

2.2 Prolongement de la transformée de Fourier pour une fonction de carré intégrable

Théorème. (de Plancherel)
L'application $\frac{1}{\sqrt{2\pi}}\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ se prolonge de manière unique en une isométrie linéaire bijective. De plus, la relation d'inversion de Fourier est satisfaite presque partout.

Exemple. La transformée de Fourier du sinus cardinal est $\pi\mathbf{1}_{[-\pi,\pi]}$.

Remarque. En général, on n'a pas de formule explicite pour la transformée de Fourier dans L^2. Cependant, si la suite de fonctions $\xi \mapsto \int_{-n}^{n} f(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x$ admet une limite presque partout, c'est $\mathcal{F}(f)$ presque partout.

Remarque. Le premier opérateur de Fourier défini est injectif mais non surjectif.

2.3 Localisation et principe d'incertitude

Théorème. (Échantillonnage de Shannon)
Soit $f \in \mathcal{S}(\mathbb{R})$ telle que le support de $\widehat{f}$ soit inclus dans $[-F,F]$ avec $F > 0$ et $2F \leq 2\pi$, alors :

$$f = \sum_{k\in\mathbb{Z}} f(k)\,\mathrm{sinc}(\pi(\,\cdot - k)).$$

Théorème. Soit $f \in \mathcal{S}(\mathbb{R})$.

$$\|xf\|_2\|\xi\widehat{f}\|_2 \geq \sqrt{\frac{\pi}{2}}\|f\|_2^2.$$

Proposition. Soit f de carré intégrable et à support compact. Sa transformée de Fourier peut être prolongée à une fonction entière sur $\mathbb{C}$.

Corollaire. Avec les mêmes notations, si f est à support compact si et seulement si $f = 0$.

3 Transformée de Fourier de mesures

3.1 Mesures positives et fonction caractéristique

Définition. Soit μ une mesure positive finie sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. On définit sa transformée de Fourier $\widehat{\mu} \in \mathcal{C}^0(\mathbb{R})$ par

$$\widehat{\mu}(t) = \int_R \mathrm{e}^{-\mathrm{i}\langle t,x\rangle}\,\mathrm{d}\mu(x).$$

Proposition. Si μ est à densité par rapport à la mesure de Lebesgue, *i.e.* $\mathrm{d}\mu(x) = f(x)\mathrm{d}x$, alors $\widehat{\mu} = \widehat{f}$.

Définition. Si μ et ν sont deux mesures positives finies sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, alors on définit leur *produit de convolution* par

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \mu * \nu(A) = \int_{(R^d)^2} \mathbf{1}_A(x+y)\,\mathrm{d}\mu(x)\mathrm{d}\nu(y).$$

C'est la mesure image de $(x,y) \mapsto x+y$ par $(\mu \otimes \nu)$.

Proposition. Avec les notations précédentes, on a la relation $\widehat{\mu * \nu} = \widehat{\mu}\widehat{\nu}$.

Proposition. (Inversion de Fourier pour les mesures positives)
Pour tous les réels $a < b$, et toute mesure μ positive finie, on a :

$$\frac{1}{2}\mu(\{a,b\}) + \mu(]a,b[) = \lim_{T\to+\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{\mathrm{e}^{-ita} - \mathrm{e}^{-itb}}{\mathrm{i}t}\widehat{\mu}(-t)\,\mathrm{d}t.$$

Corollaire. Si, pour une variable aléatoire X, on note $\varphi_X : t \mapsto \mathbb{E}(\mathrm{e}^{\mathrm{i}tX}) = \widehat{\mathbb{P}_X}(-t)$, et que l'on appelle cette fonction la *fonction caractéristique de* X, on voit que la fonction caractéristique caractérise la loi de X.

Exemple. Si X suit une loi normale $\mathcal{N}(m,\sigma^2)$, alors $\varphi_X(t) = \mathrm{e}^{\mathrm{i}mt - t^2\sigma^2/2}$ pour tout $t \in \mathbb{R}$.

Théorème. (Cramér)
Si X et Y sont deux variables aléatoires indépendantes à densité par rapport à la mesure de Lebesgue dont la somme est une loi normale, alors les deux sont de loi normale.

3.2 Problème des moments

Définition. Soit X une variable aléatoire réelle de loi μ. On dit que X ou μ admet un *moment d'ordre* $k \in \mathbb{N}^*$ lorsque x^k est μ-intégrable.

Proposition. Soit X une variable aléatoire réelle et φ_X sa fonction caractéristique.

— Si X admet un moment d'ordre $n \in \mathbb{N}^*$, alors φ_X est $\mathcal{C}^n$, et pour $k \leq n$:

$$\varphi_X^{(k)}(t) = \mathrm{i}^k \int_{\mathbb{R}} x^k \mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}\,\mathbb{P}_X(x).$$

En particulier, $\varphi_X^{(k)}(0) = \mathrm{i}^k\,\mathbb{E}(X)^k$.

— Si est k fois dérivable en 0, $k \geq 2$, X admet des moments jusqu'à l'ordre $2\lfloor \frac{K}{2} \rfloor$, donnés par la formule précédente.

Théorème. Soit X une variable aléatoire réelle et φ_X sa fonction caractéristique. On suppose que X admet des moments à tout ordre, notés $(m_n)_n$. Si de plus, $\limsup_n \frac{m_n^{1/n}}{n} < +\infty$, alors X est caractérisée par ses moments.

3.3 Convergence en loi

Définition. Soit $(X_n)_n$ une suite de variables aléatoires réelles. On dit que (X_n) *converge en loi* vers une variable aléatoire X lorsque pour toute fonction continue bornée h, $\mathbb{E}(h(X_n))$ converge vers $\mathbb{E}(h(X))$ dans $\mathbb{R}$. Remarquons que c'est plutôt une convergence de mesure dont il s'agit.

Théorème. (Paul Lévy)
Avec le mêmes notations, la suite (X_n) converge vers X si et seulement si la fonction caractéristique φ_{X_n} converge simplement vers $\varphi(X)$.

Exemple. Soit (X_n) une suite indépendante de lois respectives $\mathcal{B}(n, \frac{\lambda}{n})$. Alors (X_n) converge en loi vers une variable dont la loi est une loi de Poisson de paramètre λ.

Lemme. Pour $n \in \mathbb{N}$ et $t > 0$:

$$\left| \mathrm{e}^{\mathrm{i}t} - 1 - \mathrm{i}t - \cdots - \frac{(\mathrm{i}t)^{n-1}}{(n-1)!} \right| \leq \frac{t^n}{n!}.$$

Lemme. Si X admet un moment d'ordre n alors sa fonction caractéristique φ est de classe $\mathcal{C}^n$, et

$$\varphi^{(n)}(\xi) = \mathrm{i}^n \int_{\mathbb{R}} x^n \mathrm{e}^{\mathrm{i}\xi x} \,\mathrm{d}\,\mathbb{P}_X(x).$$

Remarque. Lorsque $\xi = 0$, on retrouve les moments (à une constante près).

Corollaire. Avec les mêmes hypothèses et notations, on a le développement de Taylor suivant :

$$\left| \varphi(\xi + t) - \varphi(\xi) - \sum_{k=1}^{n-1} \frac{t^k}{k!} \varphi^{(k)}(\xi) \right| \leq M_n \frac{|t|^n}{n!}.$$

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1 + \ldots + X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\| f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3 - 3x)g \|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

Chapitre 64

253. Utilisation de la notion de convexité en analyse.

Sommaire

1 Ensembles convexes 451
1.1 Définitions et exemples 451
1.2 Projection sur un convexe fermé 452
1.3 Hyperplans d'appui, points extrémaux 452
2 Fonctions convexes 453
2.1 Définition et caractérisations 453
2.2 Cas de la variable réelle 453
2.3 Inégalités de convexité 454
3 Optimisation, points fixes 454
3.1 Optimisation d'une fonctionnelle convexe 454
3.2 Théorèmes de point fixe 455

On désignera par E un $\mathbb{R}$-espace vectoriel normé.

1 Ensembles convexes

1.1 Définitions et exemples

Définition. Une partie C de E est *convexe* lorsque tout segment dont les extrémités sont dans E est inclus dans E.

Exemple. Une boule et un demi-espace sont convexes. Un convexe est connexe par arcs donc connexe.

Exemple. Dans $\mathcal{M}_n(\mathbb{R})$, $\mathcal{S}_n^{++}(\mathbb{R})$ est convexe mais $\mathrm{GL}_n(\mathbb{R})$ ne l'est pas.

Proposition. Une intersection de parties convexes est convexe.

Proposition. L'intérieur et l'adhérence d'un convexe sont convexes.

1.2 Projection sur un convexe fermé

On se place dans H un espace de Hilbert.

Théorème. (projection sur un convexe fermé)
Soit H un espace de Hilbert et C un convexe fermé non vide de H.

— Pour tout $x \in H$, il existe un unique point $p_C(x) \in C$ qui réalise la distance à C

— Il est caractérisé par la propriété suivante :

$$p_C(x) \in C \text{ et } \mathrm{Re}(\langle x - p_C(x), y - p_C(x)\rangle) \forall y \in C$$

— p_C est une application 1-lipschitzienne

— Si C est un sous-espace vectoriel fermé de H, alors $p_C(x)$ est caractérisé par

$$p_C(x) \in C \text{ et } (x - p_C(x)) \perp F.$$

Remarque. Seule la complétude de C est nécessaire dans la preuve.

Proposition. L'espérance conditionnelle de $X \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ sachant la tribu $\mathcal{G}$ est définie comme la projection de X sur le convexe fermé $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ et vérifie alors, pour tout Y $\mathcal{G}$-mesurable :

$$\mathbb{E}(XY) = \mathbb{E}(Y\,\mathbb{E}(X|\mathcal{G})).$$

Théorème. (supplémentaire orthogonal)
Soit F un sous-espace vectoriel fermé de H un Hilbert. Alors H est somme directe orthogonale de F et $F^\perp$.

Corollaire. Un sous-espace vectoriel F est dense dans H si et seulement si $F^\perp = 0$.

Théorème. (représentation de Riesz)
Pour tout $\varphi \in H'$, où H est un espace de Hilbert, il existe un unique a dans H tel que $\varphi = \langle \cdot, a\rangle$.

1.3 Hyperplans d'appui, points extrémaux

On se place ici en dimension finie.

Proposition. Tout convexe fermé est strictement contenu dans un demi-espace.

Définition. Soit $C \subset \mathbb{R}^n$ un convexe non vide. On dit que $c \in C$ est un *point extrémal de* C si $C \backslash \{c\}$ est encore convexe, ce qui équivaut au fait que $\frac{x+y}{2} = c \Rightarrow x = y = c$ pour $x, y \in C$.
On dit qu'un hyperplan affine H est un *hyperplan d'appui* à C en $c \in C$ si $c \in H$ et si C est inclus dans un demi-espace formé par H.

Proposition. Soit C un convexe fermé non vide. Alors si $c \in \partial C$, il existe un hyperplan d'appui à C en c.

Proposition. Soit $a \in C$, si a est dans un hyperplan d'appui H de C, a est point extrémal de C si et seulement si a est point extrémal de $H \cap C$.

Théorème. (de Krein-Milman)
Soit C un convexe compact non vide de $\mathbb{R}^n$, alors il est enveloppe convexe de ses points extrémaux.

Remarque. Les points extrémaux se montrent utiles lorsqu'il s'agit d'étudier les isométries laissant invariant un solide, car on peut alors se ramener aux permutations des points extrémaux. Si ceux-ci sont en nombre fini, on se ramène alors à l'étude de sous-groupes de $\mathfrak{S}_n$, ce qui est souvent plus commode.

2 Fonctions convexes

2.1 Définition et caractérisations

Soit Ω un (ouvert quand c'est nécessaire) convexe de E un espace vectoriel normé.

Définition. On dit que $f : \Omega \to \mathbb{R}$ est convexe si pour tous $x, y \in E$, $t \in]0,1[$, $f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$. On dit qu'elle est strictement convexe si l'inégalité est stricte dès que $x \neq y$.

Remarque. On est donc ramenés à l'étude de fonctions convexes de la variables réelle.

2.2 Cas de la variable réelle

Définition. Soit I un intervalle de $\mathbb{R}$. On dit que $f : I \to \mathbb{R}$ est *convexe* si pour tout $t \in]0,1[$ et tout couple $(x, y) \in I$, on a $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Proposition. Soit $f : I \to \mathbb{R}$ une fonction convexe.

- f est lipschitzienne sur tout segment.
- f admet des dérivées à droite et à gauche en tout point.
- Les dérivées à gauche et à droite sont des fonctions croissantes.
- L'ensemble des points en lesquels f n'est pas dérivable est dénombrable et f est continue en dehors de ces points.

Proposition. Soit $f :]a, b[\to \mathbb{R}$. Si f est dérivable sur $]a, b[$, alors f est convexe si et seulement si f' est croissante sur $]a, b[$. Si f est deux fois dérivable, alors la convexité de f équivaut à la positivité de f''.

Exemple. La fonction logarithme est concave, le sinus aussi (sur $[0, \frac{\pi}{2}]$).

Proposition. Soit $f :]a, b[\to \mathbb{R}$. Si f est convexe et dérivable sur $]a, b[$, alors on a équivalence, pour $c \in]a, b[$, entre le fait que c soit un minimum de f et un point critique de f. En particulier, si f est strictement convexe elle admet au plus un minimum.

Exemple. La fonction exp est strictement convexe mais n'admet pas de minimum.

2.3 Inégalités de convexité

Proposition. Soit $x \in [0, \pi/2]$, alors $\frac{2}{\pi}x \leq \sin(x) \leq x$.

Application. La fonction sinus cardinal n'est pas intégrable sur $\mathbb{R}$. Son intégrale est donc semi-convergente.

Proposition. Soient $x_1, \dots, x_n$ des réels positifs, alors $\sqrt[n]{x_1 \cdots x_n} \leq \frac{1}{n} \sum_{i=1}^{n} x_i$.

Proposition. (inégalité de Jensen)
Soit (Ω, μ) un espace de probabilité, et g une fonction intégrable à valeurs dans un intervalle I de $\mathbb{R}$. Soit φ une fonction convexe de I dans $\mathbb{R}$. Alors

$$\varphi\left(\int_\Omega g \, \mathrm{d}\mu\right) \leq \int_\Omega \varphi \circ g \, \mathrm{d}\mu.$$

Proposition. (inégalité de Hölder)
Si $f \in \mathrm{L}^p$ et $g \in \mathrm{L}^q$ ($\frac{1}{p} + \frac{1}{q} = 1$), alors $fg \in \mathrm{L}^1$ avec $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

3 Optimisation, points fixes

3.1 Optimisation d'une fonctionnelle convexe

Proposition. Si f est convexe sur Ω différentiable et si $c \in \Omega$ alors

— c est un minimum de f si et seulement si c est un point critique de f

— Si f est strictement convexe, alors elle admet au plus un minimum sur Ω.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y \rangle \to \langle x, y \rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}_0^1(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens faible comme précédemment.

3.2 Théorèmes de point fixe

Théorème. (de Brouwer, ADMIS)
Soit f une fonction continue définie d'un convexe de $\mathbb{R}^n$ dans lui-même. Alors f admet un point fixe.

Remarque. Ce théorème peut servir pour montrer le théorème de Perron-Frobenius.

Proposition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Alors il possède un produit scalaire invariant si et seulement s'il est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$

Théorème. (point fixe de Kakutani)
Soit G un sous-groupe compact de $\mathrm{GL}(E)$, et K une partie convexe et compacte de E stable par G. Alors il existe u un point fixe sous G dans K.

Application. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Présentation du plan

Chapitre 65

261. Loi d'une variable aléatoire :caractérisations, exemples, applications.

Sommaire

1 Loi d'une variable aléatoire 457
1.1 Définition 457
1.2 Indépendance 458
1.3 Espérance 459
2 Caractérisation de la loi 460
2.1 Fonction de répartition 460
2.2 Fonction caractéristique 460
2.3 Série génératrice 461
2.4 Problème des moments 462
3 Loi uniforme sur les permutations 463
3.1 Nombre d'inversions 463
3.2 Nombre de cycles 463
3.3 Nombre de descentes 463

Présentation du plan

1 Loi d'une variable aléatoire

1.1 Définition

Définition. Soit X une variable aléatoire d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dans un espace probabilisable $(E, \mathcal{F})$. La *loi* de la variable aléatoire X est la mesure image de $\mathbb{P}$ par X, notée $\mathbb{P}_X$ et définie par :

$$\forall B \in \mathcal{F}, \mathbb{P}_X(F) = \mathbb{P}(X^{-1}(F)) =: \mathbb{P}(X \in F).$$

Exemple. Lorsque X prend un nombre dénombrable de valeurs $\{x_n, n \in \mathbb{N}\}$, on parle de variable *discrète*, et la donnée de la loi de X équivaut à la donnée de $(\mathbb{P}(X = x_n))_n$. On a en fait $\mathbb{P}_X = \sum_n \mathbb{P}(X = x_n)\delta_{x_n}$. Une loi est alors une famille sommable de réels positifs, de somme 1.

Exemple. Si E est un ensemble fini, la *loi uniforme* sur E est la loi $\sum_{x\in E} \frac{\delta_x}{|E|}$.

Exemple. La loi de Bernoulli est la loi $p\delta_1 + (1-p)\delta_0$. La loi de Poisson est définie par $\mathcal{P}(\lambda) = \sum_n \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!}\delta_n$. On fournit plein de lois usuelles discrètes en annexe.

Définition. Si X est à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on parle de variable aléatoire réelle.

Théorème. (de Radon-Nikodym)
Soit μ une mesure définie sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. On suppose que μ est *absolument continue* par rapport à la mesure de Lebesgue λ_d, *i.e.* $\lambda(B) = 0 \Rightarrow \mu(B) = 0$ pour tout borélien B. Alors il existe une unique fonction f mesurable telle que pour tout borélien A, on ait

$$\mu(A) = \int_A f(x)\,\mathrm{d}\lambda_d(x).$$

On dit alors que μ est à densité, de densité f.

Définition. Si $\mathbb{P}_X$ est à densité par rapport à la mesure de Lebesgue, on dira que X aussi et que f est la densité de X.

Exemple. La loi uniforme sur $[a, b]$ est la loi de densité $\mathbf{1}_{[}a, b]$ sur $\mathbb{R}$.

Exemple. La loi normale $\mathcal{N}(\mu, \sigma^2)$ est la loi de densité $x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

1.2 Indépendance

Remarque. L'indépendance est une notion qui repose sur les tribus au départ, et qui n'a rien à voir avec la loi. Pour savoir si deux variables sont indépendantes, regarder la loi de l'une ou de l'autre séparément n'apportera aucune information. Il faut regarder la *loi conjointe*, c'est-à-dire la loi du couple.

Proposition. Deux variables $X : (\Omega_1, \mathcal{F})_1, \mathbb{P}_1)$ et $Y : (\Omega, \mathcal{F}_2, \mathbb{P}_2)$ sont indépendantes si et seulement si la loi de (X, Y), définie sur l'espace produit $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$ est égale à la loi produit $\mathbb{P}_X \otimes \mathbb{P}_Y$.

Proposition. Si d variables aléatoires réelles indépendantes sont à densité, de densités respectives $f_1, \ldots, f_d$ sur $\mathbb{R}$, alors le d-uplet est à densité, de densité $f_1 \otimes \ldots \otimes f_d$.

Exemple. La loi d'un couple de deux variables aléatoires indépendantes de même loi $\mathcal{N}(0, \sigma^2)$ est à densité, de densité $(x, y) \mapsto \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2+y^2}{2}}$.

Proposition. Soient X et Y deux variables aléatoires réelles indépendantes.

— Si X et Y sont discrètes, alors $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$ (convolution de suites).

— Si X et Y sont à densité, alors $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$ (convolution dans $\mathrm{L}^1(\mathbb{R})$).

Remarque. Ce n'est pas pratique pour trouver la loi d'une somme en pratique, puisque une convolution n'est pas pratique à calculer. On verra qu'on peut se ramener à étudier des outils plus commodes, qui se comportent bien avec la convolution (séries entières, transformées de Fourier).

1.3 Espérance

Définition. Soit X une variable aléatoire réelle. Pour $h : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée, $h(X)$ est alors intégrable et on pose

$$\mathbb{E}(h(X)) := \int_\Omega h(X(\omega)) \,\mathrm{d}\,\mathbb{P}(\omega) = \int_\mathbb{R} h(x) \,\mathrm{d}\,\mathbb{P}_X(x).$$

Remarque. Deux fonctions ont même loi si et seulement si pour toute fonction continue bornée, $\mathbb{E}(h(X)) = \mathbb{E}(h(Y))$.

Proposition. Si deux variables X et Y sont indépendantes et intégrables, alors XY est intégrable avec $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.

Définition. Si X est intégrable, on dit que X est *d'espérance finie*, et on définit son espérance $\mathbb{E}(X) = \int_\Omega X$. De plus, si X est p-intégrable, $p \in \mathbb{N}^*$, on dit qu'elle admet un *moment d'ordre* p, que l'on définit par $\mathbb{E}(X^p)$.

Exemple. Pour une variable X discrète à valeurs dans $\{x_n\}$, on a

$$\mathbb{E}(h(X)) \text{ existe si et seulement si } \sum_n |h(x_n)|\,\mathbb{P}(X = x_n) < \infty,$$

et alors $\mathbb{E}(h(X)) = \sum_n h(x_n)\,\mathbb{P}(X = x_n)$.

Exemple. Pour une variable X à densité de densité f, on a

$$\mathbb{E}(h(X)) \text{ existe si et seulement si } \int_\mathbb{R} |h| f < \infty,$$

et alors $\mathbb{E}(h(X)) = \int_\mathbb{R} h(x) f(x) \,\mathrm{d}x$.

Exemple. L'espérance d'une loi de Poisson $\mathcal{P}(\lambda)$ est λ, celle d'une loi normale $\mathcal{N}(\mu, \sigma)$ est μ.

Définition. Soit $(X_n)_n$ une suite de variables aléatoires réelles. On dit qu'elle *converge en loi* vers une variable aléatoire X si pour toute fonction h continue bornée : $\mathbb{E}(h(X_n))$ converge dans $\mathbb{R}$ vers $\mathbb{E}(h(X))$.

Remarque. Cette notion porte uniquement sur les mesures, les variables (X_n) ne sont *a priori* même pas définies sur le même espace. Il n'y a aucune raison pour que ce mode de convergence entraîne un autre mode de convergence de variables aléatoires.

2 Caractérisation de la loi

2.1 Fonction de répartition

Définition. Soit X une variable aléatoire réelle. Sa fonction de répartition est définie par $F_X := x \mapsto \mathbb{P}(X \leq x)$.

Proposition. La fonction de répartition d'une variable X est à valeurs dans $[0, 1]$, croissante, a pour limites en $\pm\infty$ 0 et 1, et est continue à droite, avec des limites à gauche.

Proposition. La fonction de répartition caractérise la loi.

Proposition. Une suite (X_n) converge en loi vers X si et seulement si F_{X_n} converge simplement vers F en tout point de continuité de F.

Exemple. Un maximum de lois uniformes sur $[0,1]$ converge en loi vers 1.

2.2 Fonction caractéristique

Définition. Soit μ une mesure positive finie sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. On définit sa transformée de Fourier $\widehat{\mu} \in \mathcal{C}^0(\mathbb{R})$ par

$$\widehat{\mu}(t) = \int_R \mathrm{e}^{-\mathrm{i}\langle t,x\rangle}\,\mathrm{d}\mu(x).$$

Proposition. Si μ est à densité par rapport à la mesure de Lebesgue, *i.e.* $\mathrm{d}\mu(x) = f(x)\mathrm{d}x$, alors $\widehat{\mu} = \widehat{f}$.

Définition. Si μ et ν sont deux mesures positives finies sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, alors on définit leur *produit de convolution* par

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \mu * \nu(A) = \int_{(R^d)^2} \mathbf{1}_A(x+y)\,\mathrm{d}\mu(x)\mathrm{d}\nu(y).$$

C'est la mesure image de $(x,y) \mapsto x+y$ par $(\mu \otimes \nu)$.

Proposition. Avec les notations précédentes, on a la relation $\widehat{\mu * \nu} = \widehat{\mu}\widehat{\nu}$.

Proposition. (Inversion de Fourier pour les mesures positives)
Pour tous les réels $a < b$, et toute mesure μ positive finie, on a :

$$\frac{1}{2}\mu(\{a,b\}) + \mu(]a,b[) = \lim_{T\to+\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{\mathrm{e}^{-ita} - \mathrm{e}^{-itb}}{\mathrm{i}t}\widehat{\mu}(-t)\,\mathrm{d}t.$$

Définition. Soit X une variable aléatoire réelle, on définit sa *fonction caractéristique* par $\varphi_X : t \mapsto \mathbb{P}_X(-t) = \mathbb{E}(\mathrm{e}^{\mathrm{i}tX})$.

Corollaire. La fonction caractéristique caractérise la loi de X.

Proposition. Si X et Y sont deux variables aléatoires indépendantes, alors $\varphi_{X+Y} = \varphi_X\varphi_Y$.

Exemple. Si X suit une loi normale $\mathcal{N}(m,\sigma^2)$, alors $\varphi_X(t) = \mathrm{e}^{\mathrm{i}mt - t^2\sigma^2/2}$ pour tout $t \in \mathbb{R}$.

Théorème. (Cramér)
Si X et Y sont deux variables aléatoires indépendantes à densité par rapport à la mesure de Lebesgue dont la somme est une loi normale, alors les deux sont de loi normale.

Théorème. (de Paul Lévy) Soit (X_n) une suite de variables aléatoires réelles. Alors la suite converge en loi vers X si et seulement si on a convergence simple des fonctions caractéristiques.

Théorème. (Théorème limite central)
Soit (X_n) une suite de variables L^2 i.i.d d'espérance commune μ, et de variance commune σ^2.

Alors $Z_n := \sqrt{n}\frac{\overline{X_n}-\mu}{\sigma}$ converge en loi vers une loi $\mathcal{N}(0,1)$.

De plus, si les X_n admettent pour densité commune f, que leur fonction caractéristique est dans un L^p, $p \in [1,\infty[$, et que leur moment d'ordre trois est nul, on a l'asymptotique suivante, pour $x \in \mathbb{R}$ fixé :

$$\|f_n(x) - \nu(x)\| = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de Z_n et ν celle de la loi normale centrée réduite.

Application. Sous les mêmes hypothèses, si a_n est une suite de réels strictement positifs qui vérifie $a_n \to 0$ et $a_n + \frac{1}{\sqrt{n}} \sim a_n$, alors pour $\alpha \in \mathbb{R}$:

$$\mathbb{P}\left(\alpha < Z_n < \alpha + a_n\right) \sim a_n \nu(\alpha).$$

2.3 Série génératrice

Définition. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On définit sa *série génératrice* $G_X : t \mapsto \mathbb{E}(t^X)$. Elle définie une série entière de rayon de convergence au moins égal à 1.

Proposition. La série génératrice caractérise la loi.

Exemple. On considère un processus de branchement, *i.e.* une suite (Z_n) de variables aléatoires définies par récurrence par :

$$Z_{n+1} = \sum_{k=0}^{Z_n} X_{n,k},$$

où les $X_{i,j}$ sont i.i.d de loi μ. Alors pour tout n, $G_{Z_n} = G_\mu^{\circ n}$.

Proposition. Si X et Y sont deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$, alors la série génératrice de la somme est le produit des séries génératrices.

Exemple. Si X est une loi de Poisson de paramètre λ, sa série génératrice est définie pour $t \in \mathbb{R}$ par $G_X(t) = \mathrm{e}^{\lambda(t-1)}$. Ainsi, la somme de deux lois de Poisson indépendantes est une loi de Poisson, dont le paramètre est la somme des deux.

Théorème. (de Raïkov)
Si $\mathbb{P}_1$ et $\mathbb{P}_2$ sont deux mesures de probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ telles que $\mathbb{P}_1 * \mathbb{P}_2 = \mathcal{P}(\lambda)$ pour un $\lambda > 0$, alors $\mathbb{P}_1$ et $\mathbb{P}_2$ sont des lois de Poisson. Plus simplement, si X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes dont la somme suit une loi de Poisson, alors X et Y suivent une loi de Poisson.

2.4 Problème des moments

Proposition. Soit X une variable aléatoire réelle presque sûrement bornée. Alors X est caractérisée par ses moments.

Proposition. Soit X une variable aléatoire réelle et φ_X sa fonction caractéristique.

— Si X admet un moment d'ordre $n \in \mathbb{N}^*$, alors φ_X est $\mathcal{C}^n$, et pour $k \leq n$:

$$\varphi_X^{(k)}(t) = \mathrm{i}^k \int_{\mathbb{R}} x^k \mathrm{e}^{\mathrm{i}tx} \,\mathrm{d}\,\mathbb{P}_X(x).$$

En particulier, $\varphi_X^{(k)}(0) = \mathrm{i}^k\,\mathbb{E}(X)^k$.

— Si est k fois dérivable en 0, $k \geq 2$, X admet des moments jusqu'à l'ordre $2\lfloor \frac{K}{2} \rfloor$, donnés par la formule précédente.

Théorème. Soit X une variable aléatoire réelle et φ_X sa fonction caractéristique. On suppose que X admet des moments à tout ordre, notés $(m_n)_n$. Si de plus, $\limsup_n \frac{m_n^{1/n}}{n} < +\infty$, alors X est caractérisée par ses moments.

3 Loi uniforme sur les permutations

3.1 Nombre d'inversions

Définition. Comme vu précédemment, la loi uniforme sur $\mathfrak{S}_n$ est définie par $\mathbb{P}(\{\sigma\}) : \frac{1}{n!}$.

Définition. Le *nombre d'inversions* d'une permutation σ est le nombre de couples (i, j) tels que $i < j$ et $\sigma(j) > \sigma(i)$. On note I_n qui à σ associe le nombre d'inversions de $\sigma \in \mathfrak{S}_n$.

Proposition. Si σ suit une loi uniforme sur $\mathfrak{S}_{n-1}$, et k suit une loi uniforme sur $\{1, \ldots, n\}$, alors la permutation obtenue en insérant n à la k-ième place dans σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. La loi de I_n est celle d'une somme de n variables indépendantes de loi $\mathcal{U}(\{0, k-1\})$ où $k \in \{1, \ldots, n\}$. En particulier, $\frac{6}{n^{3/2}}(I_n - \mathbb{E}(I_n))$ converge en loi vers une loi normale $\mathcal{N}(0, 1)$.

3.2 Nombre de cycles

Proposition. On va construire une permutation aléatoire selon ses cycles. On met d'abord un 1. Puis si on a i nombres utilisés, on tire une $\mathcal{B}(\frac{1}{n-i+1})$. Si elle vaut 1, on clôt le cycle. Sinon on ne touche à rien. On choisit un nombre au hasard que l'on rajoute à la suite. On obtient alors une loi uniforme sur $\mathfrak{S}_n$.

Application. On note $K_n(\sigma)$ le nombre de cycles d'une permutation $\sigma \in \mathfrak{S}_n$. Sous la loi uniforme, K_n est une somme de $\mathcal{B}(\frac{1}{k})$ indépendantes, pour $k \in \{1, \ldots, n\}$. En particulier, $\frac{K_n - \ln(n)}{\sqrt{\ln(n)}}$ converge en loi vers une loi normale $\mathcal{N}(0, 1)$.

3.3 Nombre de descentes

Définition. Soit $\sigma \in \mathfrak{S}_n$ on dit que σ présente une *descente au rang* i si $\sigma(i+1) < \sigma(i)$. On note $A_{n,k}$ le nombre de permutations de $\mathfrak{S}_n$ qui possèdent exactement k descentes.

Proposition. Soit $(U_1, \ldots, U_n) \sim \mathcal{U}(0, 1)^{\otimes n}$. Pour presque tout $\omega \in \Omega$, il existe $\sigma(\omega)$ tel que $U_{\sigma(1)} < \ldots < U_{\sigma(n)}$. Alors σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. Avec les mêmes notations, la loi du nombre de descentes est la loi $\lfloor U_1 + \ldots + U_n \rfloor$. De plus, pour $\alpha \in [0,1]$, alors, lorsque n tend vers $+\infty$:

$$A(n, \alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}} \, \mathrm{e}^{-6(\alpha - 1/2)^2}.$$

Chapitre 66

262. Convergences d'une suite de variables aléatoires. Théorèmes limite. Exemples et applications.

Sommaire

1 Convergence en probabilité, convergence presque sûre 466
1.1 Convergence en probabilité 466
1.2 Convergence presque sûre, utilisation lemmes de Borel-Cantelli 466
2 Convergence en norme 467
2.1 Convergence pour les différentes normes 467
2.2 Intégrabilité uniforme 467
3 Séries de variables indépendantes 468
3.1 Convergence générale 468
3.2 Lois des grands nombres 468
4 Convergence en loi 468
4.1 Définition, caractérisations 468
4.2 Théorème limite central 469
5 Statistiques 469
5.1 Estimation paramétrique 469
5.2 Tests non paramétriques d'adéquation à une loi 470

Présentation du plan

Les variables seront définies par défaut dans un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.

1 Convergence en probabilité, convergence presque sûre

1.1 Convergence en probabilité

Définition. On dit que la suite de variables aléatoires réelles $(X_n)_n$ *en probabilité* vers une variable aléatoire X si pour tout $\varepsilon > 0$, la suite $(\mathbb{P}(|X_n - X| > \varepsilon))_n$ tend vers 0.

Remarque. On peut faire de même dans $\mathbb{R}^d$ avec une norme quelconque.

Proposition. La limite en probabilité est unique.

Exemple. Soit (X_n) une suite de variables indépendantes de lois respectives $\mathcal{B}(1-\frac{1}{n})$. Alors (X_n) converge en probabilité vers 0.

Exemple. Le maximum de variables uniformes indépendantes sur $[0,1]$ converge en probabilité vers 0.

Proposition. La convergence en probabilité est préservée par application continue.

1.2 Convergence presque sûre, utilisation lemmes de Borel-Cantelli

Définition. Une suite $(X_n)_n$ de variables aléatoires réelles *converge presque sûrement* (pour la mesure $\mathbb{P}$) si :

$$\mathbb{P}(\{\omega \in \Omega, X_n(\omega) \to X(\omega)\} = 1).$$

Remarque. Il s'agit d'une convergence simple, on a donc unicité de la limite simple (à égalité presque sûre près). De plus, la convergence est préservée par image continue.

Théorème. Soit $(X_n)_n$ une suite de variables aléatoires.

- S'il existe une suite positive (ε) dont la série converge, avec $\sum_n \mathbb{P}(|X_{n+1}-X_n| > \varepsilon_n) < +\infty$, alors (X_n) converge presque sûrement.
- Si X est telle que pour tout $\varepsilon > 0$, la série $\sum_n \mathbb{P}(|X_n - X| > \varepsilon)$ converge, alors $(X_n)_n$ converge presque sûrement vers X.

Exemple. Dans les exemples précédents, on n'a pas convergence ps dans le premier mais dans le second si.

Théorème. La convergence presque sûre entraîne celle en probabilité avec égalité des limites. Réciproquement, une suite qui converge en probabilité converge presque sûrement à extraction près.

2 Convergence en norme

2.1 Convergence pour les différentes normes

Soit ici $p \in [1, \infty[$.

Définition. On dit que $(X_n) \in \mathrm{L}^p$ converge dans L^p si $\mathbb{E}(|X_n - X|^p)$ tend vers 0 avec n (rien de surprenant c'est la convergence associée à la norme L^p).

Remarque. Puisque l'on a des inclusions entre ces espaces, on a des implications entre ces convergences. Pour obtenir une convergence en norme à partir d'une convergence presque sûre, il faut utiliser le théorème de convergence dominée.

Proposition. (inégalité de Markov)
Soit X une variables aléatoire positive, alors pour $\lambda > 0$,

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}.$$

Proposition. La convergence L^p entraîne la convergence en probabilité. Plus généralement, des inégalités de concentration fines peuvent permettre de montrer des convergences presques sûres.

Remarque. La réciproque est fausse, pour entraîner de la convergence en norme, on va voir qu'il nous faut un contrôle uniforme sur les queues de distribution.

2.2 Intégrabilité uniforme

Définition. Une famille de variables aléatoires $(X_i)_{i\in I}$ est *uniformément intégrable* lorsque

$$\lim_{a\to\infty} \sup_{i\in I} \mathbb{E}(|X_i|\mathbf{1}_{|X_i|>a}) = 0.$$

Proposition. Si une famille est uniformément bornée par une variable intégrable positive, alors elle est uniformément intégrable.

Proposition. Si $(X_i)_i$ est bornée dans L^p pour $p > 1$ alors elle est uniformément intégrable.

Définition. Une famille de variables aléatoires (X_i) est dite *équicontinue* si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $\mathbb{P}(A) \leq \delta$ alors $\sup_i \mathbb{E}(|X_i|\mathbf{1}_A) \leq \varepsilon$.

Proposition. Une famille de variables aléatoires est uniformément intégrable si et seulement si elle est équicontinue et bornée dans L^1.

Théorème. Une suite de variables aléatoires converge dans L^1 si et seulement si elle converge en probabilité et qu'elle est uniformément intégrable.

Exemple. Pour $(X_n := \frac{1}{n^2}\delta_{n^2} + (1-\frac{1}{n})\delta_0)$, on a convergence en probabilité mais pas en norme 1.

3 Séries de variables indépendantes

3.1 Convergence générale

Théorème. (inégalité de Kolmogorov)
Soient $(X_1, \ldots, X_n)$ des variables indépendantes, on note $\sigma_i^2 := \mathrm{Var}(X_i)$. Pour $\varepsilon > 0$:

$$\mathbb{P}\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^{k} X_i\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\sum_i = 1^n\sigma_i^2.$$

Proposition. Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes centrées et admettant une variance. Si la série des variances converge, alors la série des variables aléatoires converge presque sûrement, et dans L^2.

3.2 Lois des grands nombres

Proposition. (loi faible des grands nombres)
Soit (X_i) une suite i.i.d de variables aléatoires L^2. Alors leur moyenne empirique converge dans L^2 et donc en probabilité vers leur espérance commune.

Théorème. (loi forte des grands nombres)
Soit $(X_i)_i$ une suite i.i.d de variables aléatoires admettant une espérance. Alors leur moyenne empirique converge presque sûrement vers leur espérance commune.

4 Convergence en loi

4.1 Définition, caractérisations

Définition. Soit $(X_n)_n$ une suite de variables aléatoires réelles. On dit qu'elle *converge en loi* vers une variable aléatoire X si pour toute fonction h continue bornée : $\mathbb{E}(h(X_n))$ converge dans $\mathbb{R}$ vers $\mathbb{E}(h(X))$.

Remarque. Cette notion porte uniquement sur les mesures, les variables (X_n) ne sont *a priori* même pas définies sur le même espace. Il n'y a aucune raison pour que ce mode de convergence entraîne un autre mode de convergence de variables aléatoires. Cependant, toutes les autres formes de convergence l'entraînent. Par ailleurs, si des variables définies sur un même espace de probabilité convergent en loi vers une constante, alors elles convergent en probabilité vers cette constante.

Proposition. Dans le cas discret, la convergence en loi équivaut à la convergence des probabilités atomiques.

Exemple. Une suite $\mathcal{B}(n, \lambda/n)$ converge en loi vers une loi de poisson $\mathcal{P}(\lambda)$.

Proposition. (de l'application continue)
La convergence en loi est stable par application d'une fonction continue.

Proposition. (lemme de Slutsky)
Soient (X_n), (Y_n) deux suites de variables aléatoires. On suppose que les deux convergent en loi, dont l'une vers une constante. Alors le couple converge en loi.

Proposition. Une suite (X_n) de variables réelles converge en loi vers X si et seulement si F_{X_n} converge simplement vers F en tout point de continuité de F.

Exemple. On retrouve notre résultat sur le maximum de variables uniformes sur $[0, 1]$.

Théorème. (de Paul Lévy) Soit (X_n) une suite de variables aléatoires réelles. Alors la suite converge en loi vers X si et seulement si on a convergence simple des fonctions caractéristiques.

4.2 Théorème limite central

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1+\ldots+X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3 - 3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

5 Statistiques

5.1 Estimation paramétrique

Théorème. (estimation et intervalle de confiance pour une moyenne inconnue)
Soit $(X_n)_n$ une suite de variables i.i.d de même espérance μ inconnue et de même variance σ^2 connue. Alors, si q_β est le quantile d'ordre β d'une loi normale centrée réduite :

$$\mathbb{P}\left(\mu \in \left[\overline{X_n} \pm \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right]\right) \to \alpha,$$

ce que l'on appelle en statistiques un *intervalle de confiance* asymptotique de niveau de confiance $1 - \alpha$.

Remarque. Quand on ne connaît pas la variance, on peut la remplacer par un estimateur et utiliser le lemme de Slutsky. Mais il peut subsister un problème ; la variance peut dépendre de la moyenne. Dans ce cas, on peut quelquefois s'en sortir avec le résultat suivant.

Proposition. (méthode delta)
Avec les mêmes hypothèses, si g est dérivable en μ de dérivée non nulle, alors $\sqrt{n}(g(\overline{X_n}) - g(\mu))$ converge en loi vers $\mathcal{N}(0, \sigma^2)$.

Exemple. Un exemple d'intervalle de confiance pour une loi $\mathcal{P}(\lambda)$ est

$$\sqrt{\lambda} \in \left[\sqrt{\overline{X_n}} \pm \frac{1}{2\sqrt{n}} q_{1-\frac{\alpha}{2}}\right].$$

5.2 Tests non paramétriques d'adéquation à une loi

Théorème. Soient $(X_1, \ldots, X_n)$ des variables aléatoires i.i.d à valeurs dans $\{a_1, \ldots, a_m\}$, soit $p_i = \mathbb{P}(X_1 = a_i)$ supposé strictement positif. On note $N_k(n) := \{i, X_i = a_k\}$. Lorsque n tend vers l'infini,

$$D_n := \sum_{k=1}^{m} \frac{(N_k(m) - np_k)^2}{np_k} \xrightarrow{\text{loi}} \chi^2(m-1).$$

Réciproquement, si les (X_i) suivent une loi différente sur ce même ensemble, alors D_n converge presque sûrement vers $+\infty$.

Remarque. Cela permet de bâtir le test dit du khi-deux d'adéquation à une loi.

Théorème. (Glivenko-Cantelli)
La fonction de répartition empirique associée à un n-échantillon $F_n := \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{[X_k,\infty[}$ converge presque sûrement uniformément vers la fonction de répartition F commune aux X_i.

Remarque. En particulier, on montre dans ce théorème que la statistique $D_n := \sqrt{n}\|F_n - F\|$ est libre, et donc on peut construire un test d'adéquation non asymptotique à partir de celle-ci en estimant les quantiles par Monte-Carlo.

Théorème. (Kolmogorov, ADMIS)
La statistique D_n converge en loi vers une loi tabulée dont on connaît la fonction de répartition.

Chapitre 67

264. Variables aléatoires discrètes. Exemples et applications.

Sommaire

1 Variables aléatoires discrètes 474
1.1 Définition, lois usuelles 474
1.2 Indépendance de variables aléatoires discrètes 474
2 Moments et fonction génératrice 475
2.1 Espérance et moments 475
2.2 Fonction génératrice 475
3 Des exemples des processus aléatoires discrets 476
3.1 Processus de Galton-Watson 476
3.2 Marche aléatoire sur le cercle 477
3.3 Marche aléatoire sur un réseau 477
4 Loi uniforme sur les permutations 477
4.1 Nombre d'inversions 477
4.2 Nombre de cycles 478
4.3 Nombre de descentes 478

Présentation du plan

Les variables seront définies par défaut dans un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.

1 Variables aléatoires discrètes

1.1 Définition, lois usuelles

Définition. Une variable aléatoire est *discrète* si $X(\Omega)$ est dénombrable, ou plus généralement s'il existe E dénombrable tel que $\mathbb{P}(X \in E) = 1$. La loi de X est alors

$$\mathcal{L}(X) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x)\delta_x.$$

Réciproquement, une *loi discrète* est une loi de la dorme $\sum_{x\in E} p_x\delta_x$ avec E dénombrable.

Remarque. Un couple de variables discrètes est discret, et donc toute fonction de plusieurs variables discrètes est encore une variable discrète.

Définition. (quelques lois usuelles)

- loi uniforme sur E fini $\mathcal{U}(E) = \frac{1}{|E|}\sum_{x\in E}\delta_x$.
- loi de Bernoulli $\mathbb{P}(X=0) = \mathbb{P}(X=1) = \frac{1}{2}$ (pile ou face)
- loi binomiale $\mathcal{B}(n,p) = \sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}$ (plusieurs pile ou face indépendants)
- loi de Poisson $\mathcal{P}(\lambda) = \sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}$ (événements rares)
- loi géométrique $\mathcal{G}(p) = \sum_{n=1}^{+\infty}(1-p)^{n-1}p$ (loi d'un premier succès de pile ou face)
- la loi $\zeta(s) = \frac{1}{\zeta(s)}\sum_{n=1}^{\infty} n^{-s}\delta_n$.

Proposition. La loi géométrique est l'unique loi sur $\mathbb{N}^*$ sans mémoire, c'est-à-dire que pour tout k, l, $\mathbb{P}(X > k+l | X > l) = \mathbb{P}(X > k)$.

1.2 Indépendance de variables aléatoires discrètes

Proposition. Des variables $(X_1, \ldots, X_n)$ à valeurs dans E discrètes sont indépendantes si et seulement si pour tout $(x_1, \ldots, x_n) \in E^n$, on a

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1)\cdots\mathbb{P}(X_n = x_n).$$

Proposition. Si E possède une structure de groupe commutatif, alors si X et Y à valeurs dans E sont indépendantes :

$$\mathbb{P}(X+Y=z) = \sum_{x\in E}\mathbb{P}(X=x)\,\mathbb{P}(Y=z-x).$$

Remarque. Pour calculer la loi d'une somme de variables indépendantes, on est amenés à calculer des convolutions de suites numériques.

Exemple. Si X et Y sont indépendantes et suivent des lois de Poisson de paramètres respectifs λ et μ, alors $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

Exemple. Si on met sur $\{1, \ldots, n\}$ la loi uniforme, on peut montrer que le fait d'être divisible par des nombres premiers distincts forme des événements indépendants, de quoi on déduit la formule $\varphi(n) = \frac{1}{n}\prod_{p|n}(1-1/p)$.

Exemple. De même sur $\mathbb{N}$ avec la loi $\zeta(s)$, on montre que

$$\zeta(s) = \prod_p (1-1/p)^{-s}.$$

2 Moments et fonction génératrice

2.1 Espérance et moments

Proposition. Soit X une variable discrète à valeurs dans $E \subset \mathbb{R}$ (dénombrable) et h une fonction continue bornée sur $\mathbb{R}$. Alors

$$\mathbb{E}(h(X)) = \sum_{x \in E} h(x)\, \mathbb{P}(X = x)$$

et pour h seulement continue,

$$\mathbb{E}(h(X)) \text{ existe si et seulement si } \sum_{x \in E} |h(x)|\, \mathbb{P}(X = x) < \infty \text{ et vaut alors } \sum_{x \in E} h(x)\, \mathbb{P}(X = x).$$

Exemple. La loi de Poisson à des moments à tout ordre, et son espérance et sa variance correspondent au paramètre. On met aussi en annexe des moments des lois usuelles.

Proposition. Si X à valeurs dans $\mathbb{N}$ possède un moment d'ordre p, alors

$$\mathbb{E}(X^p) = \sum_{n=0}^{\infty} ((n+1)^p - n^p)\, \mathbb{P}(X > n)$$

2.2 Fonction génératrice

On considère ici des variables à valeurs dans $\mathbb{N}$.

Définition. La *série génératrice* de X est la série entière $G_X = \sum_n \mathbb{P}(X = n) x^n$. Son rayon de convergence est au moins 1 et quand c'est défini on a $G_X(t) = \mathbb{E}(t^X)$.

Proposition. La série génératrice caractérise la loi.

Exemple. La série génératrice de la loi $\mathcal{P}(\lambda)$ est définie sur $\mathbb{C}$ par $z \mapsto \mathrm{e}^{\lambda(z-1)}$.

Proposition. Si deux variables sont indépendantes, alors le produit des fonctions génératrices correspond à la fonction génératrice de la somme.

Remarque. On retrouve le fait que si deux variables indépendantes suivent une loi de Poisson alors la somme aussi.

Théorème. (indécomposabilité de la loi de Poisson)
Réciproquement, si X et Y sont deux variables indépendantes à valeurs dans $\mathbb{N}$ telles que la somme suit une loi de Poisson, alors chacune suit une loi de Poisson également.

3 Des exemples des processus aléatoires discrets

3.1 Processus de Galton-Watson

Définition. Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. Un tel processus est appelé *processus de Galton-Watson.*

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

3.2 Marche aléatoire sur le cercle

Théorème. On considère une particule sur $\mathbb{Z}/N\mathbb{Z} = (\overline{0}, \ldots, \overline{N-1})$, initialement en $\overline{0}$. À chaque étape, la particule se déplace d'un cran vers la droite ou vers la gauche ou stationne de manière équiprobable. On peut modéliser cela par une chaîne de Markov dont la loi initiale est $\delta_{\overline{0}}$ et dont la loi de transition est $\mu = \frac{1}{3}(\delta_{-\overline{1}} + \delta_{\overline{0}} + \delta_{\overline{1}})$. La loi de X_n la position de la particule au rang n a pour loi

$$\mu_n = \mu_0 * \mu^{*n} = \mu^{*n}.$$

La chaîne de Markov en question converge vers la loi uniforme notée μ_∞ et on a une approximation de la vitesse de convergence en variation totale :

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N}\mathrm{e}^{-\frac{\pi^2 n}{N^2}}.$$

3.3 Marche aléatoire sur un réseau

Définition. On définit la *marche aléatoire symétrique* sur $\mathbb{Z}^d$ comme la chaîne de Markov de loi initiale δ_0 et de transition

$$\mathbb{P}(X_{n+1} = X_n \pm e_i) = \frac{1}{2^d}$$

pour tout i, où les (e_i) sont les vecteurs de la base canonique de $\mathbb{Z}^d$.

Théorème. La marche aléatoire revient presque sûrement à l'origine une infinité de fois si et seulement si $n \geq 2$, et dans le cas contraire elle visite l'origine un nombre fini de fois presque sûrement.

4 Loi uniforme sur les permutations

4.1 Nombre d'inversions

Définition. Comme vu précédemment, la loi uniforme sur $\mathfrak{S}_n$ est définie par $\mathbb{P}(\{\sigma\}) : \frac{1}{n!}$.

Définition. Le *nombre d'inversions* d'une permutation σ est le nombre de couples (i, j) tels que $i < j$ et $\sigma(j) > \sigma(i)$. On note I_n qui à σ associe le nombre d'inversions de $\sigma \in \mathfrak{S}_n$.

Proposition. Si σ suit une loi uniforme sur $\mathfrak{S}_{n-1}$, et k suit une loi uniforme sur $\{1, \dots, n\}$, alors la permutation obtenue en insérant n à la k-ième place dans σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. La loi de I_n est celle d'une somme de n variables indépendantes de loi $\mathcal{U}(\{0, k-1\})$ où $k \in \{1, \dots, n\}$. En particulier, $\frac{6}{n^{3/2}}(I_n - \mathbb{E}(I_n))$ converge en loi vers une loi normale $\mathcal{N}(0, 1)$.

4.2 Nombre de cycles

Proposition. On va construire une permutation aléatoire selon ses cycles. On met d'abord un 1. Puis si on a i nombres utilisés, on tire une $\mathcal{B}(\frac{1}{n-i+1})$. Si elle vaut 1, on clôt le cycle. Sinon on ne touche à rien. On choisit un nombre au hasard que l'on rajoute à la suite. On obtient alors une loi uniforme sur $\mathfrak{S}_n$.

Application. On note $K_n(\sigma)$ le nombre de cycles d'une permutation $\sigma \in \mathfrak{S}_n$. Sous la loi uniforme, K_n est une somme de $\mathcal{B}(\frac{1}{k})$ indépendantes, pour $k \in \{1, \dots, n\}$. En particulier, $\frac{K_n - \ln(n)}{\sqrt{\ln(n)}}$ converge en loi vers une loi normale $\mathcal{N}(0, 1)$.

4.3 Nombre de descentes

Définition. Soit $\sigma \in \mathfrak{S}_n$ on dit que σ présente une *descente au rang* i si $\sigma(i+1) < \sigma(i)$. On note $A_{n,k}$ le nombre de permutations de $\mathfrak{S}_n$ qui possèdent exactement k descentes.

Proposition. Soit $(U_1, \dots, U_n) \sim \mathcal{U}(0, 1)^{\otimes n}$. Pour presque tout $\omega \in \Omega$, il existe $\sigma(\omega)$ tel que $U_{\sigma(1)} < \dots < U_{\sigma(n)}$. Alors σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. Avec les mêmes notations, la loi du nombre de descentes est la loi $\lfloor U_1 + \dots + U_n \rfloor$. De plus, pour $\alpha \in [0, 1]$, alors, lorsque n tend vers $+\infty$:

$$A(n, \alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}} \, \mathrm{e}^{-6(\alpha - 1/2)^2}.$$

Chapitre 68

266. Utilisation de la notion d'indépendance en probabilités.

Sommaire

1 Notion d'indépendance 480
1.1 Événements et tribus indépendantes 480
1.2 Variables aléatoires indépendantes 480
2 Loi de variables indépendantes 480
2.1 Espérance et variables indépendantes 480
2.2 Somme de variables indépendantes 481
2.3 Théorèmes limite 481
3 Vecteurs gaussiens 482
3.1 Définition, théorème de Cochran 482
3.2 Modèle linéaire gaussien, moindres carrés 483
4 Loi uniforme sur les permutations 484
4.1 Nombre d'inversions 484
4.2 Nombre de cycles 484
4.3 Nombre de descentes 485

Présentation du plan

Les variables seront définies par défaut dans un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.

1 Notion d'indépendance

1.1 Événements et tribus indépendantes

Définition. On dit que deux événements A et B sont *indépendants* lorsque $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$. On généralise cela à un nombre fini d'événements.

Remarque. Si A et B sont indépendants, alors A et $\overline{B}$ sont indépendants.

Définition. Des événements $(A_i)_i$ sont indépendants si toute sous-famille finie est formée d'événements indépendants. En particulier, on peut définir des tribus indépendantes.

Exemple. Si on met sur $\{1,\dots,n\}$ la loi uniforme, on peut montrer que le fait d'être divisible par des nombres premiers distincts forme des événements indépendants, de quoi on déduit la formule $\varphi(n) = \frac{1}{n}\prod_{p|n}(1-1/p)$.

Exemple. De même sur $\mathbb{N}$ avec la loi $\zeta(s)$, on montre que

$$\zeta(s) = \prod_p (1-1/p)^{-s}.$$

1.2 Variables aléatoires indépendantes

Définition. On dit que deux variables aléatoires sont *indépendantes* lorsque les tribus qu'elles engendrent sont indépendantes.

Proposition. Si des variables (X_i) sont indépendantes et si les (f_i) sont des fonctions mesurables alors les $(f_i(X_i))$ sont indépendantes.

2 Loi de variables indépendantes

2.1 Espérance et variables indépendantes

Proposition. Deux variables sont indépendantes si et seulement si la loi du couple est égale à la loi produit.

Exemple. Dans le cas de variables à densité, cela équivaut à $f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$. Dans le cas discret, cela équivaut à avoir $\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y)$ pour tous x, y.

Proposition. Des variables $(X_1,\dots,X_n)$ sont indépendantes si et seulement si pour toutes fonctions continues bornées $(h_1,\dots,h_n)$:

$$\mathbb{E}(h_1(X_1)\cdots h_n(X_n)) = \mathbb{E}(h_1(X_1))\cdots\mathbb{E}(h_n(X_n)).$$

En particulier, elles sont indépendantes si et seulement si

$$\forall t_1,\dots,t_n \varphi_{(X_1,\dots,X_n)}(t_1,\dots,t_n) = \prod_i \varphi_{X_i}(t_i).$$

Proposition. Si X et Y admettent une espérance et sont indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.

Exemple. La réciproque est fausse.

2.2 Somme de variables indépendantes

Proposition. Si X et Y sont indépendantes, alors $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$.

Exemple. Si X et Y sont à densité sur $\mathbb{R}^d$, alors la densité de la somme correspond au produit de convolution des deux densités. Si X et Y son à valeurs dans $\mathbb{Z}$, alors la fonction de masse de la somme est la convolution discrète de la fonction de masse des deux.

Proposition. Soient X et Y deux variables indépendantes :

— La fonction caractéristique de leur somme est égale au produit de leurs fonctions caractéristiques

— Si elles sont à valeurs entières, on a la même chose avec les fonctions génératrices.

Exemple. Une somme de lois normales est une loi normale, idem pour la loi de Poisson.

Théorème. On a des réciproques aux résultats précédents :

— Si X et Y sont indépendantes et que leur somme suit une loi normale, alors X et Y suivent une loi normale.

— Si X et Y sont indépendantes et que leur somme suit une loi de Poisson, alors X et Y suivent une loi de Poisson.

Définition. On définit la loi du *khi-deux* à n degrés de libertés comme la loi de la somme de n carrés de variables gaussiennes centrées et réduites. On la note $\chi^2(n)$.

Proposition. La densité de $\chi^2(n)$ sur $\mathbb{R}^+$ est $x \mapsto \frac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} \mathrm{e}^{-x/2}$.

2.3 Théorèmes limite

Théorème. (inégalité maximale de Kolmogorov) : Si les $(X_i)_i$ sont des variables indépendantes avec un moment d'ordre deux. On suppose les (X_i) centrées. Alors pour $\varepsilon > 0$ et pour tout n :

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \sigma_i^2.$$

Proposition. Si les (X_i) sont centrées, et si la série des variances converge, alors la série $\sum_n X_n$ converge presque sûrement (et dans L^2).

Théorème. (loi des grands nombres avec un moment d'ordre deux)
Si $\mathbb{E}(X_n)$ tend vers m et si $\sum_{j=1}^{+\infty} \frac{\sigma_j^2}{j^2} < \infty$ alors $\overline{X_n}$ converge presque sûrement et dans L^2 vers m.

Théorème. (loi forte des grands nombres)
Soient $(X_n)_n$ des variables i.i.d d'espérance commune $m \in \mathbb{R}$. Alors $\overline{X_n}$ converge presque sûrement vers m.

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1+\ldots+X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Application. (estimation et intervalle de confiance pour une moyenne inconnue)
Soit $(X_n)_n$ une suite de variables i.i.d de même espérance μ inconnue et de même variance σ^2 connue. Alors, si q_β est le quantile d'ordre β d'une loi normale centrée réduite :

$$\mathbb{P}\left(\mu \in \left[\overline{X_n} \pm \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right]\right) \to \alpha,$$

ce que l'on appelle en statistiques un *intervalle de confiance* asymptotique de niveau de confiance $1-\alpha$.

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3 - 3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

3 Vecteurs gaussiens

3.1 Définition, théorème de Cochran

Définition. Un vecteur aléatoire est dit *gaussien* si toute combinaison linéaire de ses composantes suit une loi normale.

Définition. Si X est un vecteur aléatoire, on note $\mathrm{Cov}(X)$ la matrice de ses covariances. C'est une matrice symétrique positive (car $\langle \mathrm{Cov}(X)v, v\rangle = \mathrm{Cov}(\langle v, X\rangle) \geq 0$). On a :

$$\mathrm{Cov}(X) = \mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^\top).$$

Remarque. Si X est un vecteur gaussien, ses composantes sont des variables gaussiennes. Si X est un vecteur dont les composantes sont des lois normales indépendantes, alors X est un vecteur gaussien de matrice de variance diagonale.

Proposition. Si X est un vecteur aléatoire de $\mathbb{R}^n$ et si $A \in \mathcal{M}_n(\mathbb{R})$, alors

$$\mathbb{E}(AX) = A\,\mathbb{E}(X) \text{ et } \mathrm{Cov}(AX) = A\mathrm{Cov}(X)A^\top.$$

De même, si $v \in \mathbb{R}^n$, alors

$$\mathbb{E}(\langle v, X\rangle) = \langle v, \mathbb{E}(X)\rangle \text{ et } \mathrm{Cov}(\langle v, X\rangle) = \langle \mathrm{Cov}(X)v, v\rangle.$$

Corollaire. Soit X un vecteur gaussien. X est un vecteur indépendant si et seulement si $\mathrm{Cov}(X)$ est diagonale.

Proposition. La loi $\mathcal{N}(\mu, V)$ admet une densité sur $\mathbb{R}^n$ pour la mesure de Lebesgue si et seulement si V est définie positive, dans ce cas, on a :

$$\forall x \in \mathbb{R}^n,\ f_{(\mu,V)}(x) = \frac{1}{\sqrt{(2\pi)^d \det(V)}} \exp\left(-\frac{1}{2}\langle V^{-1}(x-\mu), (x-\mu)\rangle\right).$$

Théorème. (de Cochran)
Soit X un vecteur gaussien centré réduit sur $\mathbb{R}^d$, et $\mathbb{R}^d = \bigoplus F_i$ une somme directe orthogonale. Alors les projections orthogonales de X sur ces sous-espaces sont indépendantes, et la norme de la projection sur F_i suit une loi du khi-deux à $\dim(F_i)$ degrés de liberté.

3.2 Modèle linéaire gaussien, moindres carrés

Définition. Le *modèle linéaire gaussien homoscédastique à design fixe* est le modèle statistique $(\mathcal{N}(X\theta, \sigma^2 I_n))_{\theta\in\mathbb{R}^p}$, où $X \in \mathcal{M}_{n,p}(\mathbb{R})$ est fixée et de rang p ($p \leq n$). Le but est d'estimer le paramètre θ. On définit l'*estimateur par moindres carrés* par

$$\widehat{\theta} = \operatorname{argmin}_{u\in\mathbb{R}^p} \|Y - Xu\|^2.$$

Proposition. L'estimateur des moindres carrés vérifie

- $\widehat{\theta} = (X^\top X)^{-1} X^\top Y$
- $\mathbb{E}(\widehat{\theta}) = \theta$
- $\operatorname{Cov}(\widehat{\theta}) = \sigma^2 (X^\top X)^{-1}$

donc son risque quadratique est

$$\mathbb{E}(\|\widehat{\theta} - \theta\|^2) = \sigma^2 \operatorname{Tr}((X^\top X)^{-1}).$$

Théorème. On note $\widehat{\varepsilon} = Y - X\widehat{\theta}$ et $H = X(X^\top X)^{-1}X^\top$. Alors

- $\widehat{\theta} \sim \mathcal{N}(\theta, \sigma^2 (X^\top X)^{-1})$
- $\widehat{\varepsilon} \sim \mathcal{N}(0, \sigma^2 (I_n - H))$
- $\widehat{\theta}$ et $\widehat{\varepsilon}$ sont indépendants.

Application. Cela permet notamment de bâtir des estimateurs sur les paramètres.

4 Loi uniforme sur les permutations

4.1 Nombre d'inversions

Définition. Comme vu précédemment, la loi uniforme sur $\mathfrak{S}_n$ est définie par $\mathbb{P}(\{\sigma\}) : \frac{1}{n!}$.

Définition. Le *nombre d'inversions* d'une permutation σ est le nombre de couples (i, j) tels que $i < j$ et $\sigma(j) > \sigma(i)$. On note I_n qui à σ associe le nombre d'inversions de $\sigma \in \mathfrak{S}_n$.

Proposition. Si σ suit une loi uniforme sur $\mathfrak{S}_{n-1}$, et k suit une loi uniforme sur $\{1, \ldots, n\}$, alors la permutation obtenue en insérant n à la k-ième place dans σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. La loi de I_n est celle d'une somme de n variables indépendantes de loi $\mathcal{U}(\{0, k-1\})$ où $k \in \{1, \ldots, n\}$. En particulier, $\frac{6}{n^{3/2}}(I_n - \mathbb{E}(I_n))$ converge en loi vers une loi normale $\mathcal{N}(0,1)$.

4.2 Nombre de cycles

Proposition. On va construire une permutation aléatoire selon ses cycles. On met d'abord un 1. Puis si on a i nombres utilisés, on tire une $\mathcal{B}(\frac{1}{n-i+1})$. Si elle vaut 1, on clôt le cycle. Sinon on ne touche à rien. On choisit un nombre au hasard que l'on rajoute à la suite. On obtient alors une loi uniforme sur $\mathfrak{S}_n$.

Application. On note $K_n(\sigma)$ le nombre de cycles d'une permutation $\sigma \in \mathfrak{S}_n$. Sous la loi uniforme, K_n est une somme de $\mathcal{B}(\frac{1}{k})$ indépendantes, pour $k \in \{1, \ldots, n\}$. En particulier, $\frac{K_n - \ln(n)}{\sqrt{\ln(n)}}$ converge en loi vers une loi normale $\mathcal{N}(0,1)$.

4.3 Nombre de descentes

Définition. Soit $\sigma \in \mathfrak{S}_n$ on dit que σ présente une *descente au rang* i si $\sigma(i+1) < \sigma(i)$. On note $A_{n,k}$ le nombre de permutations de $\mathfrak{S}_n$ qui possèdent exactement k descentes.

Proposition. Soit $(U_1, \ldots, U_n) \sim \mathcal{U}(0,1)^{\otimes n}$. Pour presque tout $\omega \in \Omega$, il existe $\sigma(\omega)$ tel que $U_{\sigma(1)} < \ldots < U_{\sigma(n)}$. Alors σ suit une loi uniforme sur $\mathfrak{S}_n$.

Application. Avec les mêmes notations, la loi du nombre de descentes est la loi $\lfloor U_1 + \ldots + U_n \rfloor$. De plus, pour $\alpha \in [0,1]$, alors, lorsque n tend vers $+\infty$:

$$A(n, \alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}} \, \mathrm{e}^{-6(\alpha - 1/2)^2}.$$

Troisième partie

Développements d'algèbre et géométrie

Chapitre 69

Théorème de Perron-Frobenius

Ce théorème est le principe mathématique à la base de l'algorithme *PageRank* de Google. Pour trouver les pages les plus susceptibles d'être recherchées par un utilisateur réel, on considère un utilisateur virtuel qui se balade en grand temps, et on regarde sur quelle page il se trouve en grand temps.

Définition. Soit $A \in \mathcal{M}_n(\mathbb{R})$. A est dite *positive* (resp. *strictement positive*) et on note $A \geq 0$ (resp. $A > 0$) lorsque tous ses coefficients sont positifs (resp. strictement positifs). On définit de même un vecteur positif ou strictement positif.

Proposition. Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors A est strictement positive si et seulement si :

$$\forall x \geq 0 \text{ non nul }, Ax > 0.$$

Cette propriété donne une interprétation géométrique des matrices strictement positives : ce sont celles qui préservent le cône $\{x > 0\}$ invariant. Le théorème suivant assure que dans ce cône, il y a une direction qui va être laissée invariante par A.

Théorème. (de Perron-Frobenius)
Soit A une matrice réelle strictement positive. Alors le rayon spectral de A est l'unique valeur propre de A de module $\rho(A)$. De plus, le sous espace propre associé est de dimension 1, engendré par un vecteur strictement positif.

Démonstration. On assouplit légèrement le problème. Au lieu d'étudier directement le spectre de A, on s'intéresse à l'ensemble :

$$\mathcal{E} = \{t \in \mathbb{R}, \exists x \geq 0 \text{ non nul }, Ax - tx \geq 0\}.$$

Cet ensemble est non vide car il contient 0. On va montrer qu'il est égal à $[0, \rho(A)]$.

Tout d'abord, on vérifie que $\mathcal{E}$ n'est pas $\{0\}$. Pour cela, on considère $u = (1, \ldots, 1) > 0$ et on calcule :

$$[Au - tu]_i = \sum_{j=1}^{n} a_{i,j} - t = \sum_{j=1}^{n} a_{i,j} - nt$$

pour $t > 0$. On voit alors que puisque $A > 0$, la somme en question est strictement positive, et donc que pour t suffisamment petit, le vecteur $Au - tu$ est positif.

Montrons alors que $\mathcal{E}$ est convexe. Soient $0 < t' < t \in \mathcal{E}$. Pour x tel que $Ax - tx \geq 0$, on a :

$$Ax - t'x = Ax - tx + (t - t')x \geq 0,$$

ce qui conclut.

Maintenant, on vérifie que $\mathcal{E}$ est borné en voyant que $Ax \geq tx$ implique $t \leq \|A\|$ pour toute norme matricielle.

Enfin, on montre que $\mathcal{E}$ est fermé. Soit (t_k) une suite d'éléments de $\mathcal{E}$ qui converge vers t, et (x_k) une suite de vecteurs positifs non nuls telle que $Ax_k - t_k x_k \geq 0$. On peut supposer la suite (x_k) normalisée convergente quitte à extraire. Dès lors, on a, si x est la limite de (x_k), alors

$$Ax - tx \geq 0$$

et x est positif est non nul puisque $\{x \geq 0\}$ est fermé, donc $t \in \mathcal{E}$. $\mathcal{E}$ est donc de la forme $[0, \rho]$. Soit $x \geq 0$ non nul tel que $y = Ax - \rho x \geq 0$. Si y est non nul, on peut trouver un $\varepsilon > 0$ tel que $Ay - \varepsilon Ax > 0$. En effet, pour $1 \leq i \leq n$,

$$[Ay - \varepsilon Ax]_i = \sum_{j=1}^{n} a_{i,j}(y_j - \varepsilon x_j),$$

qui tend vers une quantité strictement positive quand ε tend vers 0 pour toute valeur de i. Dès lors,

$$Ay - \rho y = A(Ax) - (\rho + \varepsilon)Ax$$

et Ax est positif non nul donc ceci contredit la maximalité de ρ. Nécessairement y est nul c'est-à-dire ρ est valeur propre de A pour le vecteur propre $x = \frac{Ax}{\rho} > 0$.

Soit z un vecteur propre complexe associé à une valeur propre λ. Par inégalité triangulaire, si $|.|$ désigne le vecteur du module des coefficients et si z est un vecteur propre de A :

$$A|z| - |Az| = A|z| - |\lambda||z| \geq 0$$

donc $|\lambda| \in \mathcal{E} = [0, \rho]$. Ainsi $\rho = \rho(A)$. En particulier, en appliquant ce qui précède pour $\lambda = \rho$ et z vecteur propre associé à ρ, on obtient $|Az| = A|z|$. Ceci correspond au cas d'égalité dans l'inégalité triangulaire :

$$\forall k \in \{1, \ldots, n\}, [|Az|]_k = [A|z|]_k \text{ i.e. } |\sum_{j=1}^{n} a_{k,j} z_j| = \sum_{j=1}^{n} |a_{k,j} z_j|$$

Ainsi, pour tout k, il existe $\theta_k \in [0, 2\pi[$ tel que pour tout j, $|a_{k,j} z_j| = \mathrm{e}^{\mathrm{i}\theta_k} a_{k,j} z_j$. Comme $A > 0$,

on peut simplifier par $a_{k,j}$ et avoir l'existence de $\theta \in [0, 2\pi[$ tel que $z = \mathrm{e}^{\mathrm{i}\theta}|z|$. Alors

$$\langle x, z\rangle = \mathrm{e}^{\mathrm{i}\theta}\langle x, |z|\rangle \neq 0.$$

Par existence d'un supplémentaire orthogonal en dimension finie, tout vecteur du sous-espace propre associé à ρ et orthogonal à x est nul. Nécessairement, $\ker(A - \rho I_n) = \mathrm{Vect}(x)$. Enfin, si λ est une valeur propre de A de module ρ, alors $|z|$ est vecteur propre associé à ρ et comme $z = \mathrm{e}^{\mathrm{i}\theta}|z|$, alors z est colinéaire à x donc $\lambda = \rho$.

□

Remarque. On peut appliquer le résultat à $A^\top$ et avoir l'existence d'un vecteur $v > 0$ propre de $A^\top$ associé à la valeur propre ρ. On a alors nécessairement $\langle x, v\rangle > 0$.

Application. Soit $A > 0$ et $X_0 \geq 0$. On s'intéresse au système dynamique $X_{t+1} = AX_t$ pour $t \in \mathbb{N}^*$. Lorsque t tend vers l'infini, on a l'asymptotique suivante :

$$X_t \sim \rho(A)^t \langle v, X_0\rangle w$$

où $w = \frac{x}{\|x\|_1}$ et v comme ci-dessus choisi de telle sorte que $\langle v, w\rangle = 1$.

Démonstration. On considère la matrice de rang 1 $P = vw \in \mathcal{M}_n(\mathbb{R})$ dont l'image est $\mathrm{Vect}(w)$. Puisque $\langle v, w\rangle = 1$, on peut vérifier la matrice P est une matrice de projection qui commute avec A. On a donc une décomposition en sous-espaces stables de A :

$$\mathbb{R}^n = \mathrm{Im}(P) \oplus \ker(P).$$

Dans cette décomposition, on a :

$$X_0 = PX_0 + (X_0 - PX_0) =: cw + \widetilde{x}$$

avec $c = \langle v, X_0\rangle \geq 0$ et $\widetilde{x} \in \ker(P)$. On a alors pour tout $t \in \mathbb{N}$:

$$\frac{A^t}{\rho(A)^t} X_0 = cw + B^t \widetilde{x},$$

avec B^t la restriction de $\frac{A}{\rho(A)}$ à $\ker(P)$. B n'a que des valeurs propres de modules strictement inférieurs à 1 donc la suite de ses puissances tend vers 0. Ainsi :

$$\frac{A^t}{\rho(A)^t} X_0 = \frac{1}{\rho(A)^t} X_t \xrightarrow{t\to\infty} cw = \langle v, X_0\rangle w.$$

On a donc bien le résultat voulu. □

Référence : La preuve de Perron-Frobenius provient du sujet Analyse et Probabilités de l'agrégation externe de 2012. L'asymptotique provient de moi il me semble.

Chapitre 70

Déterminants de Gram et de Cauchy

On montre un résultat d'algèbre sur les matrices symétriques qui trouve une application dans un problème d'analyse.

Dans tout ce développement, E désignera un espace préhilbertien (éventuellement hermitien quitte à remplacer la transposée par l'adjoint et à changer le produit scalaire).

Définition. Soit $(x_1, \dots, x_n)$ une famille de vecteurs de E. La *matrice de Gram* associé à cette famille est la matrice :

$$M = (\langle x_i, x_j \rangle)_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{R}).$$

Son déterminant est appelé *déterminant de Gram* de la famille $(x_1, \dots, x_n)$.

Proposition. Les matrices de Gram de taille n sont exactement les matrices symétriques positives de taille n.

Démonstration. Si M est la matrice de Gram de $(x_1, \dots, x_n)$, alors $M = X^\top X$, où X est la matrice dont les colonnes sont les $(x_j)_j$, donc M est symétrique positive. Réciproquement, si A est symétrique positive, c'est la matrice de Gram de la famille des colonnes de la racine carrée de A. □

Proposition. Soit $(x_1, \dots, x_n)$ une famille de vecteurs de E. Elle est liée si et seulement si son déterminant de Gram est nul.

Démonstration. Si la famille est liée, on peut se ramener à une colonne de zéros en faisant des opérations élémentaires sur les colonnes par bilinéarité du produit scalaire. Réciproquement, si la matrice de Gram est de déterminant nul, alors ses colonnes sont liées, donc il existe des scalaires $(\lambda_2, \dots, \lambda_n)$ tels que :

$$\forall i \in [\![1, n]\!], \langle x_i, x_1 \rangle = \sum_{k=2}^{n} \lambda_k \langle x_i, x_k \rangle$$

soit par bilinéarité :

$$\forall i \in [\![1, n]\!], \left\langle x_i, x_1 - \sum_{k=2}^{n} \lambda_k x_k \right\rangle = 0.$$

Ainsi, $x_1 - \sum_{k=2}^{n} \lambda_k x_k$ appartient à la fois à $\mathrm{Vect}(x_1, \dots, x_n)$ et à son orthogonale, *i.e.* est nul, donc la famille est liée. □

Exemple. Si X est un vecteur aléatoire dans $\mathbb{R}^n$ dont chaque composante admet une variance, sa matrice de covariance correspond à la matrice de Gram de la famille $(X_1-\mathbb{E}(X_1),\dots,X_n-\mathbb{E}(X_n))$ pour le produit scalaire de l'espérance. Le déterminant de Gram associé est appelé *variance généralisée* de X.

Théorème. Soit V un sous-espace vectoriel de E admettant une base $(x_1,\dots,x_n)$. Alors pour tout $u\in E$:

$$d(u,V)^2=\frac{G(u,x_1,\dots,x_n)}{G(x_1,\dots,x_n)}$$

Démonstration. On sait que $d(u,V)^2=\|x\|^2-\|p_V(x)\|^2=\|x-p_V(x)\|^2$ par Pythagore. Notons $y=p_V(u)$ et $z=x-y$. Pour tout i, on a $\langle x_i,y\rangle=\langle x_i,u\rangle$, ainsi, si on note M la matrice de Gram de la famille $(x_1,\dots,x_n,u)$, on a :

$$M=\begin{pmatrix}\langle x_1,x_1\rangle & \dots & \langle x_1,x_n\rangle & \langle x_1,y\rangle\\ \dots & \dots & \dots & \dots\\ \langle x_n,x_1\rangle & \dots & \langle x_n,x_n\rangle & \langle x_n,y\rangle\\ \langle y,x_1\rangle & \dots & \langle y,x_n\rangle & \|y\|^2+\|z\|^2\end{pmatrix}$$

Puis, par multlinéarité du déterminant :

$$\det(M)=\begin{vmatrix}\langle x_1,x_1\rangle & \dots & \langle x_1,x_n\rangle & \langle x_1,y\rangle\\ \dots & \dots & \dots & \dots\\ \langle x_n,x_1\rangle & \dots & \langle x_n,x_n\rangle & \langle x_n,y\rangle\\ \langle y,x_1\rangle & \dots & \langle y,x_n\rangle & \|y\|^2\end{vmatrix}+\begin{vmatrix}\langle x_1,x_1\rangle & \dots & \langle x_1,x_n\rangle & 0\\ \dots & \dots & \dots & \dots\\ \langle x_n,x_1\rangle & \dots & \langle x_n,x_n\rangle & 0\\ \langle y,x_1\rangle & \dots & \langle y,x_n\rangle & \|z\|^2\end{vmatrix}$$

Or le premier déterminant vaut $G(x_1,\dots,x_n,y)=0$ car la famille en question est liée, et le deuxième vaut $\|z\|^2G(x_1,\dots,x_n)$, ce qui fournit le résultat. □

Exemple. Dans le cas d'un vecteur gaussien centré $(X=X_1,\dots,X_n,Y)\in \mathrm{L}^2$, la projection correspondant à l'espérance conditionnelle (et celle-ci étant une combinaison linéaire des (X_i) car on est dans le monde gaussien), on a l'identité suivante :

$$\mathbb{E}(\mathrm{Var}(Y|X))=\mathbb{E}((Y-\mathbb{E}(Y|X))^2)=\mathbb{E}((Y-\mathbb{E}(Y|X_1,\dots,X_n))^2)=\frac{\mathrm{Var}(Y,X)}{\mathrm{Var}(X)}.$$

Théorème. (Inégalité de Hadamard)
Soit $(x_1,\dots,x_n)$ une famille de vecteurs non nuls de E. Alors

$$G(x_1,\dots,x_n)\le\prod_{i=1}^n\|x_i\|^2,$$

avec égalité si et seulement si la famille en question est orthogonale.
Plus généralement,

$$|\det(x_1,\dots,x_n)|\le\prod_{i=1}^n\|x_i\|.$$

Démonstration. On peut toujours supposer que la famille est libre. Dans ce cas, on procède par récurrence. Si $n=1$, le résultat est immédiat, on le suppose vrai pour un entier n quelconque.

On considère alors une famille libre $(x_1, \dots, x_n)$ qui engendre V et un vecteur x_{n+1} libre avec cette famille. Notons y la projection orthogonale de x_{n+1} sur V. D'après le résultat précédent, on a

$$G(x_1, \dots, x_{n+1}) = G(x_1, \dots, x_n)\|x - y\|^2 \leq G(x_1, \dots, x_n)\|x_{n+1}\|^2,$$

d'où le résultat par hypothèse de récurrence. Le cas d'égalité à l'étape n nécessite d'avoir eu à chaque étape $\|x_{k+1} - y\|^2 = \|x_{k+1}\|^2$ *i.e.* x_{k+1} est orthogonal à tous les vecteurs d'avant *i.e.* la famille est orthogonale. Pour la généralisation, il suffit de se rappeler que l'on peut écrire la matrice de Gram comme $X^\top X$, où X est de déterminant $\det(x_1, \dots, x_n)$. □

Exemple. Si on reprend l'exemple de notre vecteur aléatoire X, on obtient :

$$\mathrm{Var}(X) \leq \prod_{i=1}^{n} \mathrm{Var}(X_i),$$

avec égalité si et seulement si le vecteur est décorrélé, ou dans le cas gaussien si et seulement si le vecteur est indépendant.

Application. Montrer que l'expression $\int_0^1 (1 + a_1 + a_2x^2 + \ldots + a_nx^n)^2 \,\mathrm{d}x$ atteint son minimum μ en un unique n-uplet.

Démonstration. On considère donc E l'ensemble des fonctions continues sur $[0, 1]$ muni du produit scalaire canonique. Le problème revient donc à calculer la distance du vecteur 1 à la famille $(x, x^2, \dots, x^n)$. On sait qu'elle est atteinte en un unique n-uplet et d'après le premier théorème,

$$\mu = \frac{G(1, x, \dots, x^n)}{G(x, \dots, x^n)}$$

Pour $1 \leq i, j \leq n$, $\langle x^i, x^j \rangle = \int_0^1 x^{i+j} \,\mathrm{d}x = \frac{1}{i+j+1}$. Les déterminants à calculer sont donc des déterminants de Cauchy (*cf.* fait suivant) :

$$\mu = \frac{\prod_{1 \leq i < j \leq n+1} (i-j)^2}{\prod_{1 \leq i,j \leq n+1} (i+j-1)} \frac{\prod_{1 \leq i,j \leq n} (i+j+1)}{\prod_{1 \leq i < j \leq n} (i-j)^2},$$

qu'on peut simplifier pour obtenir $\mu = \frac{1}{(n+1)^2}$. □

Fait. Pour des complexes $(a_j)_j, (b_j)_j$ tels que $a_i + b_j \neq 0$ pour tout (i, j),

$$\Delta_n := \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & \vdots & \vdots \\ \frac{1}{a_1+b_n} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix} = \frac{\prod_{i<j} (a_j - a_i) \prod_{i<j} (b_j - b_i)}{\prod_{1 \leq i,j \leq n} (a_i + b_j)}.$$

Démonstration. On suppose que les $(a_i)_i$ sont distincts deux à deux (sinon le déterminant est nul et la formule est vraie). On considère R la fraction rationnelle

$$R(X) = \frac{(b_1 - X) \dots (b_{n-1} - X)}{(a_1 + X) \dots (a_n + X)},$$

à laquelle on peut appliquer le théorème de décomposition en éléments simples. Il existe donc des scalaires $\lambda_1, \ldots, \lambda_n$ tels que

$$R(X) = \frac{\lambda_1}{X + a_1} + \ldots + \frac{\lambda_n}{X + a_n}.$$

Par la méthode des résidus, on obtient :

$$\lambda_k = \frac{\prod_i^{n-1}(b_i + a_k)}{\prod_{i \neq k}(a_i - a_k)} \neq 0.$$

Ainsi, si on note $(L_i)_i$ les lignes de la matrice dont on veut le déterminant, on a :

$$\Delta_n = \frac{1}{\lambda_n} \begin{vmatrix} L_1 \\ \vdots \\ \sum_{i=1}^n \lambda_i L_i \end{vmatrix} = \frac{1}{\lambda_n} \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & \frac{1}{a_1+b_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{a_1+b_n} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & \frac{1}{a_{n-1}+b_n} \\ R(b_1) & \ldots & R(b_{n-1}) & R(b_n) \end{vmatrix}$$

Puisque $R(b_k) = 0$ dès que $k \neq n$, on peut simplement développer par rapport à la dernière ligne et avoir :

$$\Delta_n = \frac{R(b_n)}{\lambda_n} \Delta_{n-1}$$

soit :

$$\Delta_n = \frac{\prod_{i=1}^{n-1}(b_i - b_n)}{\prod_{i=1}^{n}(b_n + a_i)} \frac{\prod_{i=1}^{n-1}(a_i - a_n)}{\prod_{i=1}^{n-1}(b_i + a_n)} \Delta_{n-1}.$$

Cela donne le résultat par récurrence. □

Référence. X.Gourdon, *Analyse*. Pour la partie probabilités, j'ai écrit ça rapidement et c'est donc peut-être imprécis.

Chapitre 71

Cycles d'une permutation aléatoire

On regarde dans ce développement comment générer une permutation selon ses cycles et comment en déduire la loi du nombre de cycles d'une permutation aléatoire.

Proposition. On se propose de construire une permutation comme suit :

- On ouvre un cycle (1
- On tire une suite η_j indépendante avec $\eta_j \sim \mathcal{B}\left(\frac{1}{n-j+1}\right)$
- Pour i allant de 1 à n, tant que $\eta_i = 0$ on rajoute des éléments dans le cycle en cours en prenant uniformément des éléments parmi ceux qui restent (et indépendamment des autres) et si $\eta_j = 1$ on clôt le cycle et on en ouvre un nouveau (si $j \leq n$) en ajoutant un élément uniformément au début du cycle.

La permutation σ construite est alors uniforme sur $\mathfrak{S}_n$.

Démonstration. Soit φ une permutation de $\mathfrak{S}_n$ dont le cycle contenant 1 est de longueur $1 \leq k \leq n$. Alors, si les U_i sont indépendantes

$$\mathbb{P}(\sigma(1) = \varphi(1), \ldots, \sigma^k(1) = 1) = \mathbb{P}(\eta_1 = \ldots = \eta_{j-1} = 0, \eta_k = 1) \prod_{j=1}^{k-1} \frac{1}{n-j},$$

soit donc

$$\begin{aligned}
\mathbb{P}(\sigma(1) = \varphi(1), \ldots, \sigma^k(1) = 1) &= \frac{1}{n-k+1} \prod_{j=1}^{k-1} \frac{1}{n-j} \left(1 - \frac{1}{n-j+1}\right) \\
&= \frac{1}{n-k+1} \prod_{j=1}^{k-1} \frac{1}{n-j} \frac{n-j}{n-j+1} \\
&= \frac{1}{n-k+1} \prod_{j=1}^{k-1} \frac{1}{n-j+1} = \frac{(n-k)!}{n!}.
\end{aligned}$$

Ainsi, on formule l'hypothèse de récurrence suivante

$\mathcal{H}(p)$: l'algorithme crée des permutations aléatoires sur tout ensemble de cardinal p.

Avec les mêmes notations, on peut donc utiliser l'hypothèse de récurrence et on a alors

$$\begin{aligned}\mathbb{P}(\sigma=\varphi) &= \mathbb{P}(\sigma(1)=\varphi(1),\dots,\sigma^k(1)=1)\,\mathbb{P}(\sigma(j)=\varphi(j) \text{ sur les autres valeurs})\\ &= \frac{(n-k)!}{n!}\frac{1}{(n-k)!} = \frac{1}{n!},\end{aligned}$$

donc on a bien une loi uniforme sur $\mathfrak{S}_n$. □

Corollaire. Le nombre de permutations ayant k cycles dans $\mathfrak{S}_n$ est le coefficient en X^k du polynôme $\prod_{j=0}^{n-1}(X-j)$.

Démonstration. Avec les notations de l'algorithme précédent, le nombre de cycles de σ est la somme des η_j. Ainsi, le nombre de cycles K_n d'une permutation est égal en loi à une somme de Bernoulli indépendantes de paramètres $\frac{1}{j}$ pour j allant de 1 à n. Ainsi, la série génératrice de K_n, notée G, vaut

$$G(t) = \prod_{j=1}^{n}\left(\left(1-\frac{1}{j}\right)+\frac{1}{j}t\right).$$

Ainsi, si on note $C_{n,k}$ le nombre de permutations avec k cycles dans $\mathfrak{S}_n$, on a :

$$n!G(t) = \sum_{k=1}^{n} C_{n,k}t^k = \prod_{j=1}^{n}(t+(j-1)) = \prod_{j=0}^{n-1}(t-j).$$

□

Référence : un DM de Jürgen Angst, sans doute inspiré du Feller vol. I.

Chapitre 72

Formes quadratiques, décomposition en valeurs singulières

Dans ce développement, on montre le théorème spectral à partir de la réduction des formes quadratiques. On s'intéresse après aux valeurs singulières avec plein d'applications. On peut recaser des petits bouts de tout cela un peu partout.

Proposition. (réduction des formes quadratiques)
Soit q une forme quadratique sur E un K-espace vectoriel de dimension finie où K est un corps de caractéristique différente de deux. Alors q admet une base orthogonale.

Démonstration. Quitte à se ramener à la restriction de q à un supplémentaire du noyau, on peut supposer que q est non dégénérée. On procède alors par récurrence sur la dimension de l'espace, en prenant x puis en considérant son orthogonal, auquel on applique l'hypothèse de récurrence. En ajustant les vecteurs de la base ainsi obtenue, on remarque que les coefficients de la matrice diagonale dans cette base peuvent être pris modulo les carrés de K. □

Proposition. (orthogonalisation simultanée)
On se place ici dans un espace euclidien avec un produit scalaire $\langle \cdot, \cdot \rangle$. Soit q une forme quadratique (que l'on peut supposer non dégénérée comme avant). Il existe une base orthonormée de E qui est orthogonale pour q.

Démonstration. On va encore procéder par récurrence. Notre but est ici de trouver un x anisotrope pour q tel que les orthogonaux de x pour q et le produit scalaire coïncident. Ces deux espaces étant des hyperplans, leur égalité correspond à la proportionnalité des formes linéaires dont ils sont les noyaux, à savoir $b(x, \cdot)$ et $\langle x, \cdot \rangle$. Représentons alors cette première forme linéaire avec le théorème de Riesz, il existe $a(x) \in E$ tel que $b(x, \cdot) = \langle a(x), \cdot \rangle$. Par unicité de $a(x)$, a définit un opérateur linéaire sur E, qui est auto-adjoint par symétrie de b. Ainsi, notre vecteur anisotrope existe si et seulement si notre opérateur a admet un vecteur propre, ce qui existe en considérant un vecteur propre complexe puis en conjuguant. □

Remarque. À chaque forme quadratique de E on peut associer un opérateur auto-adjoint de E et réciproquement.

Théorème. (valeurs singulières et Courant-Fischer)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$, il existe $P \in \mathrm{O}_n(\mathbb{R})$ et $Q \in \mathrm{O}_p(\mathbb{R})$ telles que

$$A = P\mathrm{diag}\left(\sigma_1(A), \ldots, \sigma_r(A), 0, \ldots, 0\right) Q^\top =: P\Sigma Q^\top,$$

avec $\sigma_1(A) \geq \ldots \geq \sigma_r(A) \geq 0$ les *valeurs singulières* de A. Si G_d est l'ensemble des sous-espaces vectoriels de $\mathbb{R}^p$ de dimension d, alors :

$$\sigma_k(A) = \min_{W \in G_{p-k+1}} \max_{\substack{x \in W \\ \|x\|=1}} \|Ax\| = \max_{W \in G_p} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.$$

Démonstration. D'après le théorème spectral, il existe une matrice $V \in \mathrm{O}_p(\mathbb{R})$ telle que $A^\top A = V\,\mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2)\,V^\top$ ($\sigma_i \geq 0$) où les valeurs propres sont des carrés car positives et où on les a rangées par ordre décroissant. Il reste alors à trouver $U \in \mathrm{O}_n(\mathbb{R})$ telle que $AV = U\Sigma$. Si on note v_i les colonnes de V, on cherche à avoir $Av_i = \sigma_i u_i$ pour $i \leq r$ et $Av_i = 0$ sinon. La deuxième condition est vérifiée puisque $\ker(A) = \ker(A^\top A)$ et pour la première on prend $u_i = \frac{1}{\sigma_i} Av_i$ et on vérifie qu'elle convient. Il reste juste à compléter notre base par une base orthonormée de $\mathbb{R}^n$.

Montrons le théorème minmax. Soit W un sous-espace de dimension k. Par Grassmann, on peut trouver un vecteur de W qui est combinaison linéaire de $e_k, \ldots, e_n$ les derniers vecteurs propres de $A^\top A$. On a alors $\|Ax\|^2 = \langle A^\top Ax, x\rangle = \sum_{i=k}^n \sigma_i^2 x_i^2 \leq \sigma_k^2$, et le minimum est atteint lorsqu'on prend pour W le sous-espace engendré par les k premiers vecteurs propres. □

Théorème. (Eckart-Young)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$ de SVD $A = P\Sigma Q^\top$. Si on note Σ_k la matrice où l'on a gardé uniquement les k premières valeurs singulières, alors $A_k := P\Sigma_k Q^\top$ est la meilleure approximation de A de rang k pour la norme spectrale.

Démonstration. Par décomposition en valeurs singulières, cela revient à minimiser $\|B - \Sigma\|$ pour $B \in \mathcal{M}_{n,p}(\mathbb{R})$ de rang k. Par la formule de Grassmann et le théorème du rang, il existe un vecteur dans $\mathrm{Vect}(e_1, ldots, e_p) \cap \ker(B)$ ((e_i) est la base canonique de $\mathbb{R}^p$). On a alors $(\Sigma_k - B)x = \sum_{i=1}^{k+1} x_i \sigma_i e_i$ d'où $\|\Sigma - B\|^2 \geq \|\sum_{i=1}^{k+1} \sigma_i^2 |x_i|^2\| \geq \sigma_{k+1}^2$. En prenant $B = \Sigma_k$, on obtient le résultat voulu. □

Définition. Pour $1 \leq k \leq p$, on définit la k-ième *norme de Ky Fan* de $A \in \mathcal{M}_{n,p}(\mathbb{R})$ comme la somme de ses k premières valeurs singulières, on la note $N_k(A)$.

Théorème. La norme k de Ky Fan vérifie la formulation variationnelle suivante :

$$N_k(A) = \max_{\substack{U^\top U = V^\top V = I_k \\ U \in \mathcal{M}_{n,k}(\mathbb{R}) \\ V \in \mathcal{M}_{p,k}(\mathbb{R})}} \mathrm{Tr}(U^\top AV).$$

C'est donc une norme.

Démonstration. On se fixe un $k \in \{1, \dots, p\}$. Soit $i \in \{1, \dots, k\}$, et W de dimension i dans $\mathbb{R}^k$, soient U, V comme dans le max. On applique le minmax une première fois à la matrice $U^\top AV$:

$$\sigma_i(U^\top AV) = \max_{W \in G_i(\mathbb{R}^k)} \min_{\substack{\|x\|=1 \\ x \in W}} \|U^\top AVx\|,$$

ce que l'on peut réécrire en considérant le changement de variable $y = Vx$:

$$\sigma_i(U^\top AV) = \max_{W \in G_i(\mathbb{R}^k)} \min_{\substack{\|y\|=1 \\ y \in V(W)}} \|U^\top Ay\|,$$

puisque pour $x \in \mathbb{R}^k$, $\|Vx\|^2 = \langle Vx, Vx\rangle = \langle V^\top Vx, x\rangle = \|x\|^2$. Or pour tout $z \in \mathbb{R}^p$, $\|U^\top z\| \leq \|z\|$ puisque $\|U^\top\|^2 = \rho(U^\top U) = \rho(UU^\top) = \|U\|_2 = 1$. Ainsi pour tout $y \in V(W)$:

$$\|U^\top Ay\| \leq \|Ay\| \text{ donc } \min_{\substack{\|y\|=1 \\ y \in V(W)}} \|U^\top Ay\| \leq \min_{\substack{\|y\|=1 \\ y \in V(W)}} \|Ay\|$$

En prenant le maximum sur $W \in G_i(\mathbb{R}^k)$, on obtient $\sigma_i(U^\top AV) \leq \max_{W \in G_i(\mathbb{R}^k)} \min_{\substack{\|y\|=1 \\ y \in V(W)}} \|Ay\|$. En considérant W_0 pour lequel le maximum de droite est atteint, on a

$$\sigma_i(U^\top AV) \leq \min_{\substack{\|y\|=1 \\ y \in V(W_0)}} \|Ay\|.$$

Mais W_0 est de dimension i dans $\mathbb{R}^k$ et V est injective de $\mathbb{R}^k$ dans $\mathbb{R}^p$ donc $V(W_0) \in G_i(\mathbb{R}^p)$. Somme toute :

$$\sigma_i(U^\top AV) \leq \min_{\substack{\|y\|=1 \\ y \in V(W_0)}} \|Ay\| \leq \max_{W' \in G_i(\mathbb{R}^p)} \min_{\substack{\|y\|=1 \\ y \in W'}} \|Ay\| = \sigma_i(A).$$

En particulier, $N_k(U^\top AV) \leq N_k(A)$.

Par ailleurs, si $B \in \mathcal{M}_k(\mathbb{R})$ on a, si $B = PDQ^\top$ est sa SVD :

$$|\operatorname{Tr}(B)| = |\operatorname{Tr}(PDQ^\top)| = |\operatorname{Tr}(DQ^\top P)| \leq \sum_{i=1}^{k} \sigma_i(B)|(Q^\top P)_{i,i}| \leq N_k(B).$$

En appliquant ce résultat à $B = U^\top AV$, on a $\operatorname{Tr}(U^\top AV) \leq N_k(U^\top AV) \leq N_k(A)$. Et le maximum est atteint lorsque U et V sont les k premiers vecteurs singuliers de A à gauche et à droite.

□

Proposition. Pour $A \in \mathcal{M}_{n,p}$, on définit sa *pseudo-inverse* en inversant les valeurs singulières dans la décomposition en valeurs singulières, et on la note $A^\dagger = V\Sigma^\dagger U^\top$. La solution de $Ax = b$ au sens des moindres carrés de norme minimale est unique et vaut $A^\dagger b$.

Démonstration. On montre avec du calcul différentiel que x est solution au sens des moindres carrés si et seulement si c'est une solution de $A^\top Ax = A^\top b$. On vérifie ensuite que notre solution

x_b est bien solution de cette équation. Par ailleurs on montre que $x - x_b \in \ker(A^\top A)$ pour x autre solution au sens des moindres carrés, et que $x_b \in \ker(A^\top A)^\perp$ donc le résultat suit par Pythagore. □

Pour finir, une preuve alternative du théorème d'Eckart-Young qui est plus forte.

Proposition. Soient X et Y deux matrices rectangulaires de $\mathcal{M}_{n,p}(\mathbb{R})$ $(n \geq p)$. Alors

$$\sigma_{i+j-1}(X+Y) \leq \sigma_i(X) + \sigma_j(Y) \qquad (i+j-1 \leq p).$$

Démonstration. On utilise les formulations variationnelles de Courant-Fischer. Soit V de dimension $p-i+1$, soit W de dimension $p-j+1$ tels que

$$\sigma_i(X) = \max_{V \cap S} \|Xx\| \text{ et } \sigma_j(Y) = \max_{W \cap S} \|Yy\|.$$

Ainsi, si F est un sous-espace de $\mathbb{R}^p$ de dimension $i+j-1$, la formule de Grassmann implique que $F \cap V \cap W \neq \{0\}$. Soit donc $z \in F \cap V \cap W \cap S$.

$$\|(X+Y)z\| \leq \|Xz\| + \|Yz\| \leq \sigma_i(X) + \sigma_j(Y),$$

d'où le résultat en prenant le min sur z puis le max sur F. □

Corollaire. Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$ de SVD $A = P\Sigma Q^\top$. Si on note Σ_k la matrice où l'on a gardé uniquement les k premières valeurs singulières, alors $A_k := P\Sigma_k Q^\top$ est la meilleure approximation de A de rang k pour la norme spectrale, et pour la norme de Frobenius.

Démonstration. Si B est une matrice de rang k, alors pour $j = k+1$ dans l'inégalité précédente, $Y = B$ et $X = A - B$, on a

$$\sigma_{i+k}(A) \leq \sigma_i(A-B) \qquad (1 \leq i \leq p-k).$$

Pour avoir le résultat pour la norme spectrale, il suffit de prendre $i = 1$, et pour la norme de Frobenius, on somme sur les i :

$$\|A-B\|_{\mathrm{F}}^2 \geq \sum_{i=1}^{p-k} \sigma_i^2(A-B) \geq \sum_{i=k+1}^{p} \sigma_i^2(A).$$

On atteint bien l'égalité pour $B = A_k$. □

Référence : La plupart de ce document est personnelle, le début sur la SVD vient de la page de Matthias Hostein, et la proposition 9 vient des inégalités de Weyl d'un oraux XENS.

Chapitre 73

Endomorphismes sur un corps fini

Proposition. (lemme de Fitting)
Soient E un K-espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Les suites $(\ker(u^k))$ et $\mathrm{Im}(u^k)$ sont respectivement croissante et décroissante à partir d'un certain rang n_0 et on a alors

$$E = \ker(u^{n_0}) \oplus \mathrm{Im}(u^{n_0}).$$

Démonstration. Les monotonies sont évidentes et elles stationnent au même rang par le théorème du rang, notons-le n_0. Soit $\ker(u^{n_0}) \cap \mathrm{Im}(u^{n_0})$ Pour $y = u^{n_0}(x)$, $u^{n_0}(y) = u^{2n_0}(x) = 0$ donc $x \in \ker(u^{n_0})$ ce qui donne exactement $y = 0$ donc la somme est directe. On conclut par égalité des dimensions. Dès lors, on voit que u restreint à $\ker(u^{n_0})$ est nilpotent et que son homologue sur $\mathrm{Im}(u^{n_0})$ est bijectif car $\mathrm{Im}(u^{n_0+1}) = \mathrm{Im}(u^{n_0})$. □

On obtient donc une décomposition dite de Fitting (F, G, v, w) où les F, G sont les espaces en question et les v, w sont les restrictions.

Proposition. La décomposition de Fitting d'un endomorphisme est unique.

Démonstration. Soit (F, G, v, w) une décomposition de Fitting de u avec w nilpotent. En écrivant la forme par blocs dans la décomposition adaptée, on montre que F et G sont bien les espaces que l'on a mis en évidence auparavant. Nécessairement, v et w sont les restrictions souhaitées. □

Théorème. Soit K fini de cardinal q. Le nombre n_d de matrices nilpotentes de $\mathcal{M}_d(K)$ est $q^{d(d-1)}$.

Démonstration. Soit X_k l'ensemble des couples (F, G) de sous-espaces de K^d tels que $\dim(F) = k$ et $F \oplus G = K^d$, soit m_k son cardinal. On a alors une action transitive de G sur X_k donnée par $g \cdot (F, G) = (g(F), g(G))$, et un stabilisateur est alors isomorphe à $\mathrm{GL}_k(K) \times \mathrm{GL}_{d-k}(K)$ d'où $m_k = \frac{g_d}{g_k g_{d-k}}$ où l'on a noté $g_i = |\mathrm{GL}_i(K)|$. La décomposition de Fitting donne alors une bijection entre $\mathcal{L}(E)$ et un couple de l'un des X_k, avec un nilpotent de $\mathcal{M}_k(K)$ et un automorphisme de $\mathrm{GL}_{n-k}(K)$, soit

$$|\mathcal{L}(E)| = \sum_{k=0}^{d} m_k \, n_k \, g_{d-k}.$$

Soit

$$\frac{n_d}{g_d} = \frac{q^{d^2}}{g_d} - \frac{q^{(d-1)^2}}{g_{d-1}}$$

d'où

$$n_d = q^{d^2} - q^{(d-1)^2}(q^d - 1)q^{d-1} = q^{d(d-1)}.$$

□

Proposition. Le nombre de matrices diagonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|}.$$

Démonstration. Soit $A \in \mathcal{M}_n(\mathbb{F}_q)$. Alors A est diagonalisable si et seulement si

$$A = \bigoplus_{i=1}^{q} \ker(A - \zeta_i I_n),$$

où $\{\zeta_i\}$ est une énumération de $\mathbb{F}_q$. Soit alors $(n_1, \ldots, n_q)$ fixé tel que $n_1 + \cdots + n_q = n$. On va chercher à compter le nombre de diagonalisables $u \in \mathcal{L}(\mathbb{F}_q)$ tels que $\ker(u - \zeta_i \mathrm{id}) = n_i$ pour tout i. Ce nombre correspond au nombre de décompositions en sous-espaces $E = \bigoplus_{i=1}^{q} F_i$ telles que $\dim(F_i) = n_i$. En effet, si on note $\mathcal{E}_{(n_i)}$ l'ensemble de telles décompositions, on construit une bijection par

$$u \mapsto (\ker(u - \zeta_i \mathrm{id}))_i, (F_i) \mapsto \sum_{i=1}^{q} \zeta_i \mathrm{id}_{F_i}.$$

Dénombrons-donc $\mathcal{E}_{(n_i)}$. Pour cela, on considère l'action de groupes :

$$(u, (F_i)) \mapsto (u(F_i)).$$

C'est bien une action de groupes, qui est transitive. On n'a donc qu'une seule orbite, égale à $\mathcal{E}_{(n_i)}$. Pour conclure, il nous faut le cardinal du stabilisateur, qui est isomorphe au produit $(\mathrm{GL}_{n_1}(\mathbb{F}_q) \times \cdots \times \mathrm{GL}_{n_q}(\mathbb{F}_q))$. □

Proposition. Le nombre de matrices trigonalisables dans $\mathcal{M}_n(\mathbb{F}_q)$ est

$$\sum_{n_1+\ldots+n_q=n} \frac{|\mathrm{GL}_n(\mathbb{F}_q)|}{|\mathrm{GL}_{n_1}(\mathbb{F}_q)|\cdots|\mathrm{GL}_{n_q}(\mathbb{F}_q)|} \prod_{i=1}^{q} q^{n_i(n_i-1)}.$$

Démonstration. On considère toujours $\mathcal{E}_{(n_i)}$ l'ensemble des décompositions en somme directe, et la bijection est alors cette fois-ci :

$$u \mapsto (\ker(u - \zeta_i \mathrm{id})^{m_i})_i, (u - \zeta_1 \mathrm{id}_{n_1}, \ldots, u - \zeta_q \mathrm{id}_{n_q})_i \text{ ou } ((w_i), (F_i))_i \mapsto \sum_i (\zeta_i \mathrm{id}_{F_i} + w_i).$$

□

Référence : Philippe Caldero, *Carnet de voyage en Algébrie*

Chapitre 74

Loi de réciprocité quadratique

Dans ce développement, on montre la loi de réciprocité quadratique par le biais des formes quadratiques. Dans tout le document, p désigne un nombre premier impair.

Définition. Soit $a \in \mathbb{Z}$ on définit le *symbole de Legendre*, noté $\left(\frac{a}{p}\right)$ comme valant 0 si p divise a, 1 si a est un carré modulo p et -1 sinon.

Proposition. (critère d'Euler)
Avec les mêmes notations :

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p.$$

En particulier, le symbole de Legendre est multiplicatif.

Démonstration. Si a est un carré, on a directement par Lagrange $a^{\frac{p-1}{2}} = 1$ dans $\mathbb{F}_p^*$. Mais puisque $\mathbb{F}_p$ est un corps, les carrés non nuls sont les seules racines de $X^{\frac{p-1}{2}} - 1$. Si $a^{\frac{p-1}{2}}$ ne vaut pas 1, c'est forcément -1 puisque son carré vaut 1. □

Proposition. Le nombre -1 est un carré modulo p si et seulement si p est congru à 1 modulo 4.

Démonstration. Cela découle directement du résultat précédent. □

Proposition. Le symbole de Legendre $\left(\frac{2}{p}\right)$ vaut $(-1)^{\frac{p^2-1}{8}}$, *i.e.* 1 si p est congru à ± 1 modulo 8 et ± 8 sinon.

Démonstration. Il s'agit de regarder le produit $\prod_{k=1}^{(p-1)/2}(2k)$ et de le découper au milieu suivant p modulo 4, d'une part on a $\left(\frac{2}{p}\right)((p-1)/2)!$ et d'autre part le résultat voulu fois $((p-1)/2)!$, par lequel on peut donc simplifier car il est inversible (par primalité de p). □

Lemme. Soit q premier impair et $a \in \mathbb{F}_q^*$.

$$|\{x \in \mathbb{F}_q, ax^2 = 1\}| = 1 + \left(\frac{a}{q}\right)$$

Démonstration. L'élément a est un carré mod q si et seulement si on inverse l'est, si et seulement si $X^2 - a$ a deux racines distinctes, et notre cardinal est ce nombre de racines, valant 2 ou 0 selon que a soit un carré ou non (car $\mathrm{car}(K) > 2$). □

Théorème. (loi de réciprocité quadratique)
Soient p, q deux nombres premiers impairs. Alors

$$\left(\frac{p}{q}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

Démonstration. La preuve se base sur deux dénombrements de

$$X := \{(x_1, \ldots, x_p) \in \mathbb{F}_q^p, \sum_{i=1}^{p} x_i^2 = 1\} = f^{-1}(1),$$

où f est la forme quadratique sur $\mathbb{F}_q$ de matrice I_p dans la base canonique. On définit alors $d = \frac{p-1}{2}$, $a = (-1)^d$ et $J \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{F}_q)$. Soit alors f' la forme quadratique de matrice

$$\begin{pmatrix} J & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & J & 0 \\ 0 & \ldots & 0 & a \end{pmatrix} \in \mathcal{M}_p(\mathbb{F}_q).$$

Cette matrice a même rang et même discriminant que I_p. Par classification des formes quadratiques sur les corps finis, M et I_p sont congruentes ; en particulier

$$|X| = |\{x \in \mathbb{F}_q^p, f'(x) = 1\}| = |\{(y_1, z_1, \ldots, y_d, z_d, t) \in \mathbb{F}_q^p, 2\sum_{i=1}^{d} y_i z_i + at^2 = 1\}|.$$

On distingue alors deux possibilités :

— si $y_1 = \ldots = y_d = 0$ alors les z_i sont quelconques et $at^2 = 1$ ce qui donne $q^d\left[1 + \left(\frac{a}{q}\right)\right]$ choix

— si un y_i est non nul alors il nous faut choisir $(y_1, \ldots, y_d) \in \mathbb{F}_q^d \backslash \{0\}$, on a $q^d - 1$ choix pour cela, puis on choisit t (q choix), puis $(z_1, \ldots, z_d)$ qui est dans un hyperplan affine fixé, on a donc q^{d-1} choix. Au total, on a $q^d(q^d - 1)$ choix.

Et au total du total, on a $|X| = q^d\left(q^d + \left(\frac{a}{q}\right)\right)$.

Passons au deuxième dénombrement, qui repose sur de la théorie des groupes. On fait agir $\mathbb{Z}/p\mathbb{Z}$ sur X par rotation. On a alors deux sortes d'orbites ; celles pour lesquelles le stabilisateur est $\mathbb{Z}/p\mathbb{Z}$ tout entier, elles sont de la forme $(x, \ldots, x)$ avec $px^2 = 1$, et celles pour lesquelles le stabilisateur est trivial. L'équation aux classes donne alors :

$$|X| = \sum_{px^2=1} \frac{|\mathbb{Z}/p\mathbb{Z}|}{|\mathbb{Z}/p\mathbb{Z}|} + \sum_{\text{les autres}} \frac{|\mathbb{Z}/p\mathbb{Z}|}{|\{1\}|} \equiv 1 + \left(\frac{p}{q}\right) \mod p.$$

Par Euler, on a $q^d \equiv \left(\frac{q}{p}\right) \mod p$ et $\left(\frac{a}{q}\right) \equiv (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$. En combinant nos deux résultats et en remarquant que nos membres sont dans $\{-1, 1\}$, on obtient la loi de réciprocité quadratique. □

Définition. Pour $n = p_1 \cdots p_r$ un produit de nombres premiers pas forcément distincts, on

définit le *symbole de Jacobi*, noté comme le symbole de Legendre, par

$$\left(\frac{a}{p_1 \cdots p_r}\right) = \left(\frac{a}{p_1}\right) \cdots \left(\frac{a}{p_r}\right).$$

Proposition. Si $a \wedge n \neq 1$, alors $\left(\frac{a}{n}\right) = 0$, et sinon il vaut ± 1. De plus, pour a, b, m, n des entiers,

$$\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right) \text{ et } \left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$$

Démonstration. Cela découle directement de la définition et donc des propriétés multiplicatives du symbole de Legendre. □

Proposition. (loi de réciprocité quadratique, version Jacobi)
Pour m et n positifs impairs :

$$\left(\frac{m}{n}\right) = \left(\frac{n}{m}\right)(-1)^{\frac{(m-1)(n-1)}{4}}, \left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}, \left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}.$$

Proposition. Soit n un entier positif impair différent de 1. Alors n est premier si et seulement si

$$\forall a, a \wedge n = 1, \left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n.$$

Démonstration. Le sens direct découle du critère d'Euler. Pour le sens réciproque, on montre d'abord que n est sans facteur carré. On écrit pour cela $n = p^2 m$ avec p premier et on pose $a = 1 + pm$. On a alors la relation de Bezout

$$(1 - pm)a + mn = 1$$

qui montre que a et n sont premiers entre eux. Par hypothèse, on obtient donc que $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$.

Par ailleurs, le binôme de Newton montre que $a^p \equiv 1 \mod n$ donc l'ordre de a modulo n est égal à 1 ou p. Puisque $a^{n-1} \equiv 1 \mod n$, il doit diviser $n-1$ donc c'est 1. Donc n est sans facteur carré.

On écrit alors $n = p_1 \cdots p_r$ avec les (p_i) distincts. Si jamais $r \geq 2$, on choisit $\alpha_1, \ldots, \alpha_r$ tels que

- $\alpha_1 = 1$
- $\alpha_i \in [\![1, p_i - 1]\!]$
- $\left(\frac{\alpha_i}{p_i}\right) = 1$ si $1 \leq i \leq r-1$
- $\left(\frac{\alpha_r}{p_r}\right) = -1$.

D'après le théorème chinois, on considère un a unique dans $\{1, \ldots, n-1\}$ tel que $a \equiv \alpha_i \mod p_i$ pour tout i. Par définition il est premier avec n sinon un des symboles de Legendre ci-dessus serait nul. Par hypothèse, $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$. En particulier $\left(\frac{a}{n}\right) \equiv \alpha_1^{\frac{n-1}{2}} \mod p_1 \equiv 1 \mod p_1$. Mais d'autre part

$$\left(\frac{a}{n}\right) = \left(\frac{\alpha_1}{p_1}\right) \cdots \left(\frac{\alpha_r}{p_r}\right) = -1,$$

ce qui est contradictoire. □

Proposition. Si n est composé, il y a au plus $\frac{\varphi(n)}{2}$ entiers compris entre 1 et $(n-1)$ tels que

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$$

Démonstration. Par multiplicativité des symboles de Jacobi, l'ensemble de tels entiers est un sous-groupe. □

Proposition. (test de Solovay-Strassen)
Le test consiste à tester l'égalité précédente. Si n est composé, il y a moins d'une chance sur 2^k pour que n passe k fois le test.

Référence : Caldero, Germoni, tome 1 et *Cours d'algèbre* de Demazure.

Chapitre 75

Quaternions et matrices orthogonales

On veut ici généraliser à la dimension trois l'homéomorphisme $\mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{U}$. Cela va se faire par le biais des quaternions.

Théorème. (définition des quaternions)
Il existe une algèbre $\mathbb{H}$ de dimension 4 sur $\mathbb{R}$ telle que 1 soit le neutre pour la multiplication, et telle que l'on ait une base $(1, \mathbf{i}, \mathbf{j}, \mathbf{k})$ telle que

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1.$$

Remarque. Il y a deux manières usuelles de réaliser $\mathbb{H}$: d'une part dans $\mathcal{M}_2(\mathbb{C})$ avec

$$I = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, K = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},$$

et d'autre part dans $\mathcal{M}_4(\mathbb{R})$ en identifiant $\mathbb{C}$ à $\mathcal{M}_2(\mathbb{R})$ par $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Définition. Soit $q \in \mathbb{H}, q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, (a, b, c, d) \in \mathbb{R}^4$. On définit la conjugaison par

$$\overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$$

et la norme par $N(q) = q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Proposition. N est une forme quadratique euclidienne sur $\mathbb{H}$, de forme polaire :

$$\varphi(q_1, q_2) = \frac{1}{2}(q_1\overline{q_2} + q_2\overline{q_1}).$$

L'ensemble P est quaternions purs est orthogonal à $\mathbb{R}$ pour ce produit scalaire.

Théorème. L'ensemble $\mathbb{H}$ des quaternions forme un corps non commutatif. Le centre de $(\mathbb{H}^*, \times)$ est égal à $\mathbb{R}$, et N réalise un homomorphisme de $\mathbb{H}^*$ dans $\mathbb{R}_+^*$.

On note $G = \ker(N) = \{q \in \mathbb{H}, N(q) = 1\}$.

Théorème. On a l'isomorphisme suivant :

$$G/\{-1,1\} \simeq \mathrm{SO}_3(\mathbb{R}).$$

Démonstration. Comme le corps $\mathbb{H}$ n'est pas commutatif, l'action de $\mathbb{H}^*$ sur $\mathbb{H}$ est non triviale. On pose donc :

$$S_q : q' \mapsto qq'q^{-1}.$$

En fait, on ne perd rien en faisant opérer G puisqu'une éventuelle constante disparaitrait, et on a alors $q^{-1} = \overline{q}$. L'application S_q est clairement $\mathbb{R}$-linéaire et bijective, et $S : q \mapsto S_q$ est un morphisme de groupes, de noyau $\mathrm{Z}(\mathbb{H}) \cup G = \{-1,1\}$.

Ensuite, on voit que S_q est une isométrie pour N. On remarque alors que S_q restreinte à $\mathbb{R}$ est l'identité, et donc comme P est l'orthogonal de $\mathbb{R}$ pour N, S_q stabilise P et on peut poser s_q la restriction de S_q à P. Alors, $s : q \mapsto s_q$ est un morphisme de G sur $\mathrm{O}(N) \simeq O_3(\mathbb{R})$ ne noyau $\{-1,1\}$. L'application s est alors continue, puisque les coefficients de s_q dans la base $(\mathbf{i},\mathbf{j},\mathbf{k})$ sont des polynômes homogènes de degré 2 en les coefficients de q dans la base $(1,\mathbf{i},\mathbf{j},\mathbf{k})$. Donc $\det \circ s$ est continue de G dans $\{-1,1\}$. Par connexité de G (identifiable à la sphère dans $\mathbb{R}^4$), on peut définir $s : q \mapsto \mathrm{SO}_3(\mathbb{R})$.

Montrons enfin que s est surjective. Soit $p \in P \cup G$, alors $s_p(p) = p$, donc s_p fixe $\mathbb{R}.p$, c'est donc une rotation d'axe $\mathbb{R}.p$. Mais puisque p est dans $P \cup G$, $p^2 = -1$ donc $s_p^2 = \mathrm{id}$ donc c'est le renversement d'axe p. On a donc tous les renversements, qui engendrent $\mathrm{SO}_3(\mathbb{R})$, donc on a bien la surjectivité, puis le résultat voulu en passant au quotient. □

Corollaire. En plongeant $\mathbb{H}$ dans $\mathcal{M}_2(\mathbb{C})$, on a :

$$G \simeq \mathrm{SU}_2(\mathbb{C})$$

donc $\mathrm{PSU}_2(\mathbb{C}) \simeq \mathrm{SO}_3(\mathbb{R})$.

Référence : Daniel Perrin, *Cours d'algèbre.*

Chapitre 76

Théorème de Gauss-Wantzel

Dans ce développement, on cherche à quelle condition sur n le polygone régulier à n côtés est constructible à la règle et au compas.

Théorème. (de Gauss-Wantzel)
Le nombre $\mathrm{e}^{2\mathrm{i}\pi/n}$ est constructible à la règle et au compas si et seulement s'il s'écrit comme une puissance de deux fois un produit de nombres premiers de Fermat distincts, c'est-à-dire de la forme $2^{2^\alpha}+1$.

Démonstration. On peut se ramener à une puissance de nombre premier par Bezout. En effet, si a et b sont premiers entre eux, on peut écrire $au+bv=1$ une relation de Bezout, et alors

$$\mathrm{e}^{\frac{2\mathrm{i}\pi}{ab}} = \left(\mathrm{e}^{\frac{2\mathrm{i}\pi}{ab}}\right)^{au+bv} = \left(\mathrm{e}^{\frac{2\mathrm{i}\pi}{b}}\right)^{u}\left(\mathrm{e}^{\frac{2\mathrm{i}\pi}{a}}\right)^{v},$$

donc si les deux sont constructibles, alors le premier l'est. Réciproquement on a aussi le résultat puisqu'un produit de nombres constructibles est constructible.

Maintenant, on peut traiter de manière élémentaire le cas des puissances de 2, puisque l'on sait construire des bissectrices. Soit donc p premier impair et $\alpha\in\mathbb{N}^*$, on est ramenés au cas $n=p^\alpha$.

Si $\omega := \mathrm{e}^{\frac{2\mathrm{i}\pi}{p^\alpha}}$ est constructible, alors le degré de son polynôme minimal est une puissance de deux notée 2^m. Mais ce polynôme minimal n'est autre que Φ_{p^α}, qui est de degré $\varphi(p^\alpha)=p^{\alpha-1}(p-1)=2^m$. Par unicité de la décomposition en produit de facteurs premiers, α vaut 1 et p vaut 2^m+1. Montrons alors qu'un nombre premier de cette forme est nécessairement un premier de Fermat. Si m' est impair et tel que $m=2^\beta m'$, alors

$$p = 1+(2^{2^\beta})^{m'} = (1+2^{2^\beta})\sum_{k=0}^{m'-1}(-2^{2^\beta})^k,$$

donc $1+2^{2^\beta}$ donc égale p, ce qu'on voulait démontrer.

Réciproquement, soit p un nombre premier de Fermat, noté de la même manière $p=1+2^{2^\beta}$ avec $m=2^\beta$. On étudie les automorphismes du groupe $\mathrm{Aut}(K)=:G$ avec $K:=\mathbb{Q}(\omega)$ (c'est un corps car le polynôme minimal en question est un polynôme cyclotomique). On va montrer que

$G \simeq (\mathbb{Z}/p\mathbb{Z})^*$. Pour cela soit $\sigma \in G$, alors

$$0 = \sigma(0) = \sigma(\Phi_p(\omega)) = \Phi_p(\sigma(\omega)),$$

où l'on a utilisé le morphisme de Frobenius. Donc $\sigma(\omega)$ est une racine de Φ_p, donc un élément de la forme ω^k avec $k \in \{1, \dots, p-1\}$. Il existe donc k_σ tel que $\sigma(\omega) = \omega^{k_\sigma}$. On pose alors naturellement $\psi : G \to (\mathbb{Z}/p\mathbb{Z})^*$ qui à σ associe k_σ, qui est un morphisme de groupes (c'est une vérification). Elle est injective car si deux automorphismes coïncident sur $\mathbb{Q}$ et $\{\omega\}$ alors ils coïncident sur K et est surjective puisque cela revient à passer au quotient le morphisme d'évaluation en un ω^k (ils ont tous le même polynôme minimal). Ainsi G est cyclique d'ordre 2^m. Remarquons qu'alors pour σ_0 un générateur de G, $(\sigma_0^i(\omega))_{0 \leq i \leq p-2}$ est une base de K.

Puisque notre groupe est cyclique, on peut bâtir une suite de sous-groupes

$$G =: G_0 \supset \cdots \supset G_m = \{\mathrm{id}_K\},$$

avec G_i engendré par $\sigma_0^{2^i}$. À partir de cette suite de sous-groupes, on construit les sous-corps de K suivants :

$$K_i := \{z \in K, \forall \sigma \in G_i, \sigma(z) = z\} = \{z \in K, \sigma_0^{2^i}(z) = z\},$$

où la deuxième égalité est due au fait que l'on a pris un générateur de G_i. Si $z \in K_0$, on a, en utilisant le fait que $\sigma_0(z) = z$ et en décomposant selon la base que l'on a exhibée, $z \in \mathbb{Q}$ donc $K_0 = \mathbb{Q}$. On vérifie alors que

$$z := \sum_{l=0}^{2^{m-i-1}-1} \sigma_0^{2^{i+1}l}(\omega) \in K_{i+1} \backslash K_i.$$

Finalement : $2^m = p - 1 = [\mathbb{Q}(\omega) : \mathbb{Q}] = \prod_{i=1}^m [K_i : K_{i-1}]$. Chaque terme est plus grand que 2 est le produit est une puissance de deux donc chaque degré vaut 2, on peut appliquer le théorème de Wantzel. □

Remarque. En général, le théorème de Wantzel est toujours énoncé pour des *réels* constructibles. Voyons donc comment le généraliser pour des complexes constructibles, et donc pouvoir éviter des vérifications pénibles.

On part de la remarque suivante : le point $M(x, y)$ est constructible si et seulement si les points $A(x, 0)$ et $B(0, y)$ sont constructibles.

Définition. On définit E(i) l'ensemble des complexes constructibles (où E est l'ensemble des réels constructibles).

Lemme. C'est un sous-corps de $\mathbb{C}$ stable par racine carrée.

Démonstration. Cela se montre comme sur $\mathbb{R}$, en remarquant que les parties réelle, imaginaire d'une racine carrée de z s'expriment comme des sommes, produits, racines carrées réelles des parties réelle, imaginaire de z. □

Théorème. (de Wantzel, version complexe)
Soit $z \in \mathbb{C}$. Alors $z \in \mathrm{E(i)}$ si et seulement s'il existe une tour d'extensions quadratiques (de sous-corps de $\mathbb{C}$) $(L_0, \dots, L_p)$ avec $\mathbb{Q} = L_0$ et $z \in L_p$.

Démonstration. Pour le sens direct, la preuve de Wantzel réel donne une tour d'extensions réelles $(K_0, \dots, K_q)$ avec $K_0 = \mathbb{Q}$ et $\mathrm{Re}(z)$ et $\mathrm{Im}(z)$ qui sont dans K_q. On pose alors $K_{q+1} = L_p(\mathrm{i})$, qui est bien une extension de L_p de dimension 2 ($\pi_{\mathrm{i},Lp} = X^2 + 1$). Ainsi, $(K_0, \dots, K_{q+1})$ donne la tour voulue.

Réciproquement, par récurrence, si $K_j \subset \mathrm{E(i)}$, considérons $x \in K_{j+1} \backslash K_j$. Puisque $[K_{j+1} : K_j] = 2$, il existe b, c dans K_j tels que $x^2 + bx + c = 0$. Soit alors δ une racine complexe de $\Delta = b^2 - 4c$, on a $\frac{-b \pm \delta}{2}$ mais $\Delta \in K_j \subset \mathrm{E(i)}$ qui est stable par racine carrée donc $\delta \in \mathrm{E(i)}$ et x également. □

Référence : le livre de Carréga sur les constructions à la règle et au compas. On retrouve la preuve dans le document de Matthias Hostein également.

Chapitre 77

Sous-groupes compacts de matrices réelles

Dans ce développement, on donne une condition nécessaire et suffisante pour qu'un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$ soit compact.

Définition. Soit G un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. On dit que $B \in \mathcal{S}_n^{++}(\mathbb{R})$ est invariant pour G (en tant que produit scalaire) si $G \subset \mathrm{O}(B)$ *i.e.* si pour tout $M \in G, M^\top BM = B$.

Proposition. Un sous-groupe G de $\mathrm{GL}_n(\mathbb{R})$ possède un produit scalaire invariant si et seulement si G est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Démonstration. Soit B symétrique définie positive invariante pour G, cela veut dire que $G \subset \mathrm{Stab}_B$ (pour l'action de congruence). La matrice B ayant une racine carrée symétrique, elle est dans l'orbite de l'identité, donc par principe de conjugaison G est conjugué au stabilisateur de l'identité, qui n'est autre que $\mathrm{O}_n(\mathbb{R})$. Réciproquement, si $G = P^{-1}HP$ pour P inversible et H un sous-groupe de $\mathrm{O}_n(\mathbb{R})$, on écrit $P = OR$ la décomposition polaire de P et $B = R^2$ donne la produit scalaire voulu. □

Théorème. (point fixe de Kakutani)
Soit E un espace euclidien et G un sous-groupe compact de $\mathrm{GL}(E)$ et K un convexe compact non vide de E tel que $g(K) \subset K$ pour tout g dans G. Alors il existe $z \in K$ tel que $g(z) = z$ pour tout $g \in G$.

Démonstration. Soit $N : x \mapsto \max_{g \in G} \|g(x)\|$, qui est bien définie car G est compact. C'est une norme ; la séparation et l'homogénéité sont directes, et l'inégalité triangulaire se voit en prenant un g_0 qui atteint le maximum. De plus elle est invariante pour G. Pour avoir égalité dans l'inégalité triangulaire, il faut et il suffit que $\|g_0(x+y)\| = \|g_0(x)\| + \|g_0(y)\|$, ce qui est vrai si et seulement si $g_0(x)$ et $g_0(y)$ sont positivement liés, si et seulement si x et y le sont car g_0 est un isomorphisme. La norme N est continue. On peut donc considérer par compacité un élément de K de norme N minimale, on l'appelle z et sa norme m. Montrons que z est unique. Pour cela si y vérifie la même chose, on peut considérer $a := \frac{y+z}{2} \in K$ (par convexité), et on a alors

$$m \leq N(a) = \frac{1}{2}N(y+z) \leq m$$

donc $N(y+z) = N(y)+N(z)$ donc y et z sont positivement liés. Puisqu'ils sont de même norme ils sont égaux. Soit $g \in G$; on a $N(g(z)) = N(z) = m$ et par unicité : $g(z) = z$. □

Théorème. Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Démonstration. Soit H un sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ soit $\rho : M \mapsto (S \mapsto M^\top SM)$. C'est un morphisme de groupe du groupe opposé de H vers $\mathrm{GL}(\mathcal{S}_n(\mathbb{R}))$ qui est continue car polynomiale en les coefficients de M donc $G = \rho(H)$ est un sous-groupe compact de $\mathrm{GL}(\mathcal{S}_n(\mathbb{R}))$. Soit alors $K := \mathrm{Conv}(\{M^\top M, M \in H\})$ qui est convexe compact par Carathéodory (donc inclus dans $\mathcal{S}_n^{++}(\mathbb{R})$). On vérifie que K est stable par G. Le point fixe de Kakutani et la première proposition permettent de conclure. □

Référence : *Thèmes de géométrie*, Alessandri. On le trouve aussi sur la page de Matthias Hostein.

Chapitre 78

Forme normale de Smith

On montre dans ce développement le théorème de Smith, habituellement rencontré dans des anneaux euclidiens, mais qui peut être généralisé sans grand changement de la philosophie dans un cadre principal. Ce théorème est fondamental en théorie des modules, et permet donc de retrouver des résultats sur les groupes abéliens finis et les endomorphismes d'un espace de dimension finie.

Théorème. (forme normale de Smith) Soit $M \in \mathcal{M}_{n,p}(A)$ où A est un anneau principal. Il existe $(P,Q) \in \mathrm{GL}_n(A) \times \mathrm{GL}_p(A)$ telles que

$$PMQ = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Avec $d_1 | \dots | d_r$. De plus, les (d_i) sont uniques à association près. Cette forme diagonale est appelée *forme (normale) de Smith* de la matrice M, et les (d_i) sont appelés *facteurs invariants.*

Démonstration. On va procéder à une preuve algorithmique. On suppose que l'on a une boîte noire qui permet de calculer les relations de Bezout. L'idée est de procéder récursivement. On cherche à obtenir, avec des opérations élémentaires sur les lignes et des colonnes (*i.e.* des multiplications par des matrices inversibles simples à gauche et à droite) une matrice de la forme

$$M' = \begin{pmatrix} d_1 & 0 \\ 0 & N \end{pmatrix},$$

avec $N \in \mathcal{M}_{n-1,p-1}$ dont les coefficients sont tous divisibles par d_1 puis de recommencer cette étape sur la matrice N.

Lançons-nous donc dans l'algorithme qui va nous permettre d'arriver à la forme ci-dessus. On note $M = (a_{i,j})$. Notre but est de remplacer successivement le coefficient $a_{1,1}$ par un diviseur strict, de telle sorte à ensuite pouvoir annuler les coefficients sur les lignes et les colonnes avec de simples transvections, comme on le ferait sur un corps. Pour faire cela, on utilise ce qu'on a dans

un anneau principal : la relation de Bezout. À chaque étape, on parcourt les cinq choix suivants et on effectue le premier qu'il est possible de faire.

— S'il existe $j > 1$ tel que $a_{1,1}$ ne divise pas $a_{1,j}$, on calcule la relation de Bezout $ua_{1,1}+va_{1,j} = \delta$ avec δ un pgcd de $a_{1,1}$ et $a_{1,j}$. On cherche alors une matrice élémentaire qui va nous permettre de remplacer $a_{1,1}$ par δ. On multiplie alors à droite par la matrice

$$\widetilde{Q} = \begin{pmatrix} u & \dots & -\frac{a_{1,j}}{\delta} & \dots \\ \vdots & I_{j-2} & \vdots & \\ v & \dots & \frac{a_{1,1}}{\delta} & \dots \\ \vdots & & \vdots & I_{p-j} \end{pmatrix},$$

où les coefficients écrits sont sur les lignes et colonnes d'indices 1 et j, et avec des 1 sur le reste de la diagonale. Cette matrices est bien inversible puisque son déterminant peut être calculé avec la formule de base et vaut

$$\det(\widetilde{Q}) = \frac{1}{\delta}(ua_{1,1} + va_{1,j}) = 1 \in A^\times.$$

La multiplication à droite par cette matrice correspond aux opérations élémentaires suivantes :

$$\begin{aligned} C_1 &\leftarrow uC_1 + vC_j \\ C_j &\leftarrow -\frac{a_{1,j}}{\delta}C_1 + \frac{a_{1,1}}{\delta}C_j. \end{aligned}$$

On a donc bien remplacé $a_{1,1}$ par un diviseur strict. Remarquons que le coefficient sur la colonne j a été remplacé par 0.

— S'il existe $i > 1$ tel que $a_{1,1}$ ne divise pas $a_{i,1}$, on multiplie à gauche par

$$\widetilde{P} = \begin{pmatrix} u & \dots & v & \dots \\ \vdots & I_{i-2} & \vdots & \\ -\frac{a_{i,1}}{\delta} & \dots & -\frac{a_{1,1}}{\delta} & \dots \\ \vdots & & \vdots & I_{n-i} \end{pmatrix},$$

et on remplace de manière similaire $a_{1,1}$ par un diviseur strict et $a_{i,1}$ par zéro.

— S'il existe un coefficient $a_{1,j}$ non nul tel que $a_{1,1}$ divise $a_{1,j}$ on peut l'annuler avec la transvection

$$C_j \leftarrow C_j - \frac{a_{1,j}}{a_{1,1}}C_1.$$

— De même sur les lignes, s'il existe un coefficient $a_{i,1}$ non nul tel que $a_{1,1}$ divise $a_{i,1}$, on l'annule aussi avec une transvection :

$$L_i \leftarrow L_i - \frac{a_{i,1}}{a_{1,1}}L_1.$$

— Enfin s'il existe (i, j) avec $i > 1, j > 1$ tel que $a_{1,1}$ ne divise pas $a_{i,j}$, on fait apparaître $a_{i,j}$

en position $(i,1)$ par la transvection

$$C_1 \leftarrow C_1 + C_j.$$

Montrons maintenant la terminaison de notre algorithme. Commençons par remarquer que l'on est sûrs d'arriver au cinquième point de la liste précédente, puisque les quatre étapes précédentes remplacent $a_{1,1}$ par un diviseur strict (et il n'en a qu'un nombre fini car on est dans un anneau principal donc factoriel ou ajoutent un zéro en ligne ou colonne 1). Mais lorsque l'on est au cinquième point, on remplace au moins un zéro de C_1 qui n'est pas divisible par $a_{1,1}$, donc $a_{1,1}$ va encore décroître strictement. Somme toute, la suite des valeurs successives de $a_{1,1}$ ne pouvant pas décroître strictement une infinité de fois, elle stationne. Lorsqu'elle stationne, on a bien que des zéros ailleurs sur la ligne et la colonne 1. On peut donc itérer de procédé comme indiqué précédemment et on obtient bien notre forme de Smith.

Réglons maintenant le problème d'unicité de la forme normale de Smith. On note $\mathcal{I}_k(B)$ l'idéal engendré par les mineurs de taille k de la matrice B. Si U est une matrice de taille convenable, alors $\mathcal{I}_k(UB) \subset \mathcal{I}_k(B)$ d'après la formule de Cauchy-Binet, et de même pour la multiplication à droite. Ainsi, si M et $M' = UMV$ sont deux matrices équivalentes, $\mathcal{I}_k(M') \subset \mathcal{I}_k(M)$. Par symétrie on a égalité. Ainsi, si on a deux jeux de facteurs invariants $d_1|\dots|d_r$ et $d'_1|\dots|d'_s$, alors d'abord $r = s$ car si $s > r$ on aurait $\langle 0\rangle = \langle d'_1 \cdots d'_{r+1}\rangle$ donc un des d'_j serait nul par intégrité ce qui est exclu. Ensuite, on a donc

$$\begin{aligned}\langle d_1\rangle &= \langle d'_1\rangle\\ \langle d_1 d_2\rangle &= \langle d'_1 d'_2\rangle\\ &\dots\\ \langle d_1\cdots d_r\rangle &= \langle d'_1\cdots d'_r\rangle.\end{aligned}$$

De proche en proche, on obtient bien l'association des facteurs invariants. □

Proposition. (formule de Cauchy-Binet)
Soit R un anneau principal et $A \in \mathcal{M}_{m,n}(R)$ et $B \in \mathcal{M}_{n,q}(R)$. On note $C = AB \in \mathcal{M}_{m,q}(R)$. Alors pour tous $\{i_1,\dots,i_p\} \subset [\![1,m]\!]$ et $\{j_1,\dots,j_p\} \subset [\![1,q]\!]$:

$$|C|_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_p}} = \sum_{1\le k_1<\dots<k_p\le n} |A|_{\substack{i_1,\dots,i_p\\ k_1,\dots,k_p}} |B|_{\substack{k_1,\dots,k_p\\ j_1,\dots,j_p}}.$$

Démonstration. On remarque que

$$(C)_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_p}} = (A)_{\substack{i_1,\dots,i_p\\ 1,\dots,n}} (B)_{\substack{1,\dots,n\\ j_1,\dots,j_p}}.$$

On se ramène donc à un déterminant carré d'un produit de deux matrices rectangulaires. On

change alors les notations, soient

$$A = (a_{i,j}) \in \mathcal{M}_{p,n}(R)$$
$$B = (b_{i,j}) \in \mathcal{M}_{n,p}(R)$$
$$C = AB \in \mathcal{M}_p(R).$$

On cherche à calculer $\det(C)$. On note $A = (A_1|\dots|A_n)$ les colonnes de A.

$$\begin{aligned}
\det(C) &= \det\left(\sum_{k_1=1}^{n} b_{k_1,1}A_{k_1}, \dots, \sum_{k_p=1}^{n} b_{k_p,p}A_{k_p}\right) \\
&= \sum_{k_1,\dots,k_p=1}^{n} b_{k_1,1}\cdots b_{k_p,p}|A|_{\substack{1,\dots,p\\k_1,\dots,k_p}} \\
&= \sum_{1\le k_1<\dots<k_p\le n}\ \sum_{\{l_1,\dots,l_p\}=\{k_1,\dots,k_p\}} b_{l_1,1}\cdots b_{l_p,p}|A|_{\substack{1,\dots,p\\l_1,\dots,l_p}} \\
&= \sum_{1\le k_1<\dots<k_p\le n}\ \sum_{\sigma\in\mathfrak{S}_{\{k_1,\dots,k_p\}}} b_{\sigma(k_1),1}\cdots b_{\sigma(k_p),p}|A|_{\substack{1,\dots,p\\\sigma(k_1),\dots\sigma(k_p)}} \\
&= \sum_{1\le k_1<\dots<k_p\le n}\ \sum_{\sigma\in\mathfrak{S}_{\{k_1,\dots,k_p\}}} b_{\sigma(k_1),1}\cdots b_{\sigma(k_p),p}\varepsilon(\sigma)|A|_{\substack{1,\dots,p\\k_1,\dots,k_p}} \\
&= \sum_{1\le k_1<\dots<k_p\le n} |A|_{\substack{1,\dots,p\\k_1,\dots,k_p}} \sum_{\sigma\in\mathfrak{S}_{\{k_1,\dots,k_p\}}} \varepsilon(\sigma)b_{\sigma(k_1),1}\cdots b_{\sigma(k_p),p} \\
&= \sum_{1\le k_1<\dots<k_p\le n} |A|_{\substack{1,\dots,p\\k_1,\dots,k_p}}\,|B|_{\substack{k_1,\dots,k_p\\1,\dots,p}}.
\end{aligned}$$

On utilise d'abord la multilinéarité du déterminant, avant de regrouper les indices selon les valeurs prises, puis à réindexer selon les permutations et à utiliser le caractère alterné du déterminant. □

Remarque. (Quelques mots sur la complexité de l'algorithme)
On note considère linéaire en k le coût d'une opération élémentaire sur une ligne ou une colonne de taille k (disons inférieur à αk). On note b le coût d'une opération de Bezout (ou d'un test de divisibilité). Dans le pire des cas, pour passer la première étape de la récurrence, on fait $(n-1)+(p-1)$ opérations élémentaires et relations de Bezout puis lorsque l'on arrive au cinquième point de l'algorithme $a_{1,1}$ ne divise aucune des éléments de la sous-matrice en bas à droite, et on est alors obligés de faire revenir une à une les colonnes de cette sous-matrice en C_1, ce qui rajoute au plus $(n-1)(p-1)$ calculs de relations de Bezout et opérations élémentaires. Ainsi, on a la borne suivante (si $C(n,p)$ est le nombre d'opérations pour une matrice de taille $n\times p$) :

$$C(n,p) \le (\alpha\max(n,p)+b)(n+p-2+(n-1)(p-1))+C(n-1,p-1) \le A(\alpha N+b)N^2+C(n-1,p-1)$$

On a somme toute une complexité en $\mathrm{O}((\alpha N+b)N^3)$.

Remarque. Dans un anneau euclidien, la relation de Bezout s'obtient par l'algorithme d'Euclide. D'ailleurs, on peut effectuer directement l'algorithme d'Euclide dans la matrice pour économiser de la mémoire dans l'ordinateur.

Application. (Réduction de Frobenius d'un endomorphisme)
Soit $u \in \mathcal{L}(E)$, où E est un K-espace vectoriel de dimension n. En appliquant l'algorithme des facteurs invariants à la matrice $XI_n - M$, où M est la matrice de u dans une base quelconque, on obtient des polynômes $P_1 | \cdots | P_r$ tels que $XI_n - M$ soit équivalente dans $K[X]$ à

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & P_1(X) & & & \\ & & & & \ddots & & \\ & & & & & P_r(X) \end{pmatrix}.$$

En effet, il ne peut pas y avoir de zéros puisque le déterminant de notre matrice est $\chi_M \neq 0$.

On considère alors une "action" de $K[X]$ par $P \cdot x := P(u)(x)$. Elle est compatible avec la structure d'anneau de $K[X]$. On peut alors définir un morphisme compatible avec cette action de $K[X]^n$ dans E par $(Q_1, \ldots, Q_n) \mapsto \sum_{i=1}^n Q_i \cdot e_i$ où $(e_1, \ldots, e_n)$ est une base de E. Ce morphisme est clairement surjectif. De plus, l'image de $XI_n - M$ est incluse dans son noyau. On peut alors passer au quotient, comme on le ferait sur un espace vectoriel, en préservant la structure définie par $\cdot$. On note $W := \mathrm{Im}(XI_n - M)$, et on pose $\overline{\varphi} : K[X]^n/W \to E$ qui est surjectif. C'est aussi une application K-linéaire. Par récurrence sur le degré de polynômes (U_i) fixés, on montre qu'il existe des vecteurs (x_i) de E tels que $(U_1, \ldots, U_n) = (x_1, \ldots, x_n) \bmod W$, ce qui donne l'injectivité de $\overline{\varphi}$ par égalité des dimensions. On a donc un isomorphisme entre E et $K[X]^n/W$. Remarquons alors que si deux matrices A et A' sont équivalentes sur un anneau R, alors $R^n/\mathrm{Im}(A) \simeq R^n/\mathrm{Im}(A')$. Ainsi avec notre forme normale de Smith on obtient un isomorphisme

$$\psi : E \to K[X]/(P_1) \times \cdots \times K[X]/(P_r).$$

Soit alors $x_i = \psi^{-1}((0, \ldots, \overline{1}, \ldots, 0))$ où le 1 est en position i. Si on note $\deg(P_i) =: d_i$, alors

$$\mathcal{B} := (x_1, u(x_1), \ldots, u^{d_1-1}(x_1), x_2, \ldots, x_r, u(x_r), \ldots, u^{d_r-1}(x_r))$$

est une K-base de E, et la matrice de U dans cette base est diagonale par blocs, les blocs étant les matrices compagnon des polynômes $P_1, \ldots, P_r$. On retrouve donc le théorème de réduction de Frobenius, avec en plus une méthode algorithmique pour calculer les facteurs invariants.

Application. (Systèmes linéaires dans $\mathbb{Z}$)
La forme normale de Smith permet de résoudre des systèmes linéaires à coefficients dans $\mathbb{Z}$. En effet, si on veut résoudre $Ax = b$ dans $\mathbb{Z}^n$, on peut écrire $A = PDQ$ sa forme de Smith et alors on réécrit le problème en $D(Qx) = P^{-1}b$ et cela devient alors plus simple. Remarquons que l'on pourrait seulement échelonner selon les colonnes, on parle alors de forme normale de Hermite.

Référence : La preuve de la forme normale de Smith provient de la page Wikipédia *Théorème des facteurs invariants* et d'un cours de Christian Blanchet accessible ici. Le lien avec la réduction de Frobenius provient du cours de Vincent Guirardel.

Chapitre 79

Caractères et marches aléatoires dans les groupes abéliens

Dans ce développement, on s'intéresse à la transformée de Fourier discrète qui, comme dans le cas continu, permet de simplifier des calculs de lois de sommes de variables aléatoires indépendantes. Pour introduire la transformée de Fourier sur un groupe, on doit introduire ses caractères et son groupe dual (de la même manière que la transformée de Fourier d'une fonction L^1 définit un élément de son dual).

Dans toute la suite G désignera un groupe abélien fini de cardinal supérieur à 2.

Définition. Un *caractère* de G est un morphisme de G vers $(\mathbb{C}^*, \times)$. Le groupe formé des caractères de G est un groupe pour la multiplication est noté $\widehat{G}$ est appelé *groupe dual* de G. Puisque G est fini, les caractères sont en fait à valeurs dans les racines de l'unité, et en particulier sont donc de module 1.

Exemple. Les caractères de $\mathbb{Z}/N\mathbb{Z}$ sont les $\overline{x} \mapsto \mathrm{e}^{\frac{2\mathrm{i}kx\pi}{N}}$ pour $\overline{k} \in \mathbb{Z}/N\mathbb{Z}$. On remarque que $\widehat{\mathbb{Z}/N\mathbb{Z}} \simeq \mathbb{Z}/N\mathbb{Z}$.

Remarque. Tout groupe abélien fini est isomorphe à son dual et canoniquement isomorphe à son bidual par $g \mapsto (\chi \mapsto \chi(g))$ par le théorème de structure des groupes abéliens finis.

Définition. Soit H un groupe abélien fini, on peut définir un produit scalaire sur $\mathbb{C}^H$ par la formule :

$$\langle f_1, f_2 \rangle_H = \frac{1}{|H|} \sum_{h \in H} \overline{f_1(h)} f_2(h).$$

Définition. Soit $f \in \mathbb{C}^G$. On définit $\widehat{f} \in \mathbb{C}^{\widehat{G}}$ par la formule :

$$\widehat{f}(\chi) := \sum_{g \in G} \overline{\chi(g)} f(g) = |G| \langle \chi, f \rangle_G.$$

Proposition. Les caractères forment une base orthonormée de $\mathbb{C}^G$.

Démonstration. Soient χ et ψ deux caractères de G. Alors remarquons que

$$\langle \chi, \psi \rangle_G = \langle \chi\overline{\psi}, 1 \rangle_G = \langle \chi\psi^{-1}, 1 \rangle_G.$$

Ainsi, pour montrer que la base est orthonormée, il suffit de montrer que $\langle \varphi, 1\rangle_G = \delta_{\varphi,1}$ pour tout caractère φ. Soit donc un tel caractère, et $h \in H$ différent de 1.

$$\langle \varphi, 1\rangle_G = \frac{1}{|G|}\sum_{g\in G}\varphi(g) = \frac{1}{|G|}\sum_{g\in G}\varphi(gh) = \varphi(h)\frac{1}{|G|}\sum_{g\in G}\varphi(g),$$

Si $\varphi(h)$ est différent de 1 pour un h, *i.e.* si $\varphi \neq 1$, on obtient $\frac{1}{|G|}\sum_{g\in G}\chi(g) = 0$ d'où le résultat. □

Proposition. (Égalité de Bessel-Parseval)
Pour toute fonction $f \in \mathbb{C}^G$, puisque la famille $\widehat{G}$ est une base orthonormée, on a

$$\sum_{g\in G}|f(g)|^2 = |G|\sum_{g\in G}|\langle \chi, f\rangle|^2 = \frac{1}{|G|}\sum_{\chi\in\widehat{G}}|\widehat{f}(\chi)|^2.$$

Proposition. (Inversion de Fourier)
Soit $f \in \mathbb{C}^G$. Alors pour tout $g \in G$:

$$f(g) = \frac{1}{|G|}\sum_{\chi\in\widehat{G}}\chi(g)\widehat{f}(\chi)$$

Démonstration. Il suffit de décomposer f dans la base orthonormée $\widehat{G}$:

$$f = \sum_{\chi\in\widehat{G}}\langle \chi, f\rangle_G\,\chi = \frac{1}{|G|}\sum_{\chi\in\widehat{G}}\widehat{f}(\chi)\chi.$$

□

Remarque. On aurait pu aussi directement écrire :

$$\frac{1}{|G|}\sum_{\chi\in\widehat{G}}\chi(g)\widehat{f}(\chi) = \frac{1}{|G|}\sum_{\chi\in\widehat{G}}\sum_{h\in G}\chi(g)\overline{\chi(h)}f(h) = \frac{1}{|G|}\sum_{h\in H}\left(f(h)\sum_{\chi\in\widehat{G}}\overline{\chi(h)}\chi(g)\right).$$

Ainsi, par isomorphisme canonique avec le bidual, on a $\sum_{\chi\in\widehat{G}}\overline{\chi(h)}\chi(g) = |G|\delta_{g,h}$, d'où le résultat. Toujours avec cet isomorphisme canonique avec le bidual, on peut écrire $f(g) = \langle g^{-1}, \widehat{f}\rangle_{\widehat{G}} = \frac{1}{|\widehat{G}|}\widehat{\widehat{f}}(g^{-1}) = \frac{1}{|G|}\widehat{\widehat{f}}(g^{-1})$.

Définition. On peut définir un produit de convolution (associatif, commutatif et distributif par rapport à +) sur $\mathbb{C}^H$ pour H un groupe abélien fini par la formule suivante :

$$(f_1 * f_2)(h) := \sum_{k\in G} f_1(k)f_2(hk^{-1}).$$

Proposition. Soient f_1 et f_2 deux fonctions de $\mathbb{C}^G$. Alors $\widehat{f_1 * f_2} = \widehat{f_1}\widehat{f_2}$.

Démonstration. Soit $\chi \in \widehat{G}$:

$$\begin{aligned}\widehat{f_1 * f_2}(\chi) &= \sum_{g\in G} \overline{\chi(g)}(f_1 * f_2)(g) \\ &= \sum_{g\in G}\sum_{h\in G} \overline{\chi(g)} f_1(h) f_2(gh^{-1}) \\ &= \sum_{g\in G}\sum_{h\in G} \overline{\chi(gh^{-1})\chi(h)}\, f_1(h) f_2(gh^{-1}) \\ &= \sum_{g'\in G} \overline{\chi(g')} f_2(g') \sum_{h\in G} \overline{\chi(h)} f_1(h) \\ &= \widehat{f}_1(\chi)\widehat{f}_2(\chi).\end{aligned}$$

où l'on a posé $g' = gh^{-1}$ qui réalise une bijection sur G. □

On a maintenant tout ce qu'il nous faut pour faire des probabilités. On considère une particule sur $\mathbb{Z}/N\mathbb{Z} = (\overline{0}, \ldots, \overline{N-1})$, initialement en $\overline{0}$. À chaque étape, la particule se déplace d'un cran vers la droite ou vers la gauche ou stationne de manière équiprobable. On peut modéliser cela par une chaîne de Markov dont la loi initiale est $\delta_{\overline{0}}$ et dont la loi de transition est $\mu = \frac{1}{3}(\delta_{-\overline{1}} + \delta_{\overline{0}} + \delta_{\overline{1}})$. La loi de X_n la position de la particule au rang n a pour loi

$$\mu_n = \mu_0 * \mu^{*n} = \mu^{*n}.$$

Théorème. La chaîne de Markov en question converge vers la loi uniforme notée μ_∞ et on a une approximation de la vitesse de convergence en variation totale :

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N}\mathrm{e}^{-\frac{\pi^2 n}{N^2}}.$$

Démonstration. Tout l'intérêt de la transformée de Fourier discrète est de pouvoir passer cette équation dans le monde de Fourier :

$$\widehat{\mu_n} = \widehat{\mu}^n.$$

Calculons alors $\widehat{\mu}$. On utilise pour cela l'isomorphisme entre $\mathbb{Z}/N\mathbb{Z}$ et son dual :

$$\widehat{\mu}(\overline{k}) = \sum_{l=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}kl\pi}{N}} \mu(\overline{l}) = \frac{1}{3}\left(1 + \mathrm{e}^{-\frac{2\mathrm{i}k\pi}{N}} + \mathrm{e}^{\frac{2\mathrm{i}k\pi}{N}}\right) = \frac{1}{3}\left(1 + 2\cos\left(\frac{2k\pi}{N}\right)\right).$$

Ainsi, par inversion de Fourier, pour $\overline{l} \in \mathbb{Z}/n\mathbb{Z}$:

$$\mu_n(\overline{l}) = \frac{1}{N}\sum_{k=0}^{N-1} \mathrm{e}^{\frac{2\mathrm{i}k\pi}{N}} \left(\frac{1}{3}\left(1 + 2\cos\left(\frac{2k\pi}{N}\right)\right)\right)^n.$$

Or $\frac{1}{3}|1 + 2\cos(\frac{2k\pi}{n})| = 1$ si et seulement si $k = 0$ donc quand n tend vers l'infini : $\mu_n(\overline{l}) \sim \frac{1}{N}$. Ainsi la chaîne de Markov converge en loi vers la mesure uniforme $\mu_\infty = \frac{1}{N}\sum_{k=0}^{N-1} \delta_{\overline{k}}$ dont la

transformée de Fourier est $\delta_{\overline{0}}$. Par Parseval, on a donc

$$\|\mu_n - \mu_\infty\|_2^2 = \frac{1}{N}\|\widehat{\mu_n} - \widehat{\mu_\infty}\|_2^2 = \sum_{k=0}^{n-1} \left| \left(\frac{1}{3}\left(1 + 2\cos\left(\frac{2k\pi}{N}\right)\right)\right)^n - \delta_0(k) \right|^2.$$

Les termes en $k = 0$ se simplifient et on a donc :

$$\|\mu_n - \mu_\infty\|_2^2 = \sum_{k=1}^{N-1} \left(\frac{1}{3}\left(1 + 2\cos\left(\frac{2k\pi}{N}\right)\right)\right)^{2n}.$$

Le terme est maximal pour $k = 1$ donc

$$\|\mu_n - \mu_\infty\|_2 \leq \left(\frac{1 + 2\cos(\frac{2\pi}{N})}{3} \right)^n.$$

Par Cauchy-Schwarz, on a alors

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N} \left(\frac{1 + 2\cos(\frac{2\pi}{N})}{3} \right)^n.$$

Pour $x \in]0, \frac{\pi}{2}[$, on peut montrer que $\frac{1+2\cos(x)}{3} \leq 1 - \frac{x^2}{4} \leq \mathrm{e}^{-\frac{x^2}{4}}$. Ainsi, pour $N > 4$, on a la majoration suivante :

$$\|\mu_n - \mu_\infty\|_{\mathrm{VT}} \leq \sqrt{N}\mathrm{e}^{-\frac{\pi^2 n}{N^2}}.$$

□

Référence. Cours de chaînes de Markov de Pierre-Loïc Méliot, accessible ici.

Chapitre 80

Matrices inversibles à coefficients entiers

Dans ce développement, on s'intéresse aux sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$, et on montre que leur cardinal peut être borné indépendamment du groupe.

Proposition. Soit $M \in \mathrm{GL}_n(\mathbb{Z})$ d'ordre fini d. Alors M est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité.

Démonstration. Par définition, le polynôme $X^d - 1$ annule M et ce polynôme est scindé à racines simples. Ainsi M est diagonalisable et ses racines figurent parmi les racines de $X^d - 1$, *i.e.* parmi les racines d-ièmes de l'unité. □

Théorème. Soit M une matrice de $\mathrm{GL}_n(\mathbb{Z})$ d'ordre fini. L'ensemble des ordres possibles pour M est fini.

Démonstration. Avec la proposition précédente, on sait que le polynôme caractéristique de M n'a que des racines dans $\mathbb{U}_d$ où d est l'ordre de M. À l'aide des relations coefficients racines, on obtient, si $\chi_M = X^n + \sum_{k=0}^{n-1} a_k X^k = (X - z_1) \cdots (X - z_n)$:

$$a_{n-k} = (-1)^k \sum_{1 \leq i_1 < \ldots < i_k \leq n} z_{i_1} \cdots z_{i_k},$$

et en prenant le module :

$$|a_{n-k}| \leq \binom{n}{k}.$$

Ainsi, puisque χ_M est un polynôme de $\mathbb{Z}[X]$ à coefficients bornés, il n'y a qu'un nombre fini de valeurs possibles pour chaque coefficient, donc un nombre fini de polynômes caractéristiques possibles, indépendamment de l'ordre !

Or l'ordre de la matrice M correspond au ppcm des ordres (dans $(\mathbb{C}^*, \times)$) de ses racines. Comme M ne peut avoir qu'un nombre fini de polynômes caractéristiques possibles, il ne peut y avoir qu'un nombre fini de tels ppcm, donc les ordre possibles pour M sont bien finis. □

Théorème. Soit G un sous-groupe fini de $\mathrm{GL}_n(\mathbb{Z})$. Alors son cardinal est borné par 3^{n^2}.

Démonstration. On veut injecter, en réduisant les coefficients, G dans $\mathcal{M}_\setminus(\mathbb{Z}/m\mathbb{Z})$ où m est un entier plus grand que 3. Soient g et h deux éléments de G. On suppose que $g-h$ vaut 0 modulo m. Ainsi on peut écrire $g-h=mB$ avec $B\in\mathcal{M}_n(\mathbb{Z})$. On a alors

$$A := h^{-1}B = (h^{-1}g - I_n)/m \in \mathcal{M}_n(\mathbb{Z}).$$

Puisque $h^{-1}g$ est un élément de G, il est diagonalisable et ses valeurs propres sont des racines de l'unité. Ainsi A est également diagonalisable et ses valeurs propres sont bornées par $\frac{2}{m} < 1$ par inégalité triangulaire. Ainsi la suite des puissances de A tend vers 0 par continuité du produit matriciel. Ainsi les suites des coefficients de A tendent vers 0, mais comme ce sont des suites entières donc elles stationnent. En somme A est nilpotente.

La matrice A est alors diagonalisable et nilpotente, ce qui montre que son polynôme minimal est X est donc que A est nulle. Ainsi, en remontant les calculs, on remarque que l'on a montré que la réduction modulo m est une injection de G dans $\mathcal{M}_n(\mathbb{Z}/m\mathbb{Z})$. En particulier, pour $m=3$, on obtient le résultat voulu. □

Référence : Sujet X-ENS maths A de 2021.

Chapitre 81

Théorème de Dirichlet faible

Ce développement a pour but de montrer le théorème de la progression arithmétique de Dirichlet dans une version faible. Le théorème général énonce que si a et b sont premiers entre eux, il existe une infinité de nombres premiers congrus à a modulo b, et sa version faible correspond à $a = 1$.

Théorème. (de la progression arithmétique de Dirichlet, version faible)
Soit $n \in \mathbb{N}^*$. Il existe une infinité de nombres premiers congrus à 1 modulo n.

Démonstration. Fixons $n \in \mathbb{N}^*$. Un nombre p est congru à 1 modulo n si et seulement si n divise $(p-1)$. Ceci équivaut à l'existence d'un élément a d'ordre n dans $\mathbb{Z}/p\mathbb{Z}$. On cherche donc $a \in \mathbb{N}^*$ tel que $\overline{a}^n - 1 = 0$ (la barre désigne la classe modulo p). Ceci se réécrit :

$$\prod_{d|n} \overline{\Phi_d(a)} = 0.$$

Chercher une racine de Φ_n dans $\mathbb{Z}/p\mathbb{Z}$ qui n'est pas racine des autres Φ_d semble alors être une bonne idée, puisqu'alors l'ordre de cet élément a ne pourra être que n. De plus, on sait que le polynôme $X^n - 1$ n'a que des racines simples dans $\mathbb{Z}/p\mathbb{Z}$, et donc que si on est racine de Φ_n, on n'est pas racine des autres.

Si on récapitule, il nous faut donc trouver des nombres premiers p qui divisent $\Phi_n(a)$ arbitrairement grands. On va prendre N un entier plus grand que 3 et $a = N!$. Puisque $a \geq 3$:

$$|\Phi_n(a)| = \prod_{\zeta \in \mu_n^*(\mathbb{C})} |a - \zeta| \geq 2,$$

on s'en convainc par un dessin. On peut donc considérer p un facteur premier de $\Phi_n(a)$. Si $p \leq N$ alors p divise $a = N!$ donc p divise tout polynôme de $X\mathbb{Z}[X]$ en a donc $\Phi_n(a) - \Phi(0)$ donc $\Phi(0)$ qui vaut 1 ou -1. Nécessairement, $p > N$, ce qu'il fallait démontrer. □

Référence : J'ai réarrangé la preuve à ma sauce, qui vient de XENS Algèbre 1.

Chapitre 82

Corps finis et polynômes irréductibles

Dans ce développement, on s'intéresse à la proportion de polynômes irréductibles de degré donné dans $\mathbb{F}_q[X]$, qui peut être pratique à connaître si on veut donner rapidement une description de $\mathbb{F}_q$ en machine comme un corps de rupture pour pouvoir faire des calculs.

Théorème. Soit $\mathcal{P}_q(d)$ l'ensemble des polynômes irréductibles de degré d sur $\mathbb{F}_q$. Alors pour $n \in \mathbb{N}^*$:

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in \mathcal{P}_q(d)} P(X).$$

Démonstration. Soit $P \in \mathcal{P}_q(d)$ et $K := \mathbb{F}_q[X]/(P)$. C'est un corps de cardinal q^d. Si d est un diviseur de n, on a alors $x^{q^d} = x$ pour tout x élément de K. Ainsi $P|(X^{q^n} - X)$ dans $\mathbb{F}_q[X]$. On peut donc faire le produit, qui divise bien $X^{q^n} - X$.

Réciproquement, soit P un facteur irréductible de degré d de $X^{q^n} - X$ dans $\mathbb{F}_q[X]$, P est scindé sur $\mathbb{F}_{q^n}$ par unicité du corps de décomposition de $X^{q^n} - X$ sur $\mathbb{F}_q$. Ainsi pour x racine de P :

$$[\mathbb{F}_{q^n} : \mathbb{F}_q] = n = [\mathbb{F}_{q^n} : \mathbb{F}_q(x)][\mathbb{F}_q(x) : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}_q(x)]\, d$$

car P est irréductible. Donc d divise n. Par ailleurs, chaque facteur ne peut apparaître qu'une fois parce que $(X^{q^n} - X)' = -1$ donc $X^{q^n} - X$ est à racines simples dans un corps de décomposition. Ainsi on a exactement l'égalité voulue. □

Corollaire. Si $I(q, d)$ est le nombre de polynômes irréductibles de degré d sur $\mathbb{F}_q$, on a en passant au degré :

$$q^n = \sum_{d|n} d\, I(d, q)$$

d'où, par inversion de Möbius :

$$I(q, n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d.$$

Remarque. Puisque $r^n := \sum_{d|n, d \neq n} \mu\left(\frac{n}{d}\right) q^d \leq \sum_{d=0}^{\lfloor n/2 \rfloor} q^d \leq \frac{q^{\lfloor n/2 \rfloor + 1} - 1}{q-1} = \mathrm{O}(q^n)$ quand q tend vers l'infini, on a

$$I_(q, n) = \frac{q^n + r_n}{n} \sim_{q \to \infty} \frac{q^n}{n}.$$

Concluons avec deux propositions utilisées et/ou en partie redémontrées.

Proposition. (injections des corps finis)
Soit q une puissance d'un nombre premier. Alors $\mathbb{F}_{q^d} \hookrightarrow \mathbb{F}_{q^n}$ si et seulement si $d|n$.

Démonstration. Le sens direct vient du fait que si on a l'injection alors on peut munir le grand corps d'une structure d'espace vectoriel sur le petit. Le fait qu'il soit forcément de dimension finie permet de conclure. Pour la réciproque, on utilise le fait que si $d|n$ alors q^d-1 divise q^n-1, et donc dans ce cas que $X^{q^d-1}-1$ divise $X^{q^n}-1$. En multipliant par X, on obtient le fait que les racines de $X^{q^d}-X$ sont dans $\mathbb{F}_{q^n}$, et ce sont exactement les éléments de $\mathbb{F}_{q^d}$. □

Proposition. Soit q une puissance d'un nombre premier. Alors $\mathbb{F}_{q^n}$ est le corps de décomposition de $X^{q^n}-X$ sur $\mathbb{F}_q$.

Démonstration. Puisque $(\mathbb{F}_{q^n})^*$ est un groupe de cardinal q^n-1, tous les éléments de $\mathbb{F}_{q^n}$ sont racine de $X^{q^n}-X$. Ainsi, $X^{q^n}-X$ est scindé à racines simples sur $\mathbb{F}_{q^n}$ et ses racines sont exactement les éléments de ce corps. □

Référence : Complément de Théo Untrau. On le trouve dans un livre d'exercices de Francinou.

Chapitre 83

Automorphismes du groupe symétrique

On se propose de montrer dans ce développement le théorème suivant.

Théorème. Tous les automorphismes de $\mathfrak{S}_n$ sont intérieurs pourvu que n soit différent de 6.

Démonstration. Pour montrer que notre automorphisme $\varphi \in \mathrm{Aut}(\mathfrak{S}_n)$ est intérieur, on va montrer qu'il envoie les transpositions sur les transpositions. En effet, si c'est le cas, il envoie la transposition $\tau_i = (1\,i)$ sur une transposition pour tout i. De plus, pour $i \neq j$, les images par φ de τ_i et τ_j ne commutent pas donc sont à supports non disjoints. On peut alors poser $\varphi(\tau_2) = (\alpha_1\,\alpha_2)$ et $\varphi(\tau_3) = (\alpha_1\,\alpha_3)$.

Mais alors, pour $i > 3$, si $\varphi(\tau_i) = (\alpha_2\,\alpha_3)$, alors puisque $(\alpha_1\,\alpha_2)(\alpha_1\,\alpha_3)(\alpha_2\,\alpha_3) = (\alpha_1\,\alpha_3)$ on aurait, en appliquant φ^{-1} :

$$(1\,2)(1\,3)(1\,i) = (1\,3)$$

ce qui n'est pas. Ainsi, $\varphi(\tau_i) = (\alpha_1, \alpha_i)$ pour des $\alpha_i \in [\![1, n]\!]$. Par injectivité de φ, α définit une permutation de $\mathfrak{S}_n$. Cette permutation vérifie

$$\varphi(\sigma) = \alpha\,\sigma\,\alpha^{-1} \quad \forall \sigma \in \mathfrak{S}_n,$$

et φ est alors intérieur. Il reste donc à montrer que notre φ envoie bien les transpositions sur les transpositions. Les transpositions correspondent à la classe de conjugaison d'une transposition. Or un automorphisme envoie une classe de conjugaison d'un élément d'ordre 2 sur une classe de conjugaison d'un élément d'ordre 2, *i.e.* d'un produit de k transpositions à supports disjoints. Ainsi, par le lemme de dénombrement des classes de conjugaison, on a, pour un certain $k \leq \frac{n}{2}$:

$$\begin{aligned}
&\binom{n}{2} = \frac{n!}{2^k k!(n-2k)!} \\
&\Leftrightarrow \frac{(n-2)!}{2^{k-1} k!(n-2k)!} = 1 \\
&\Leftrightarrow \binom{n-2}{2k-2} \times \frac{(2k-3)(2k-5)\ldots 3}{k} = 1.
\end{aligned}$$

Si $k > 3$, $2k - 3 > k$ donc l'égalité est fausse. Ainsi, pour que φ ne soit pas intérieur, il faut que k vaille 2 ou 3. Si il vaut 2, on a $(n-2)(n-3) = 4$ ce qui est impossible et s'il vaut 3, alors

$\binom{n-2}{4} = 1$ ce qui impose $n - 6$. □

Un petit mot rapide sur le lemme.

Lemme. Le cardinal de la classe de conjugaison d'une permutation est donné en fonction de son type $(k_1, \ldots, k_n)$ par la formule :

$$|\mathcal{O}_{k_1,\ldots,k_n}| = \frac{n!}{\prod_{i=1}^{n} k_i! i^{k_i}}.$$

Démonstration. Pour voir cela, on se ramène au calcul du cardinal du centralisateur associé. Ce centralisateur correspond exactement aux permutations qui commutent avec notre permutation donnée σ. Pour commuter avec σ, d'après la formule de conjugaison des cycles, il faut et il suffit d'effectuer une permutation cyclique sur chacun des k_i cycles de longueur i d'où un facteur $\prod_{i=1}^{n} i^{k_i}$, et de permuter les k_i i-cycles entre eux, ce qui correspond au produit $\prod_{i=1}^{n} k_i!$. On conclut alors par la relation orbite-stabilisateur. □

Référence : Daniel Perrin, *Cours d'algèbre*. On peut en trouver une version détaillée sur la page de Rémi Moreau.

Chapitre 84

Décomposition de Dunford

Soient $P_1, \dots, P_r$ des polynômes premiers entre eux deux à deux. Soit $Q_i = \prod_{j \neq i} P_i$, alors les (Q_i) sont premiers entre eux dans leur ensemble donc on peut écrire une relation de Bezout

$$Q_1 U_1 + \ldots + Q_r U_r = 1,$$

et on vérifie alors que $Q_i U_i(u)$ est le projecteur sur $\ker(P_i(u))$ parallèlement aux autres (ok par le lemme des noyaux).

Théorème. (décomposition de Jordan-Chevalley-Dunford)
Soit $u \in \mathcal{L}(E)$ trigonalisable. Alors $u = d + n$ avec d diagonalisable, n nilpotent et $dn = nd$. Ce couple est unique et ce sont des polynômes en u.

Démonstration. D'après le lemme des noyaux appliqué au polynôme minimal, qui est scindé par hypothèse :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} C_\lambda,$$

avec $C_\lambda \ker((u - \lambda \mathrm{id}_E)^{m_\lambda})$ le sous-espace associé à la valeur propre λ. Soit $p_\lambda \in K[u]$ le projecteur sur C_λ parallèlement à la somme directe des autres. Alors puisque les C_λ sont stables par u :

$$\begin{aligned} u &= \sum_\lambda u \circ p_\lambda \\ &= \sum_\lambda (u - \lambda \mathrm{id} + \lambda \mathrm{id}) \circ p_\lambda \\ &= \sum_\lambda \lambda \mathrm{id} \circ p_\lambda + \sum_\lambda (u - \lambda \mathrm{id}) \circ p_\lambda =: d + n \end{aligned}$$

On voit alors que d et n conviennent et sont des polynômes en u. Si $u = d' + n'$ est une autre décomposition de Dunford, d' et n' commutent à u donc à d et n donc d et d' sont codiagonalisables donc $d - d'$ est diagonalisable et n et n' commutent donc $n' - n = d - d'$ est nilpotent donc nul, et on a bien unicité. □

Lemme. Pour $M \in \mathcal{M}_n(\mathbb{C})$, on définit

$$\ln(I_n - M) = \sum_{k=1}^{n-1} \frac{M^k}{k} \in \mathbb{C}[M].$$

Alors $\exp^{(}\ln(I_n - M)) = I_n - M$.

Démonstration. On considère les séries formelles et on obtient le résultat. □

Théorème. L'exponentielle de matrices est surjective de $\mathcal{M}_n(\mathbb{C})$ dans $\mathrm{GL}_n(\mathbb{C})$.

Démonstration. Soit $M \in \mathrm{GL}_n(\mathbb{C})$ et $M = D + N = D(I_n + D^{-1}N)$ sa décomposition de Dunford (on sait que D est inversible car son polynôme caractéristique est le même que celui de M).

Par surjectivité de exp sur $\mathbb{C}$, on peut trouver des antécédents aux valeurs propres de D donc un antécédent de exp de D. Par interpolation de Lagrange, cet antécédent A peut être un polynôme en D donc en M.

Avec le lemme on obtient $\exp(\ln(I_n - D^{-1}N)) = I_n - D^{-1}N$. On pose alors $B = \ln(I_n - D^{-1}N)$, c'est un polynôme de M car tout le monde est un polynôme en M. Donc A et B commutent donc $\exp(A + B) = \exp(A)\exp(B) = M$. □

Remarque. (refaisons proprement le lemme des noyaux)
Soient $u \in \mathcal{L}(E)$ et $P_1, \ldots, P_r$ des polynômes premiers entre eux deux à deux. On considère alors $Q_i = \prod_{j \neq i} P_i$ et on écrit une relation de Bezout sur ceux-ci :

$$U_1 Q_1 + \ldots + U_r Q_r = 1.$$

On suppose dans la suite que $P_1 \cdots P_r(u) = 0$ (on peut toujours s'y ramener de toute façon). Dès lors, pour $x \in E$, on a la décomposition

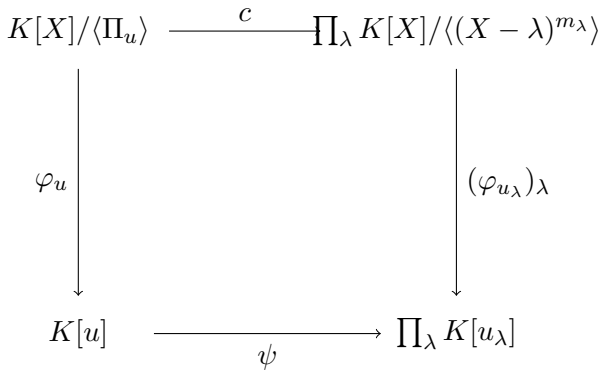
$$x = \sum_{i=1}^{r} U_i Q_i(u) =: \sum_{i=1}^{r} \pi_i(x),$$

avec $\pi_i(x) \in \ker(P_i(u))$. Les (π_i) sont des endomorphismes de E, et on vérifie que $\pi_i \circ \pi_j = 0$ si $j \neq i$ donc que ce sont des projecteurs. On a donc une somme directe, et les projecteurs sur les espaces associés sont des polynômes en u.

Remarque. Considérons le diagramme suivant. On peut montrer que c'est un diagramme commutatif. Ainsi, si on cherche $D \in K[X]$ tel que $D(u) = d$ de la décomposition de Dunford, on peut voir se problème en remontant une à une les flèches du diagramme :

- On cherche d tel que la restriction de d à C_λ soit $\lambda\mathrm{id}$
- Cela revient à évaluer chaque u_λ en un polynôme tel que $D_\lambda(u_\lambda) = \lambda\mathrm{id}$
- Cela revient à trouver D un antécédent de $(\lambda)_{\lambda \in \mathrm{Sp}(u)}$ par le théorème chinois, ce que l'on sait faire de manière algorithmique par Bezout.

Référence : Dunford et l'application sont une combinaison personnelle de toutes les preuves que je connaissais. Le diagramme provient d'un document sur AgregMaths de Thibault Monneret.


$K[X]/\langle \Pi_u \rangle$
c
$\prod_\lambda K[X]/\langle (X-\lambda)^{m_\lambda} \rangle$
φ_u
$(\varphi_{u_\lambda})_\lambda$
$K[u]$
ψ
$\prod_\lambda K[u_\lambda]$

Chapitre 85

Une équation de Mordell

Une *équation de Mordell* est une équation diophantienne de la forme

$$y^2 = x^3 + k$$

pour $k \in \mathbb{Z}$. Mordell à démontré en 1920 que ces équations avaient toujours un nombre fini de solutions.

Pour résoudre ces équations, on les met souvent sous la forme

$$x^3 = (y - \sqrt{k})(y + \sqrt{k})$$

avec $\sqrt{k} = \mathrm{i}\sqrt{-k}$ si k est négatif. Cette forme invite donc à se placer dans l'anneau $\mathbb{Z}[\sqrt{k}]$. Lorsque celui-ci possède de bonnes propriétés, on a bon espoir de pouvoir résoudre l'équation.

Théorème. L'anneau $\mathbb{Z}[\mathrm{i}\sqrt{2}]$ est euclidien donc principal donc factoriel.

Démonstration. On considère la norme $N(z = a + \mathrm{i}\sqrt{2}b) = a^2 + 2b^2$. C'est une application multiplicative à valeurs positives et qui vaut 0 si et seulement si $z = 0$. On montre que les inversibles de $Z[\mathrm{i}\sqrt{2}]$ sont les éléments de norme inversible dans $\mathbb{Z}$ donc les éléments de norme 1 donc ± 1. Si z et w sont deux éléments de $\mathbb{Z}[\mathrm{i}\sqrt{2}]$ avec w non nul, on note $\frac{z}{w} =: x + \mathrm{i}y\sqrt{2}$. On considère alors $q := c + \mathrm{i}d\sqrt{2}$ avec c et d les entiers les plus proches respectivement de x et y. On a alors :

$$\left|\frac{z}{w} - q\right|^2 = |x - c|^2 + 2|y - d|^2 \leq \frac{1}{4} + \frac{2}{4} < 1$$

donc si $r := z - qw$, alors

$$N(r) = N(w)\left|\frac{z}{w} - q\right| < N(w).$$

Ainsi $z = qw + r$ est une division euclidienne de z par w pour la jauge N, ce qui montre que notre anneau est bien euclidien. □

Théorème. L'équation de Mordell pour $k = -2$ a pour uniques solutions $(3, 5)$ et $(3, -5)$.

Démonstration. Supposons qu'il existe une solution (x, y). On cherche alors des premières informations sur x et y. Supposons que x soit pair, on a :

$$y^2 \equiv 6 \mod 8$$

on peut alors vérifier que 6 n'est pas un carré dans $\mathbb{Z}/8\mathbb{Z}$ donc x est impair. Maintenant, on se ramène comme annoncé à $\mathbb{Z}[\mathrm{i}\sqrt{2}]$. On part donc de

$$x^3 = y^2 + 2 = (y + \mathrm{i}\sqrt{2})(y - \mathrm{i}\sqrt{2}).$$

Montrons alors que $y \pm \mathrm{i}\sqrt{2}$ sont premiers entre eux dans $\mathbb{Z}[\mathrm{i}\sqrt{2}]$. On considère pour cela $p \in \mathbb{Z}[\mathrm{i}\sqrt{2}]$ un diviseur commun à $y + \mathrm{i}\sqrt{2}$ et $y + \mathrm{i}\sqrt{2}$. Ainsi p divise leur différence donc $p|2\mathrm{i}\sqrt{2}$. En passant à la norme et en utilisant sa multiplicativité, on obtient $N(p)|8$. Or $N(p)$ est un nombre impair puisque l'on a par ailleurs $N(p)|N(y + \mathrm{i}\sqrt{2}) = x^3$. Nécessairement, $N(p) = 1$ donc p est inversible, on a donc bien la primalité relative. Par unicité de la décomposition en produit d'irréductibles, $y + \mathrm{i}\sqrt{2}$ et $y - \mathrm{i}\sqrt{2}$ sont donc associés à des cubes puisqu'ils sont premiers entre eux et que leur produit est un cube. On écrit alors

$$y + \mathrm{i}\sqrt{2} = u\alpha^3 \text{ et } y - \mathrm{i}\sqrt{2} = v\beta^3$$

avec u et v inversibles et α et β dans $\mathbb{Z}[\mathrm{i}\sqrt{2}]$. On peut donc écrire $y + \mathrm{i}\sqrt{2} = (m + \mathrm{i}n\sqrt{2})^3$ et identifier parties réelle et imaginaire :

$$y = m^3 - 6mn^2 = m(m^2 - 6n^2) \text{ et } 1 = 3m^2n - 2n^3 = (3m^2 - 2n^2)n.$$

On voit directement que $n = \pm 1$. Si $n = 1$, on alors $3m^2 - 2 = 1$ donc $m = \pm 1$ *i.e.* $y = \pm 5$. Si $n = -1$ alors $1 = 3m^2$ ce qui n'est pas possible. Par conséquent $y = \pm 5$ donc $x = 3$. On vérifie réciproquement que ce sont bien des solutions. □

Remarque. La preuve du caractère euclidien puis la résolution de l'équation remplissent le format du développement. On aurait pu aussi résoudre de la même manière le cas $k = -1$ mais cela serait trop rapide et cela permet d'introduire un anneau encore différent.

Théorème. L'équation de Mordell pour $k - 1$ admet pour unique solution $(1, 0)$.

Démonstration. On se place dans $\mathbb{Z}[\mathrm{i}]$ et tout est alors similaire jusqu'à arriver à $y + \mathrm{i} = (m + \mathrm{i}n)^3$. On identifie alors encore les parties réelle et imaginaire ; les valeurs changent légèrement, mais on conclut de manière simple comme dans le cas précédent. □

Référence : Duverney, *Théorie des nombres.*

Quatrième partie

Développement d'analyse et probabilités

Chapitre 86

Fonction gamma et formule des compléments

Dans ce développement, on définit la fonction gamma sur $\mathbb{C}\setminus -\mathbb{N}$ et on montre la formule des compléments.

Définition. On définit la fonction gamma sur $\Omega_0 := \{z \in \mathbb{C}, \mathrm{Re}(z) > 0\}$ par :

$$\Gamma(z) = \int_0^{+\infty} t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t.$$

Proposition. La fonction gamma est bien définie sur Ω_0, et définit une fonction holomorphe sur tout ce domaine.

Démonstration. Pour montrer que Γ est bien définie, on écrit $|t^{z-1}| = |\mathrm{e}^{(z-1)\ln(t)}| = \mathrm{e}^{(\mathrm{Re}(z)-1)\ln(t)} = t^{\mathrm{Re}(z)-1}$, qui est intégrable au voisinage de 0 et négligeable devant $t^2\mathrm{e}^{-t}$ au voisinage de l'infini. Ainsi, le module de l'intégrande est intégrable sur $\mathbb{R}_+$, et on est bons.

Pour montrer qu'elle est holomorphe, on va passer par le théorème d'holomorphie sous l'intégrale. Il est clair que pour t fixé, $t^{z-1}\mathrm{e}^{-t} = \mathrm{e}^{(z-1)\ln(t)}\mathrm{e}^{-t}$ est holomorphe en z, et que pour z fixé, cette même expression est continue en t. Soient maintenant $\beta > \alpha > 0$. Pour z de partie réelle supérieure à α, et inférieure à β, on a :

$$|t^{z-1}\mathrm{e}^{-t}| = t^{\mathrm{Re}(z)-1}\mathrm{e}^{-t} \leq t^{\alpha-1}\mathrm{e}^{-t}\mathbf{1}_{]0,1]}(t) + t^{\beta-1}\mathrm{e}^{-t}\mathbf{1}_{[1,+\infty[}(t) =: g(t),$$

où g est intégrable et ne dépend pas de z. Ainsi, on peut dominer par une fonction intégrable qui ne dépend pas de z sur tout compact et conclure. □

Remarque. En effectuant une intégration par parties, on peut remarquer que Γ satisfait l'équation fonctionnelle :

$$\Gamma(z+1) = z\Gamma(z),$$

et qu'en particulier pour $n \in \mathbb{N}$, $\Gamma(n+1) = n!$. En effet :

$$\Gamma(z+1) = \left[-t^z\mathrm{e}^{-t}\right]_0^{\infty} + z\int_0^{\infty} t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t = z\Gamma(z).$$

Théorème. La fonction gamma se prolonge en une fonction méromorphe sur $\Omega := \mathbb{C}\backslash -\mathbb{N}$, avec un pôle simple en $-n$ pour tout $n \in \mathbb{N}$, de résidu $\frac{(-1)^n}{n!}$.

Démonstration. On va se servir de l'équation fonctionnelle précédente. Soit $\Omega_n = \{z, \mathrm{Re}(z) > -n\}\backslash\{0, -1, \ldots, -(n-1)\}$. Pour $z \in \Omega_0$, on peut écrire :

$$\Gamma(z) = \frac{\Gamma(z+n)}{(z-n+1)\cdots(z+1)z}.$$

Et on peut alors remarquer que cette expression définit une fonction holomorphe sur Ω_n, qui coïncide donc avec Γ sur Ω_0. Ceci valant pour tout n, on crée une suite de prolongements qui coïncident deux à deux car ils sont toujours égaux sur Ω_0. L'expression utilisée pour prolonger gamma montre que ses seuls pôles sont les entiers négatifs. De plus, elle permet de calculer les résidus :

$$\mathrm{Res}(\Gamma, -n) = \lim_{z\to -n} (z+n)\Gamma(z) = \lim_{z\to -n} \frac{\Gamma(z+n+1)}{(z-n+1)\cdots(z+1)z} = \frac{\Gamma(1)}{(-1)\cdots(-n)} = \frac{(-1)^n}{n!}.$$

□

Théorème. (Formule des compléments) Pour tout $z \in \mathbb{C}\backslash\mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Démonstration. Si on montre cette formule pour $z \in]0,1[$, on pourra conclure par le principe des zéros isolés car $]0,1[$ est un sous-ensemble de Ω qui admet un point d'accumulation, puisque les deux membres de l'égalité sont holomorphes sur $\mathbb{C}\backslash\mathbb{Z}$. Soit donc $s \in]0,1[$. D'après le théorème de Fubini :

$$\Gamma(s)\Gamma(1-s) = \int_0^\infty \int_0^\infty u^{-s}\mathrm{e}^{-u}t^{s-1}\mathrm{e}^{-t}\,\mathrm{d}u\mathrm{d}t = \int_0^\infty \int_0^\infty (vt)^{-s}\mathrm{e}^{-vt}t^s\mathrm{e}^{-t}\,\mathrm{d}v\mathrm{d}t = \int_0^\infty \frac{v^{s-1}}{1+v}\,\mathrm{d}v.$$

où l'on a posé $v = \frac{u}{t}$, et où l'on a éventuellement utilisé la symétrie du problème à la dernière égalité. Ainsi, si on pose $v = \mathrm{e}^x$, on se ramène finalement à calculer l'intégrale suivante :

$$I(s) = \int_{\mathbb{R}} \frac{\mathrm{e}^{sx}}{1+\mathrm{e}^x}\,\mathrm{d}x.$$

Pour calculer cette intégrale, on va utiliser le théorème des résidus. On considère le contour rectangulaire γ_R dont les sommets sont $-R, R, R+2\mathrm{i}\pi, -R+2\mathrm{i}\pi$, parcouru dans le sens direct. À l'intérieur de ce contour, $g : z \mapsto \frac{\mathrm{e}^{sx}}{1+\mathrm{e}^x}$ définit une fonction méromorphe, avec un pôle simple en $\mathrm{i}\pi$, dont le résidu vaut :

$$\mathrm{Res}(g, \mathrm{i}\pi) = \lim_{z\to \mathrm{i}\pi} (z-\mathrm{i}\pi)\frac{\mathrm{e}^{sz}}{1+\mathrm{e}^z} = \lim_{z\to \mathrm{i}\pi} \mathrm{e}^{sz}\frac{z-\mathrm{i}\pi}{\mathrm{e}^z - \mathrm{e}^{\mathrm{i}\pi}} = \mathrm{e}^{\mathrm{i}s\pi}\mathrm{e}^{-\mathrm{i}\pi} = -\mathrm{e}^{\mathrm{i}s\pi}.$$

D'après le théorème des résidus, on a, pour $R > 0$:

$$\begin{aligned}\int_{\gamma_R} g(z)\,\mathrm{d}z &= -2\mathrm{i}\pi \mathrm{e}^{-\mathrm{i}s\pi} \\ &= \int_{-R}^{R} g(x)\,\mathrm{d}x - \int_{-R}^{R} g(x+2\mathrm{i}\pi)\,\mathrm{d}x + \int_0^{2\pi} g(R+\mathrm{i}\theta)\mathrm{i}\,\mathrm{d}\theta - \int_0^{2\pi} g(-R+\mathrm{i}\theta)\mathrm{i}\,\mathrm{d}\theta.\end{aligned}$$

Le premier terme tend vers $I(s)$ l'intégrale de g sur $\mathbb{R}$ car g est intégrable donc on peut faire une convergence dominée. Le deuxième terme vaut :

$$-\int_{-R}^{R} \frac{\mathrm{e}^{sx+2\mathrm{i}\pi s}}{1+\mathrm{e}^{2\mathrm{i}\pi+x}}\,\mathrm{d}x = -\mathrm{e}^{2\mathrm{i}\pi s}\int_{-R}^{R} g(x)\,\mathrm{d}x,$$

et tend donc vers $-\mathrm{e}^{2\mathrm{i}\pi s}I(s)$ lorsque R tend vers l'infini. Les deux autres termes se majorent directement sans convergence dominée par une exponentielle décroissante donc tendent vers 0. Somme toute, on a donc :

$$I(s)(1-\mathrm{e}^{2\mathrm{i}\pi s}) = -2\mathrm{i}\pi\mathrm{e}^{\mathrm{i}\pi s}.$$

En utilisant les formules d'Euler, on obtient finalement :

$$I(s) = -2\mathrm{i}\pi\frac{\mathrm{e}^{\mathrm{i}\pi s}}{1-\mathrm{e}^{2\mathrm{i}\pi s}} = \frac{\pi}{\sin(\pi s)}.$$

On a donc bien montré la formule des compléments. □

Corollaire. La fonction Γ ne s'annule pas. De plus, la fonction $\frac{1}{\Gamma}$ est entière.

Démonstration. Si gamma s'annule, la formule des compléments implique que $\frac{1}{\sin}$ s'annule sur $\mathbb{C}$, ce qui n'est pas. Ainsi, on peut écrire pour $z \in \mathbb{C}\backslash\mathbb{Z}$:

$$\frac{1}{\Gamma(z)} = \Gamma(1-z)\frac{\sin(\pi z)}{\pi},$$

avec l'expression de droite, on voit que $\frac{1}{\Gamma}$ possèdes des singularités effaçables en les entiers négatifs, donc qu'elle se prolonge en une fonction holomorphe sur $\mathbb{C}\backslash\mathbb{N}^*$. Or en les entiers positifs l'expression de droite n'est pas définie mais $\frac{1}{\Gamma}$ l'est. Ainsi, par prolongement analytique, on peut bien définir $\frac{1}{\Gamma}$ sur $\mathbb{C}$ tout entier. □

Référence : Stein et Sharkachi, *Complex analysis.*

Chapitre 87

Théorème de Charkovski

Dans ce développement, on présente un résultat élémentaire d'analyse réelle, le théorème de Charkovski, qui s'énonce souvent ainsi :

"Trois-cycle implique chaos."

En effet, il stipule qu'une fonction définie d'un intervalle fermé dans lui-même possède un point de toute période dès lors qu'elle possède un point de période trois. Précisons le propos en introduisons les définitions nécessaires à l'énoncé du théorème.

Dans tout le développement, f désigne une fonction continue définie sur un segment fermé D et à valeurs dans D. La notation f^k désigne la fonction $f \circ \ldots \circ f$, avec k termes.

Définition. Un point $x \in D$ est de période p ou p-périodique lorsque :

$$f^p(x) = x \text{ et } f^k(x) \neq x \text{ pour tout } k \in \{1, \ldots, p-1\}.$$

Pour $p = 1$, cela correspond à un point fixe.

Proposition. Soit g une fonction continue définie d'un segment $[a, b]$ dans lui-même. Alors g admet un point fixe.

Démonstration. Soit $h(x) = g(x) - x$. La fonction h est continue sur $[a, b]$ et $h(a)$ est positif alors que $h(b)$ est négatif. D'après le théorème des valeurs intermédiaires, h doit s'annuler sur $[a, b]$, ce qui revient exactement à dire que g possède un point fixe. □

Définition. Soient I et J deux segments inclus dans D. On dit que I recouvre J (par f) lorsque $J \subset f(I)$. On note alors $I \rightsquigarrow J$.

Avec cette notation, la proposition précédente se réécrit : si $I \rightsquigarrow I$, alors f admet un point de période 1. On peut généraliser ce résultat à des périodes plus grandes.

Lemme. Si $I \rightsquigarrow J$, alors il existe un segment $K \subset I$ tel que $f(K) = J$.

Démonstration. Soit $J = [a, b]$ et α et β des antécédents respectifs de a et b dans I. On considère alors :

$$u = \max\{x \in [\alpha, \beta], f(x) = a\} \text{ et } v = \min\{x \in [u, \beta], f(x) = b\},$$

et $K = [u, v]$. On sait que $f(u) = a$ et $f(v) = b$ donc par le théorème des valeurs intermédiaires, $J \subset f(K)$. Réciproquement, soit $y \in f(K)$ si $y < a$, alors on a un $x \in]u, v[$ tel que $f(x) < a$, et par le théorème des valeurs intermédiaires, il existe un $u' > u$ tel que $f(u') = a$, ce qui contredit la maximalité de u. On fait pareil de l'autre côté, ce qui achève la preuve du lemme. □

Proposition. Supposons que l'on ait une chaîne $I_0 \rightsquigarrow I_1 \rightsquigarrow \cdots \rightsquigarrow I_{n-1} \rightsquigarrow I_0$. Alors il existe $x \in I_0$ tel que $f^n(x) = x$, et $f^k(x) \in I_k$ pour $k \in \{0, \ldots, n-1\}$.

Démonstration. D'abord, on exhibe avec le lemme un segment $K_{n-1} \subset I_{n-1}$ tel que $f(K_{n-1}) = I_0$. Ensuite, on exhibe $K_{n-2} \subset I_{n-2}$ tel que $f(K_{n-2}) = K_{n-1}$. De proche en proche, on se retrouve avec une suite $K_0 \rightsquigarrow K_1 \rightsquigarrow \cdots \rightsquigarrow K_{n-1} \rightsquigarrow I_0$, où l'on a à chaque fois égalité. Ainsi, comme $K_0 \subset I_0 = f^n(K_0)$, il existe d'après la première proposition un $x \in K_0$ tel que $f^n(x) = x$. Mais comme $x \in K_0$ et que l'on a la chaîne précédente, on a $f^k(x) \in K_k \subset I_k$ pour $k \in \{0, \ldots, n-1\}$, ce qui achève la preuve. □

Théorème. (de Charkovski)
Si f possède un point de période 3, alors elle possède un point à toute période.

Démonstration. Soit a un point de période trois, et $b = f(a)$, $c = f(b)$ (et alors $f(c) = a$). Par symétrie on peut supposer $a < b$ et $a < c$. Il y a alors deux cas à traiter. Dans tous les cas, on sait que l'on a un point fixe.

Commençons par le cas $a < b < c$. On note $I_0 = [a, b]$ et $I_1 = [b, c]$, d'après le théorème des valeurs intermédiaires, on a $I_0 \rightsquigarrow I_1 \rightsquigarrow I_0$. Avec la proposition précédente, on peut considérer un $x \in I_0$ tel que $f^2(x) = x$. Si $f(x) = x$, alors nécessairement $x \in I_1 \cap I_0$ donc $x = b$, ce qui contredit $f(x) = x$. Donc ce x est bien un point de période 2. Soit maintenant $n \geq 4$. On considère la suite $I_0 \rightsquigarrow I_1 \rightsquigarrow \ldots \rightsquigarrow I_1 \rightsquigarrow I_0$, avec $(n-1)$ fois I_1. On considère à nouveau un x comme dans la proposition précédente. S'il existe un $k \in \{1, \ldots, n-1\}$ tel que $f^k(x) = x$, alors $x \in I_0 \cap I_1$, ce qui entraîne à nouveau $x = b$ donc $f^2(x) = a \notin I_1$, ce qui contredit les hypothèses. Ainsi, f admet bien un point de toute période.

Dans le cas $a < c < b$, avec $I_0 = [a, c]$ et $I_1 = [c, b]$ on a toujours les mêmes chaînes d'intervalles. Dans le cas $n = 2$, si x est comme dans la proposition avec $f(x) = x$, alors $x = c$ et x ne peut pas être de période 2. Dans le cas $n \geq 4$, et pour un x de la proposition avec $f^k(x)$, $k \in \{1, \ldots n-1\}$, on a $x = c$, mais alors $f^3(x) = a \notin I_1$, ce qui contredit les hypothèses. □

Remarque. Ce qui coince dans le cas où f a un point périodique de période 4 survient dans le cas où l'orbite n'est pas croissante. Dans le cas $x < f(x) < f^2(x) < f^3(x)$, on peut faire la même preuve et avoir un point périodique à toute période.

Exemple. La fonction $x \mapsto (x + \frac{1}{2})\mathbf{1}_{[0,\frac{1}{2}]} + (4 - 2x)\mathbf{1}_{[\frac{1}{2},1]}$, aussi simple soit-elle, admet des points périodiques à toutes les périodes.

Par contre, la fonction $x \mapsto (1 - x)$ n'a que des points de période 2, et on voit que l'existence d'un point de période trois est importante.

Référence : Francinou, Gianella, Nicolas : *Oraux X-ENS Analyse 1.*

Chapitre 88

Théorème de Krein-Milman

Définition. Soit $C \subset E$ convexe. On dit que $c \in C$ est un *point extrémal* de C si $C\backslash\{c\}$ est convexe, ou de manière équivalente si $\frac{x+y}{2} = c$ entraîne $x = y = c$ pour $x, y \in C$ (c ne peut pas s'écrire comme milieu strict de points de C).

Démonstration. Supposons que c ne peut pas s'écrire comme milieu strict de points de C et montrons que $C\backslash\{c\}$ est convexe. Soient $x, y \in C$ qui ne sont pas c. On sait déjà que pour tout $t \in [0,1]$, $tx + (1-t)y \in C$. Maintenant, s'il existe $t_0 < 1/2$ [1] tel que $t_0 x + (1-t_0)y = c$, alors c serait le milieu du segment $[x, 2t_0 + (1-2t_0)]$ ce qui n'est pas.
Réciproquement, si c est le milieu de $[x, y]$, avec x et y différents de c, alors $C\backslash\{c\}$ n'est pas connexe puisque le segment $[x, y]$ contiendrait un point qui n'est pas dans $C\backslash\{c\}$ (en l'occurrence c, j'avoue c'est lourd). □

Définition. On dit que H est un hyperplan d'appui $a^\perp$ à C en $c \in C$ si $c \in H$ et si C est inclus dans $\{x \in E, \langle x - c, a\rangle \geq 0\}$ ou $\{x \in E, \langle x - c, a\rangle \leq 0\}$.

Proposition. Soit C un convexe fermé non vide. Si $c \in \partial C$, il existe un hyperplan d'appui à C en c.

Démonstration. On raisonne "par dualité" en cherchant une suite de directions orthogonales ; plutôt que de trouver une suite d'hyperplans qui converge vers notre hyperplan d'appui, on va regarder la suite des directions orthogonales, ce qui est plus dans nos cordes...

Soit $c \in \partial C$. En particulier, $c \in \overline{E\backslash C}$. Donc il existe (x_k) une suite d'éléments de $(E\backslash C)$ qui converge vers c. D'après le théorème de projection sur un convexe fermé, si on note p la projection sur C, l'hyperplan affine H_k orthogonal à $x_k - p(x_k)$ passant par x_k est tel que C est inclus dans un demi-espace formé par H_k [2]. Si on note alors $u_k = \frac{x_k - p(x_k)}{\|x_k - p(x_k)\|}$, on peut, quitte à extraire, supposer que u_k converge vers u de norme 1 (on vient de faire "converger" les hyperplans, et on va considérer l'hyperplan limite). Soit alors H d'équation $\langle x, u\rangle = \langle c, u\rangle$, et $z \in C$. Pour tout k

$$\langle u_k, z - p(x_k)\rangle \leq 0$$

1. On peut toujours se ramener à ce cas quitte à échanger x et y, et non ce n'est pas $1/2^2$.
2. Cela peut sembler obscur dit comme ça, mais c'est exactement une des conséquences du théorème, que Karine appelle caractérisation par l'angle obtus.

et donc quand k tend vers l'infini (avec la continuité de p et du produit scalaire) :

$$\langle u, z - p(c)\rangle = \langle u, z - c\rangle \leq 0,$$

ce qui montre que H est bien un hyperplan d'appui. □

Proposition. Soit $a \in C$, on suppose que a est dans un hyperplan d'appui H de C. Alors a est point extrémal de C si et seulement si a est point extrémal de $H \cap C$.

Démonstration. Si a est extrémal, alors $C\backslash\{a\}$ est convexe. De plus :

$$(C\backslash\{a\}) \cap H = (C \cap H)\backslash\{a\}$$

puisque $a \in H$, or comme $C\backslash\{a\}$ et H sont convexes, $C\backslash\{a\} \cap H$ l'est donc $(C \cap H)\backslash\{a\}$ est convexe, ce qui veut dire par définition que a est point extrémal de $H \cap C$. Réciproquement, si a est extrémal dans $H \cap C$, on considère u et v dans C et $a = \frac{u+v}{2}$. Soit φ tel que $H = \{\varphi = \lambda\}$. On peut supposer que $\varphi(u)$ (et donc $\varphi(v)$) est inférieur à λ. Mais alors par linéarité, $\varphi(a) = \frac{\varphi(u)+\varphi(v)}{2} = \lambda$ et donc $\varphi(u) = \varphi(v) = \lambda$ donc $u, v \in H$. Ceci permet de conclure. □

Ce qui est dur dans ces preuves et de bien penser à toutes les caractérisations d'un hyperplan, d'un point extrémal il faut avoir une bonne vision des notions et du problème.

Théorème. (de Krein-Milman)
Soit C un convexe compact non vide de $\mathbb{R}^d$. Alors il est l'enveloppe convexe de ses points extrémaux.

Démonstration. On procède par récurrence sur la dimension (la proposition précédente nous invite à le faire). Le résultat est vrai en dimension 1 (les convexes de $\mathbb{R}$ sont les segments, dont on connaît les points extrémaux). On suppose le résultat vrai en dimension inférieure ou égale à $d-1$.

Soit donc $c \in C$. On cherche à l'écrire comme barycentres de points extrémaux :

- c est à la frontière de C. Soit H un hyperplan d'appui en c et $C' = C \cap H$, on applique l'hypothèse de récurrence avec notre proposition précédente et on est bons.
- c est à l'intérieur de C. Soit D une droite qui passe par c. Alors $D \cap C$ est compact convexe inclus dans une droite donc c'est un segment $[a, b]$ avec a et b des points de ∂C. On conclut alors par associativité des barycentres.

□

Référence : Un des oraux XENS. Pour une application, il y a quelque chose de super sur la page de Thomas Cavallazzi.

Chapitre 89

Théorème de Hadamard-Lévy

Le théorème d'inversion globale est insatisfaisant dans le sens où prouver l'injectivité d'une fonction revient à exhiber sa réciproque. Le théorème d'Hadamard-Lévy demande une hypothèse qui est souvent plus simple à vérifier.

Définition. Une fonction $f : \mathbb{R}^n \to \mathbb{R}^n$ est dite *propre* lorsque pour l'image réciproque de tout compact est compacte.

Proposition. Une fonction continue est propre si et seulement si $\|f(x)\| \xrightarrow[\|x\|\to\infty]{} \infty$.

Démonstration. Si f ne tend pas vers l'infini quand la norme tend vers l'infini, alors il existe un ensemble non borné A tel que $f(A)$ est inclus dans une boule fermée, et alors l'image réciproque de cette boule fermée ne peut pas être compacte *i.e.* f n'est pas propre. Réciproquement, si f n'est pas propre, il existe un compact K tel que $f^{-1}(K)$ ne soit pas compact. Mais f étant continue, cet ensemble est pourtant fermé donc forcément il n'est pas bornée et on a bien le résultat. □

Théorème. (Hadamard-Lévy)
Soit $f \in \mathcal{C}^2(\mathbb{R}^n)$, alors f est un $\mathcal{C}^1$-difféomorphisme sur $\mathbb{R}^n$ si et seulement si f est propre et pour tout $x \in \mathbb{R}^n$, $\mathrm{d}f(x)$ est inversible.

Démonstration. Tout d'abord, si f est un $\mathcal{C}^1$-difféomorphisme, alors en différentiant $f^{-1} \circ f = \mathrm{id}$, on obtient que $\mathrm{d}f(x)$ est inversible. Par ailleurs, f est propre puisque f^{-1} est continue donc l'image par f^{-1} (qui correspond à l'image réciproque par f) de tout compact est compacte.

Réciproquement, on suppose que f est propre et que sa différentielle en tout point est inversible. On souhaite appliquer le théorème d'inversion globale, mais pour cela il nous faut l'injectivité de f. On veut montrer que pour tout z, l'ensemble $f^{-1}(\{z\})$ est réduit à un élément. Quitte à tout décaler, on se ramène à $z = 0$. Toute la preuve repose sur l'introduction du flot suivant :

$$F : x \mapsto -\mathrm{d}f(x)^{-1} \cdot f(x),$$

qui est une application $\mathcal{C}^1$ puisque f est $\mathcal{C}^2$ et que l'inversion est continue. On considère alors le problème de Cauchy

$$y' = F(y), \quad y(0) = q \in \mathbb{R}^n.$$

D'après le théorème de Cauchy-Lipschitz, ce problème admet une unique solution maximale

$$\varphi(\cdot, q) : [0, T^*[\mapsto \mathbb{R}^n.$$

On veut montrer que les trajectoires restent dans un compact pour avoir des solutions globales. Mais puisque f est propre, cela revient à montrer que l'image par f de ces trajectoires est bornée. Soit donc

$$g : t \mapsto f(\varphi(t, q)),$$

qui et $\mathcal{C}^1$ par composition et qui vérifie :

$$g'(t) = \mathrm{d}f(\varphi(t, q))(\partial_t \varphi(t, q)) = -g(t).$$

Ainsi pour tout t, $f(\varphi(t, q)) = \mathrm{e}^{-t} f(q)$, *i.e.* le flot est à valeurs dans $f^{-1}(\overline{B}(0, \|f(q)\|))$, compact car f est propre. Donc $T^* = +\infty$. En particulier, il existe une suite (t_k) tendant vers l'infini et un vecteur y tels que

$$\varphi(t_k, q) \xrightarrow[k\to\infty]{} y.$$

Par continuité de f, $f(y) = 0$ donc f est surjective.

Maintenant, on veut avoir la convergence globale du flot vers y. Pour cela, on veut passer f de l'autre côté de l'expression obtenue précédemment. Pour cela, on applique le théorème d'inversion locale : on sait que f est un difféomorphisme d'un ouvert U^y dans une boule fermée B^y de rayon δ_y. On sait, d'après ce qui précède, qu'il existe un $t_0 > 0$ tel que $\varphi(t_0, q) \in U^y$. On voit alors que

$$\{t > t_0, \varphi(t, q) \in U^y\} = \{t > t_0, \varphi(t, q) = (f_{|U_y})^{-1}(\mathrm{e}^{-t} f(q))\}.$$

En effet, l'inclusion réciproque est immédiate en appliquant f, et l'inclusion directe vient du fait que pour tout $t > t_0$, $\mathrm{e}^{-t} f(q) = f(\varphi(t, q)) \in B^y$ par décroissance de $u \mapsto \mathrm{e}^{-u}$ donc on peut appliquer $f^{-1}_{|U_y}$. Or l'ensemble de gauche est ouvert et celui de droite est fermé, donc par connexité, ces deux ensembles sont égaux à $[t_0, \infty[$. Ainsi, on peut faire tendre t vers l'infini dans l'expression

$$\varphi(t, q) = (f_{|U^y})^{-1}(\mathrm{e}^{-t} f(q)),$$

et obtenir que le flot converge globalement vers y le seul zéro de f dans U^y.

Finalement, si on note W^y l'ensemble des conditions initiales qui font converger le flot vers y, on obtient une partition :

$$\mathbb{R}^n = \bigcup_{y \in f^{-1}(\{0\})} W^y.$$

Soit $q \in W^y$. On sait qu'il existe un $T > 0$ tel que $\varphi(T, q) \in \mathrm{B}(y, \eta_y)$, où η_y est choisi pour avoir $\mathrm{B}(y, 2\eta_y) \subset U^y$. Par continuité du flot, on peut trouver un $\delta > 0$ tel que $\|\varphi(T, q) - \varphi(T, q')\| < \eta_y$ dés que $\|q - q'\| < \delta$. On a donc $\varphi(T, q') \in \mathrm{B}(y, 2\eta_y) \subset U^y$. Ainsi la trajectoire partant de q' converge vers y donc $\mathrm{B}(q, \delta) \subset W^y$, qui est donc une partie ouverte. Comme les $(W^y)_y$ forment une partition d'ensembles non vides ouverts de $\mathbb{R}^n$, ils ne font qu'un, c'est-à-dire que f est

injective donc bijective donc un $\mathcal{C}^1-$difféomorphisme par inversion globale. □

Application. Il existe $\varepsilon > 0$ tel que pour tout $a \in \mathbb{R}^n$, pour tout $L \in \mathrm{GL}_n(\mathbb{R})$ qui vérifie $\|a\| < \varepsilon$ et $\|L - I_n\| < \varepsilon$, alors il existe un difféomorphisme f de $\mathbb{R}^n$ tel que :

— $\forall \|x\| < 1, f(x) = L(x) + a$

— $\forall \|x\| > 2, f(x) = x$.

Démonstration. On considère une fonction plateau $\chi \in \mathcal{D}(\mathbb{R})$ qui vaut 1 sur $[-1, 1]$ et 0 en dehors de $[-2, 2]$. Soit $\|\cdot\|$ la norme euclidienne et $f : \mathbb{R}^n \to \mathbb{R}^n$ définie par

$$f(x) = x + \chi(\|x\|^2)(Lx + a - x),$$

où L est inversible.
Cette application est clairement $\mathcal{C}^1$. En dehors d'une boule, cette application est l'identité. Ainsi, on a bien $\|f(x)\| \to +\infty$ quand $\|x\|$ tend vers l'infini. Ainsi f est propre. Pour $\|x\| < 1$ et $\|x > 2\|$ se calcule directement et est inversible. Pour un x de norme comprise entre 1 et 2 :

$$\mathrm{d}f(x)(h) = h + \chi(\|x\|^2)(Lh - h) + 2\langle x, h\rangle\chi'(\|x\|^2)(Lx + a - x) \qquad \forall h \in \mathbb{R}^n.$$

Ainsi

$$\|\mathrm{d}f(x) - \mathrm{Id}\| \leq (1 + 4\|\chi\|_\infty)\|L - \mathrm{Id}\| + 4\|\chi'\|_\infty\|a\|.$$

Ainsi, pour L assez proche de l'identité et a assez proche de 0, la différentielle de f est inversible pour tout x. On peut donc appliquer le théorème d'Hadamard-Lévy et conclure. □

Référence : Claude Zuily et Hervé Queffélec, *Analyse pour l'agrégation.* L'application provient de Thomas Cavallazzi.

Chapitre 90

Lemme de Morse

Dans ce développement, on montre le Lemme de Morse, qui permet d'étudier le comportement d'une fonction de plusieurs variables au voisinage d'un point critique non dégénéré. Pour commencer, on précise le résultat d'orthogonalisation simultanée sur $\mathbb{R}$, en voyant que la réduction peut être $\mathcal{C}^1$ au voisinage d'une forme quadratique non dégénérée.

Proposition. (Orthogonalisation simultanée $\mathcal{C}^1$)

Soit A_0 une matrice symétrique réelle inversible. Alors il existe un voisinage V de A_0 dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ telle que

$$A = \rho(A)^\top A_0 \rho(A) \qquad \forall A \in V.$$

Démonstration. On considère l'application f définie par $f(M) = M^\top A_0 M$ de $\mathcal{M}_n(\mathbb{R})$ dans $\mathcal{S}_n(\mathbb{R})$. C'est une application $\mathcal{C}^1$ car polynomiale en les coefficients de M. De plus, un calcul simple donne :

$$\mathrm{d}f(I_n).H = H^\top A_0 + A_0 H,$$

donc $\ker(\mathrm{d}f(I_n)) = A_0^{-1}\mathcal{A}_n(\mathbb{R})$. Soit $F = A_0^{-1}\mathcal{S}_n(\mathbb{R})$, qui est un supplémentaire de $\ker(\mathrm{d}f(I_n))$, et considérons ψ la restriction de f à F. Étant donné que I_n est dans F, l'expression de $\mathrm{d}\psi(I_n)$ est la même que pour f. De plus, d'après le théorème du rang, $\mathrm{d}\psi(I_n)$ est maintenant bijective. Par inversion locale, il existe un voisinage U de I_n dans F (que l'on peut supposer inclus dans $\mathrm{GL}_n(\mathbb{R})$ car c'est un ouvert), et V un voisinage de A_0 dans $\mathcal{S}_n(\mathbb{R})$ tels que ψ soit un $\mathcal{C}^1$-difféomorphisme de U sur V. Si on pose $\rho := \psi^{-1}$, on a alors :

$$A = \psi(\rho(A)) = (\rho(A))^\top A_0 \rho(A) \qquad \forall A \in U.$$

□

Théorème. (Lemme de Morse) Soit U un ouvert de $\mathbb{R}^n$ qui contient 0 et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^3$. On suppose que la différentielle de f en 0 est nulle et que la hessienne de f en 0 est non-dégénérée, de signature $(p, n-p)$.

Alors il existe un $\mathcal{C}^1$-difféomorphisme φ entre deux voisinages V_0 et $W_0 = \varphi(V_0)$ de 0 dans $\mathbb{R}^n$

tels que $\varphi(0) = 0$ et :

$$f(x) - f(0) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

Démonstration. Quitte à restreindre U, on peut le supposer égal à une boule ouverte. Sur cette boule (convexe) on peut appliquer la formule de Taylor avec reste intégral à l'ordre deux (avec $\mathrm{d}f(0) = 0$) :

$$f(x) - f(0) = \int_0^1 (1-t)x^\top \mathrm{H}f(tx)x \,\mathrm{d}t =: x^\top Q(x)x,\ Q(x) \in \mathcal{S}_n(\mathbb{R}).$$

La fonction Q est $\mathcal{C}^1$ car f est $\mathcal{C}^3$ et par dérivation sous l'intégrale. De plus, $Q(0) = \frac{1}{2}\mathrm{H}f(0)$ qui est inversible et de signature $(p, n-p)$ par hypothèse. Le lemme précédent assure alors qu'il existe un voisinage V de $Q(0)$ dans $\mathcal{S}_n(\mathbb{R})$ et une application $\rho \in \mathcal{C}^1(V, \mathrm{GL}_n(\mathbb{R}))$ tels que :

$$\forall A \in V, \quad A = (\rho(A))^\top Q(0)\rho(A).$$

Par continuité de Q, il existe W un voisinage de 0 dans $\mathbb{R}^n$ tel que $Q(W) \subset V$. On a donc :

$$\forall x \in W, \quad f(x) - f(0) = (\rho(Q(x))x)^\top Q(0)\rho(Q(x))x.$$

D'après le théorème d'inertie de Sylvester, si Δ est la matrice diagonale avec p 1 et $(n-p)$ -1, alors il existe $M \in \mathrm{GL}_n(\mathbb{R})$ telle que $Q(0) = M^\top \Delta M$. Pour $x \in W$, on pose alors $\varphi(x) = M\rho(Q(x))x$, qui est $\mathcal{C}^1$ par composition, et qui vérifie donc :

$$f(x) - f(0) = \varphi(x)^\top \Delta \varphi(x) = \varphi_1(x)^2 + \ldots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \ldots - \varphi_n(x)^2.$$

Il reste à voir que φ est un difféomorphisme local. Pour cela, on remarque que $\mathrm{d}\varphi(0) = M\rho(Q(0)) = M \in \mathrm{GL}_n(\mathbb{R})$, et on applique le théorème d'inversion locale. □

Référence : François Rouvière, *Petit guide de calcul différentiel.*

Chapitre 91

Méthode de Laplace généralisée

Dans ce développement, on généralise la méthode de Laplace sur $\mathbb{R}^n$.

Théorème. Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ de classe $\mathcal{C}^3$ et $a : \mathbb{R} \to \mathbb{R}$ continue et intégrable. On suppose que φ admet un minimum global en x_0 avec $\mathrm{H}\varphi(x_0)$ inversible (et donc strictement positive). On suppose de plus que

$$\liminf_{\|x\|\to\infty} \varphi(x) > \varphi(x_0).$$

Alors on a l'équivalent (quand t tend vers l'infini) :

$$\int_{\mathbb{R}^d} a(x)\mathrm{e}^{-t\varphi(x)}\,\mathrm{d}x \sim \frac{a(x_0)}{\sqrt{\det(\mathrm{H}\varphi(x_0))}}\left(\frac{2\pi}{t}\right)^{d/2}\mathrm{e}^{-t\varphi(x_0)}.$$

Démonstration. D'après le lemme de Morse appliqué en x_0 à φ. Il existe alors f un $\mathcal{C}^1$ difféomorphisme d'un voisinage V de x_0 vers un voisinage W de x_0 tel que pour tout $x \in V$:

$$\varphi(x) - \varphi(x_0) = \frac{1}{2}\|f(x) - x_0\|^2.$$

On considère alors δ tel que $B(x_0, \delta) \subset W$, et on applique le changement de variable $y = f(x)$ pour $x \in f^{-1}(B(x_0, \delta))$:

$$\int_{f^{-1}(B(x_0,\delta))} a(x)\mathrm{e}^{-t\varphi(x)}\,\mathrm{d}x = \mathrm{e}^{-t\varphi(x_0)}\int_{B(x_0,\delta)} a(f^{-1}(y))\mathrm{e}^{-\frac{t\|y-x_0\|^2}{2}}\frac{\mathrm{d}y}{|\mathrm{jac}f(f^{-1}(y))|}.$$

Le changement de variable $z = \sqrt{t}(y - x_0)$ donne alors

$$\int_{f^{-1}(B(x_0,\delta))} a(x)\mathrm{e}^{-t\varphi(x)}\,\mathrm{d}x = \mathrm{e}^{-t\varphi(x_0)}t^{-d/2}\int_{B(x_0,\sqrt{t}\delta)} a(f^{-1}(x_0+z/\sqrt{t}))\mathrm{e}^{-\frac{\|z\|^2}{2}}\frac{\mathrm{d}y}{|\mathrm{jac}f(f^{-1}(x_0+\frac{z}{\sqrt{t}}))|}.$$

Cette intégrale se traite alors par convergence dominée, la domination découle de la régularité de f et de son inverse. On a alors l'équivalent voulu d'après le calcul de l'intégrale de Gauss et parce que $|\mathrm{d}f(x_0)| = \sqrt{\det(\mathrm{H}f(x_0))}$ (*cf.* preuve du lemme de Morse). Pour le reste de l'intégrale, il suffit d'écrire $\varphi(x) \geq \varphi(x_0) + \varepsilon$ et utiliser le fait que a est intégrable. □

Chapitre 92

Indécomposabilité de la loi de Poisson

On montre ici le fait que si une somme de deux variables indépendantes à valeurs dans $\mathbb{N}$ suit une loi de Poisson, chacune des variables suit une loi de Poisson. Si on voit cela en termes d'événements rares, cela veut dire qu'un événement moins rare qu'une loi de Poisson doit être une loi de Poisson. Cela peut aussi se voir avec l'inégalité de Le Cam.

Définition. On définit la loi de Poisson de paramètre $\lambda > 0$ sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par :

$$\mathcal{P}(\lambda) = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!}\delta_n.$$

C'est une loi qui modélise des événements rares. Elle admet des moments à tout ordre et sa variance et son espérance sont toutes deux égales à λ. Plus généralement, sa série génératrice est entière sur $\mathbb{C}$ et vaut :

$$G_{\mathcal{P}(\lambda)}(z) = \mathrm{e}^{\lambda(z-1)} \quad \forall z \in \mathbb{C}.$$

Remarque. La convolution de deux lois de Poisson de paramètres respectifs λ et μ est égale à la loi de Poisson de paramètre $(\lambda+\mu)$. L'objectif de ce développement est de montrer la réciproque de ce résultat.

Théorème. Si $\mathbb{P}_1$ et $\mathbb{P}_2$ sont deux mesures de probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ telles que $\mathbb{P}_1 * \mathbb{P}_2 = \mathcal{P}(\lambda)$ pour un $\lambda > 0$, alors $\mathbb{P}_1$ et $\mathbb{P}_2$ sont des lois de Poisson. Plus simplement, si X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes dont la somme suit une loi de Poisson, alors X et Y suivent une loi de Poisson.

Démonstration. On considère donc deux variables X et Y comme dans l'énoncé et leurs séries génératrices respectives, que l'on note f et g. *A priori*, leur rayon de convergence est plus grand que 1 :

$$f(z) := \sum_{n=0}^{\infty} p_n z^n, g(z) := \sum_{n=0}^{\infty} q_n z^n \quad \forall z \in D(0,1).$$

La première étape consiste à montrer que f et g sont en fait entières. Si λ est le paramètre de la loi (de Poisson) de $Z = X + Y$, alors par indépendance de X et Y :

$$f(z)g(z) = \mathrm{e}^{\lambda(z-1)} \quad \forall z \in D(0,1).$$

On a donc, par unicité du développement en série entière :

$$\forall n \in \mathbb{N}, \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \sum_{k=0}^{n} p_k q_{n-k} \geq \max(p_n q_0, p_0 q_n).$$

En particulier, $p_n \leq \mathrm{e}^{-\lambda}\frac{\lambda^n}{n! q_0}$ et idem pour q_n donc l'identité précédente est valables sur $\mathbb{C}$, *i.e.* f et g sont entières. Maintenant, on chercher des majorations fortes de f et g. On considère pour cela $z \in \mathbb{C}^*$ et puisque les coefficients dans le développement en série entière de f et g sont positifs, on a :

$$|f(z)| \leq f(|z|) \text{ donc } f(|z|) \leq q_0^{-1} f(|z|) g(|z|) = \mathrm{e}^{\lambda(|z|-1)}.$$

En faisant de même pour g, on obtient l'existence d'une constante $C > 0$ telle que $\max(|f(z)|, g(|z|)) \leq C\mathrm{e}^{\lambda(|z|-1)}$. Par ailleurs, les fonctions f et g étant entières et sans zéro (sinon fg en aurait un) donc elles s'écrivent e^F et e^G, où F et G sont entières (*cf. infra*). On a donc $\max(\mathrm{e}^{\mathrm{Re}(F(z))}, \mathrm{e}^{\mathrm{Re}(G(z))}) \leq C\mathrm{e}^{\lambda(|z|-1)}$. On peut donc passer au logarithme (réel) et appliquer le lemme de la partie réelle d'Hadamard (*cf. infra*), qui nous donne $F(z) = \alpha z + a$ et $G(z) = \beta z + b$ où α, β, a, b sont des réels. Puisque $f(1) = g(1) = 1$, on peut écrire $f(z) = \mathrm{e}^{\alpha(z-1)}$ donc $X \sim \mathcal{P}(\alpha)$ et de même $Y \sim \mathcal{P}(\beta)$. □

Montrons maintenant les deux résultats que l'on a utilisés dans cette preuve.

Fait. Soit f un fonction entière qui ne s'annule pas sur $\mathbb{C}$, alors f s'écrit e^F avec F entière.

Démonstration. On peut considérer F une primitive de la fonction $\frac{f'}{f}$, qui est entière car $\frac{f'}{f}$ l'est puisque f ne s'annule pas sur $\mathbb{C}$. On peut alors dériver $\frac{\mathrm{e}^F}{f}$:

$$\left(\frac{\mathrm{e}^F}{f}\right)' = \frac{(f'/f)\mathrm{e}^F f - \mathrm{e}^F f'}{f^2} = 0.$$

Ainsi les deux fonctions e^F et f sont égales quitte à ajuster la valeur de F en un point. □

Théorème. (de la partie réelle d'Hadamard)
Soit $f = u + \mathrm{i}v = \sum_{a_n} z^n$ une fonction entière et

$$A(r) := \sup_{|z|=r} u(z).$$

Alors pour $n \geq 1$ et $r > 0$, $|a_n| \leq 2\frac{A(r)-u(0)}{r^n}$. Si de plus $A(r) = O_{+\infty}(r^d)$ pour $d \in \mathbb{N}$ alors f est un polynôme de degré inférieur ou égal à d.

Démonstration. L'idée est d'avoir une formule de Cauchy pour u. On part donc des identités :

$$a_n r^n = \frac{1}{2\pi}\int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{-\mathrm{i}n\theta}\,\mathrm{d}\theta \text{ et } 0 = \frac{1}{2\pi}\int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{\mathrm{i}n\theta}\,\mathrm{d}\theta, \quad n \geq 1.$$

En conjuguant la deuxième que l'on ajoute à la première, on obtient :

$$a_n r^n = \frac{1}{\pi}\int_0^{2\pi} u(r\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta.$$

On peut alors faire apparaître artificiellement $A(r)$:

$$a_n r^n = \frac{1}{\pi} \int_0^{2\pi} (u(r\mathrm{e}^{\mathrm{i}\theta}) - A(r))\mathrm{e}^{-\mathrm{i}n\theta} \, \mathrm{d}\theta,$$

puis on prend le module en se rappelant que u et A sont des fonctions à valeurs réelles et que la deuxième est plus grande que la première :

$$|a_n| r^n \leq \frac{1}{\pi} \int_0^{2\pi} (A(r) - u(r\mathrm{e}^{\mathrm{i}\theta})) \, \mathrm{d}\theta.$$

On conclut donc puisque $a_0 = f(0) = \frac{1}{2\pi} \int_0^{2\pi} f(r\mathrm{e}^{\mathrm{i}\theta}) \, \mathrm{d}\theta$. On obtient ensuite le résultat en faisant tendre le rayon vers l'infini.

Référence : Hervé et Martine Queffélec, *Analyse complexe*.

Chapitre 93

Théorème de Fejér

Dans ce développement, on montre le théorème de Fejér classique, *i.e.* l'approximation en norme uniforme d'une fonction par les moyennes de Cesàro des sommes partielles de Fourier, puis le théorème de Fejér-Lebesgue, qui est quant à lui une version L^p de cette approximation.

Définition. Pour $N \in \mathbb{N}^*$, on définit le noyau de Dirichlet :

$$D_N(t) = \sum_{n=-N}^{N} \mathrm{e}^{\mathrm{i}nt} = \frac{\sin((N+1/2)t)}{\sin(t/2)}$$

ainsi que le noyau de Fejér :

$$K_N(t) = \frac{1}{N}\sum_{n=0}^{N-1} D_n(t) = \frac{\sin(Nt/2)^2}{N\sin(t/2)^2}.$$

Ainsi, les sommes partielles de Fourier s'écrivent : $S_N(f) = (D_N * f)$, et leur moyenne de Cesàro s'écrit $(K_N * f)$. On montre facilement que K_N est une approximation de l'unité sur $[0, 2\pi]$ (en fait, elle converge au sens des distributions vers $\delta_{\mathbb{Z}}$).

Théorème. (de Fejér)
Soit f une fonction continue et 2π-périodique, alors $(K_N * f)_N$ converge uniformément vers f.

Démonstration. On note, pour $\delta > 0$,

$$\omega(\delta) = \sup\{|f(x) - f(y)|, |x - y| \leq \delta\}.$$

On a alors, pour $\delta > 0$ et $x \in \mathbb{R}$:

$$\begin{aligned}
|(K_N * f)(x) - f(x)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x-t) - f(x))K_N(t)\,\mathrm{d}t\right| \text{ car l'intégrale du noyau vaut 1.} \\
&\leq \frac{1}{2\pi}\int_{|t|<\delta} |f(x-t) - f(x)|K_N(t)\,\mathrm{d}t + \frac{1}{2\pi}\int_{|t|>\delta} |f(x-t) - f(x)|K_N(t)\,\mathrm{d}t \\
&\leq \frac{\omega(\delta)}{2\pi} + \frac{2\|f\|_\infty}{2\pi}\int_{|t|>\delta} K_N(t)\,\mathrm{d}t
\end{aligned}$$

Or pour $\delta < |t| \leq \pi$, $\sin(\delta/2) \leq \sin(|t|/2) \leq 1$ donc $K_N(t) \leq \frac{1}{N \sin(\delta/2)^2}$, ce qui donne finalement :

$$|(K_N * f)(x) - f(x)| \leq \omega(\delta) + \frac{2\|f\|_\infty}{N \sin(\delta/2)^2}.$$

En prenant la borne supérieur en x, il vient :

$$\|K_N * f - f\|_\infty \leq \omega(\delta) + \frac{2\|f\|_\infty}{N \sin(\delta/2)^2},$$

puis quand N tend vers l'infini :

$$\limsup_{N \to +\infty} \|K_N * f - f\|_\infty \leq \omega(\delta).$$

Ceci valant pour tout δ, on peut le faire tendre vers 0 ; puisque f est continue et 2π-périodique, elle est uniformément continue par le théorème de Heine, ce qui implique que $\omega(\delta)$ tend vers 0 avec δ. Somme toute, on a bien le résultat voulu. □

On passe maintenant au cas L^p, $p \in [1, \infty[$. Il existe un résultat plus général qui entraîne le théorème de Fejér-Lebesgue.

Théorème. Soit $(\rho_n)_n$ une approximation de l'unité, et $f \in \mathrm{L}^p_{2\pi}$. Alors, on a la convergence suivante dans $\mathrm{L}^p_{2\pi}$:

$$\rho_n * f \xrightarrow[n \to \infty]{} f.$$

Démonstration. Soit $p \in [1, +\infty[$. Pour presque tout $x \in \mathbb{R}$,

$$\begin{aligned}
|\rho_n * f(x) - f(x)| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) |f(x) - f(x-y)| \, \mathrm{d}y \\
&\leq \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) |f(x-y) - f(x)|^p \, \mathrm{d}y \right)^{1/p} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) \, \mathrm{d}y \right)^{1/q} \\
&= \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) |f(x-y) - f(x)|^p \, \mathrm{d}y \right)^{1/p}.
\end{aligned}$$

Où l'on a utilisé l'inégalité de Hölder. Ce calcul est utile pour majorer la norme suivante :

$$\begin{aligned}
\|\rho_n * f - f\|_p^p &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) |f(x-y) - f(x)|^p \, \mathrm{d}y \right) \mathrm{d}x \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) |f(x-y) - f(x)|^p \, \mathrm{d}x \right) \mathrm{d}y \text{ par Fubini} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_n(y) \|\tau_y f - f\|_p^p \, \mathrm{d}y
\end{aligned}$$

Par continuité des translations dans L^p, il existe, pour $\varepsilon > 0$ fixé, un $\delta > 0$ tel que pour $|y| \leq \delta$, la norme $\|\tau_y f - f\|_p$ soit majorée par ε. On coupe l'intégrale au niveau de δ comme précédemment. On peut alors majorer l'intégrale au voisinage de 0 par ε puisque l'intégrale de l'approximation de l'unité vaut 1, et pour l'autre intégrale on écrit $\|\tau_y f - f\|_p \leq 2\|f\|_p$ pour se débarrasser du y puis on prend n suffisamment grand pour que l'intégrale de l'approximation soit assez petite (ce

n existe par définition d'une approximation de l'unité). Somme toute, on a bien la convergence vers 0 de la norme, donc la convergence voulue. □

Remarque. La continuité des translations permet d'utiliser le premier théorème pour montrer le deuxième.

Corollaire. (Fejér-Lebesgue)
Puisque le noyau de Fejér est une approximation de l'unité, les sommes de Fejér convergent dans L^p vers la fonction considérée. En particulier, les polynômes trigonométriques sont denses dans L^p.

Remarque. Le premier théorème de Fejér énoncé est fondamental, puisqu'il permet de savoir que si la série de Fourier de f converge simplement en un point t, c'est forcément vers $f(t)$ d'après le lemme de Cesàro. Ainsi, en ajoutant l'hypothèse $\mathcal{C}^1$ par morceaux sur f, on obtient que f est somme de sa série de Fourier.

Par ailleurs, le théorème de Fejér ne risque pas d'être vrai dans L^∞, puisque les polynômes trigonométriques sont continus et qu'une limite uniforme de fonctions continues est continue.

Référence : Claude Zuily et Hervé Queffélec, *Analyse pour l'agrégation.*

Chapitre 94

Théorème de Banach-Steinhaus et séries de Fourier

On s'intéresse dans ce développement au théorème de Banach-Steinhaus et à une de ses applications à la théorie des séries de Fourier. On passe par le lemme de Baire, qui possède de très nombreuses autres applications.

Théorème. (de Baire)
Soit (E, d) un espace métrique complet. Toute intersection d'ouverts denses est dense.

Démonstration. Soit $\Omega = \bigcap_{n\in\mathbb{N}^*} O_n$ une intersection d'ouverts denses de E. Pour ω un ouvert de E, montrons que $\Omega \cap \omega$ est non vide. Comme O_1 est dense, $\omega \cap O_1$ contient une boule fermée $\overline{B_1} := \overline{\mathrm{B}}(x_1, r_1)$, et on peut supposer $0 < r_1 < 1$. Comme O_2 est dense, la boule ouverte B_1 rencontre O_2, donc rencontre une boule fermée $\overline{B_2}$, de centre x_2 et de rayon $r_2 < \frac{1}{2}$. Par récurrence, on a des boules ouvertes emboîtées (B_n) telles que

$$B_n \subset \overline{B_n} = \overline{B}(x_n, r_n) \subset B_{n-1} \cap O_n \subset B_{n-1}.$$

avec $r_n < \frac{1}{n}$. Alors la suite (x_n) est de Cauchy car $d(x_n, x_{n+p}) < \frac{1}{n}$ (puisque $x_{n+p} \in B_n$). À la limite, on a alors $x \in \overline{B_n}$ pour tout n donc $x \in \Omega$ et de plus $x \in \overline{B_1} \subset \omega$ donc $x \in \Omega \cap \omega$. □

Application. Une fonction $f : \mathbb{R} \to \mathbb{R}$ ne peut pas être continue en les rationnels et discontinue en les irrationnels. Cependant, l'inverse est possible (fonction de Thomae).

Remarque. Un formulation équivalente du théorème de Baire est que toute union dénombrable de fermés d'intérieur vide est d'intérieur vide.

Théorème. (de Banach-Steinhaus)
Soit $(E, \|\cdot\|_E)$ un espace de Banach et $(F, \|\cdot\|_F)$ un espace vectoriel normé. Soit $(T_i)_{i\in I}$ une famille (pas forcément dénombrable) d'applications linéaires et continues. On suppose que pour tout x, la famille $(T_i(x))_i$ est bornée dans F. Alors la famille (T_i) est bornée en norme d'opérateur.

Démonstration. Pour $(n, i) \in \mathbb{N} \times I$, on pose $X_{n,i} = \{x \in E, \|T_i(x)\|_F \leq n\}$. C'est un fermé de E comme image réciproque d'un fermé. Ainsi, $Y_n := \bigcap_{i\in I} X_{n,i}$ est un fermé de E comme intersection de fermés. Comme E est complet et $\bigcup_{n\in\mathbb{N}} Y_n = E$ est d'intérieur non vide, il existe

par le théorème de Baire un n_0 tel que Y_{n_0} est d'intérieur non vide. Soit alors $x_0 \in E$ et $r > 0$ tels que $\mathrm{B}_E(x_0, r) \subset Y_{n_0}$, c'est-à-dire :

$$\forall z \in \mathrm{B}_E(0,1), \forall i \in I, \|T_i(x_0 + rz)\|_F \leq n_0.$$

Par linéarité et inégalité triangulaire, il vient :

$$\begin{aligned}\forall z \in \mathrm{B}_E(0,1), \forall i \in I, \|T_i(z)\|_F &= \left\| \frac{1}{r} \left(T_i(x_0 + rz) - T_i(x_0)\right) \right\|_F \\ &\leq \frac{\|T_i(x_0 + rz)\|_F + \|T_i(x_0)\|_F}{r} \\ &\leq \frac{2n_0}{r}.\end{aligned}$$

On a donc $\sup_i \|T_i\|_{E,F} \leq \frac{2n_0}{r} < \infty$. □

Remarque. En adaptant simplement la preuve, on peut montrer que dans le cas où l'un des ouvert $\bigcup_{i\in I}\{x \in E, \|T_i(x)\| > n\}$ n'est pas dense, alors on a la même borne qui donne $\sup \|T_i\| < \infty$, ce qui donne par contraposée et par Baire que $\sup_i \|T_i\|$ est infini sur un ensemble dense.

Corollaire. Sous les mêmes hypothèses, si (T_n) converge simplement vers T, alors $(\|T_n\|)$ est bornée, T est linéaire continue et la norme de T est plus petite que la limite inférieure des normes des T_n.

Application. Il existe une fonction continue 2π-périodique dont la série de Fourier ne converge pas en $t = 0$.

Démonstration. On munit l'espace E des fonctions continues 2π-périodiques de la norme infinie, qui en fait bien un Banach. Pour tout N, on définit la forme linéaire $\Lambda_N : f \mapsto (f * D_N)(0)$, où D_N est le noyau de Dirichlet. Pour tout N, Λ_N est une forme linéaire continue de norme $\|D_N\|_1$ (en effet, des majorations simples donnent l'inégalité, et pour l'égalité on approche la fonction "signe de D_N" par des fonctions continues).

Ensuite, on montre que $\|D_N\|_1 \to +\infty$. Ainsi, par le corollaire précédent, Λ_N ne peut pas converger simplement. Il existe une fonction dont la série de Fourier ne converge pas en 0. La remarque suivant la preuve de Banach-Steinhaus montre alors même que ces fonctions forment un ensemble dense. □

Référence : Cours de Karine Beauchard, éventuellement trouvable sur son site web. Sinon, X. Gourdon doit faire des preuves de Baire et Banach-Steinhaus dans son livre d'analyse.

Chapitre 95

Théorème de Liapounov

Dans ce développement, on montre le théorème de linéarisation de Lyapounov, qui permet de conclure quant à la stabilité d'un système en étudiant son système linéarisé.

Théorème. (de Lyapounov)
Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une fonction de classe $\mathcal{C}^1$ telle que $f(0) = 0$. On suppose que les valeurs propres de $A := \mathrm{d}f(0)$ ont toutes une partie réelle strictement négative. Alors 0 est un point d'équilibre asymptotiquement stable de l'équation $y' = f(y)$.

Démonstration. Par le théorème de Jordan, on sait que 0 est un point d'équilibre asymptotiquement stable pour me système linéarisé $z' = Az$. On considère le produit scalaire suivant sur $\mathbb{R}^n$:

$$b(x, y) := \int_0^\infty \langle \mathrm{e}^{tA}x, \mathrm{e}^{tA}y\rangle \,\mathrm{d}t.$$

Il est bien défini par l'inégalité de Cauchy-Schwarz et l'estimation de la norme de e^{tA} qui peut être obtenue par Jordan encore une fois. On note q la forme quadratique associée. Puisque c'est une forme quadratique, on a pour tout $x \in \mathbb{R}^n$ et tout $h \in \mathbb{R}^n$:

$$\mathrm{d}q(x)(h) = 2b(x, h).$$

Remarquons qu'en particulier

$$\mathrm{d}q(x)(Ax) = 2b(x, Ax) = 2\int_0^\infty \langle \mathrm{e}^{tA}x, \mathrm{e}^{tA}Ax\rangle \,\mathrm{d}t = -\|x\|^2.$$

Soit maintenant (y, I) l'unique solution maximale à notre problème de Cauchy pour la condition initiale $y(0) = y_0$. Par composition, si on note $r(y) = f(y) - Ay$:

$$\begin{aligned}\frac{\mathrm{d}q(y(s))}{\mathrm{d}s} &= \mathrm{d}q(y(s))(y'(s)) \\ &= 2b(y(s), Ay(s)) + 2b(y(s), r(y(s))) \\ &= -\|y(s)\|^2 + 2b(y(s), r(y(s))) \\ &\leq -cq(y(s)) + 2\sqrt{q(y(s))}\sqrt{q(r(y(s)))}\end{aligned}$$

Par équivalence des normes et Cauchy-Schwarz. Par définition de la différentielle, $r(y(s)) =$

$\mathrm{o}(\|y(s)\|)$ donc

$$\frac{\mathrm{d}q(y(s))}{\mathrm{d}s} = -cq(y(s)) + \mathrm{o}(q(y(s))).$$

Ainsi, il existe une constante $\beta > 0$ et $\alpha > 0$ tels que $q(y(s)) \leq \alpha$ entraîne

$$\frac{\mathrm{d}q(y(s))}{\mathrm{d}s} \leq -\beta q(y(s)) < 0.$$

Supposons que $q(y_0) < \alpha$. On considère alors $\{t > 0, q(y(t)) \leq \alpha\}$. C'est un ensemble fermé par construction et ouvert puisque si t_0 est dedans alors la dérivée de $q \circ y$ en ce point va être strictement négative donc la fonction va être décroissante sur un petit voisinage. Cet ensemble est donc connexe donc c'est tout I. Ainsi les solutions sont dans la boule fermée de centre 0 et de rayon $q(y_0)$ donc par théorème des bouts les solutions sont globales et alors pour tout $t \geq 0$:

$$q(y(t)) \leq \mathrm{e}^{-\beta t} q(y_0).$$

On peut alors conclure par équivalence des normes. □

Référence : Rouvière, *petit guide du calcul différentiel*

Chapitre 96

Théorème limite central et extensions aux densités

Dans ce développement, on montre le théorème limite central, fondamental en probabilités et en statistiques, et on montre comment il peut s'étendre aux densités. Enfin, on le raffine avec la connaissance des premiers moments de notre mesure.

Lemme. Pour $n \in \mathbb{N}$ et $t > 0$:

$$\left| \mathrm{e}^{\mathrm{i}t} - 1 - \mathrm{i}t - \cdots - \frac{(\mathrm{i}t)^{n-1}}{(n-1)!} \right| \leq \frac{t^n}{n!}.$$

Démonstration. C'est l'inégalité de Taylor-Lagrange. □

Lemme. Si X admet un moment d'ordre n alors sa fonction caractéristique φ est de classe $\mathcal{C}^n$, et

$$\varphi^{(n)}(\xi) = \mathrm{i}^n \int_{\mathbb{R}} x^n \mathrm{e}^{\mathrm{i}\xi x} \,\mathrm{d}\,\mathbb{P}_X(x).$$

Démonstration. Par théorème de dérivation sous l'intégrale. □

Remarque. Lorsque $\xi = 0$, on retrouve les moments (à une constante près).

Corollaire. Avec les mêmes hypothèses et notations, on a le développement de Taylor suivant :

$$\left| \varphi(\xi + t) - \varphi(\xi) - \sum_{k=1}^{n-1} \frac{t^k}{k!} \varphi^{(k)}(\xi) \right| \leq M_n \frac{|t|^n}{n!}.$$

Théorème. (théorème limite central)
Soit (X_i) une suite i.i.d de variables aléatoires centrées et réduites (*a fortiori* dans L^2), et soit $\overline{X_n} := \frac{X_1 + \ldots + X_n}{n}$. Alors $\sqrt{n}\overline{X_n}$ converge en loi vers la loi normale centrée réduite.

Démonstration. D'après le corollaire, on a un développement limité de la fonction caractéristique φ des X_i :

$$\varphi(x) = \varphi(0) + x\varphi'(0) + \frac{x^2}{2}\varphi''(0) + \mathrm{o}(x^2).$$

En prenant $x = \frac{\xi}{\sqrt{n}}$ et en prenant à la puissance n :

$$\varphi_{\sqrt{n}\overline{X_n}}(\xi) = \left(1 - \frac{1}{2n}\xi^2 + \mathrm{o}(1/n)\right)^n \xrightarrow[x\to\infty]{} \mathrm{e}^{-\frac{1}{2}\xi^2}.$$

On reconnaît la fonction caractéristique de la gaussienne et on conclut donc avec le théorème de Lévy. □

Théorème. (théorème limite central pour les densités)
On suppose que X_1 est à densité, de densité f. Si φ est intégrable alors la suite des densités f_n de $\sqrt{n}\overline{X_n}$ converge uniformément vers la densité de la loi normale centrée réduite, notée g.

Démonstration. Par inversion de Fourier,

$$f_n(x) - g(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \left(\varphi\left(\frac{\xi}{\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2}\right)\mathrm{e}^{\mathrm{i}\xi x}\,\mathrm{d}\xi.$$

Soit :

$$|f_n(x) - g(x)| \leq \frac{1}{2\pi}\int_{\mathbb{R}} \left|\varphi\left(\frac{\xi}{\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2}\right|\,\mathrm{d}\xi.$$

On utilise ici un fait général, pour majorer la norme infinie on peut passer en Fourier et cela revient à majorer la norme 1 de la transformée de Fourier. D'après le théorème limite central classique, on sait que l'intégrande en question converge simplement vers 0. Cependant, on n'a *a priori* pas de domination simple pour appliquer le théorème de convergence dominée, car ça coince à l'infini. On va pallier ce problème en découpant $\mathbb{R}$. Soit $a > 0$, à calibrer éventuellement plus tard ; l'intégrale

$$\int_{-a}^{a} \left|\varphi\left(\frac{\xi}{\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2}\right|\,\mathrm{d}\xi$$

converge vers 0 quand n tend vers l'infini par convergence dominée. Pour le reste de l'intégrale, il y a deux sous-cas. En effet, quand ξ va être très grand, on s'attend à ce que l'intégrande soit petite et pour ξ de l'ordre de $\sqrt{n}$, on s'attend à pouvoir contrôler φ au vu de son développement limité au voisinage de 0. Mettons tout cela en forme ; d'après le développement de φ, il existe $\delta > 0$ tel que $|\varphi(\xi)| \leq \mathrm{e}^{-\frac{1}{4}\xi^2}$ dès que $|\xi| < \delta$. Pour $a < |\xi| < \delta\sqrt{n}$:

$$\left|\varphi\left(\frac{\xi}{\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2}\right| \leq (\mathrm{e}^{-\frac{1}{4n}\xi^2})^n + \mathrm{e}^{-\frac{1}{2}\xi^2} \leq 2\mathrm{e}^{-\frac{\xi^2}{4}}$$

pour tout n. En prenant a assez grand (et indépendamment de n), cette borne peut être rendue plus petite que ε fixé *a priori*. Enfin,

$$\int_{|\xi|>\delta\sqrt{n}} \left|\varphi\left(\frac{\xi}{\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2}\right|\,\mathrm{d}\xi \leq \left(\sup_{|\xi|\geq\delta}|\varphi|\right)^{n-1}\int_{\mathbb{R}}\left|\varphi\left(\frac{\xi}{\sqrt{n}}\right)\right|\,\mathrm{d}\xi + \int_{|\xi|>\delta\sqrt{n}} \mathrm{e}^{-\frac{1}{2}\xi^2}\,\mathrm{d}\xi.$$

D'après le lemme de Riemann-Lebesgue, le suprémum peut être rendu strictement inférieur à 1 pour δ assez grand. L'intégrale à côté se calcule par changement de variable et vaut $\sqrt{n}\int_{\mathbb{R}}|\varphi|$. Par croissances comparées, ce premier terme tend vers 0. Le deuxième terme également comme reste d'une intégrale convergente. □

Remarque. On peut affaiblir les hypothèses, φ peut n'être intégrable qu'à partir d'une certaine puissance.

Théorème. (premier terme du développement d'Edgeworth)
Soit (X_i) une suite i.i.d de variables à densité centrées de variance σ^2, et avec un moment d'ordre trois μ_3. On suppose que leur fonction caractéristique commune φ est intégrable. Alors

$$\|f_n - g - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3-3x)g\|_\infty = \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

où f_n est la densité de $\sqrt{n}\frac{\overline{X_n}}{\sigma}$.

Démonstration. Au début, on procède comme précédemment, on passe en Fourier et on cherche à dominer l'intégrande :

$$\left|f_n(x) - g(x) - \frac{\mu_3}{6\sigma^3\sqrt{n}}(x^3-3x)g(x)\right| \leq \frac{1}{2\pi}\int_{\mathbb{R}}\left|\varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2} - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\mathrm{e}^{-\frac{1}{2}\xi^2}\right| \mathrm{d}\xi.$$

Comme dans le cas précédent, par Riemann-Lebesgue, l'intégrale de notre intégrande pour $|\xi| > \sigma\delta\sqrt{n}$ (où $\delta > 0$ est fixé) tend vers 0 exponentiellement vite quand n tend vers l'infini. On peut donc écrire notre intégrale comme

$$\int_{\mathbb{R}}\left|\varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right)^n - \mathrm{e}^{-\frac{1}{2}\xi^2} - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\mathrm{e}^{-\frac{1}{2}\xi^2}\right| \mathrm{d}\xi =$$
$$\int_{|\xi|<\sigma\delta\sqrt{n}} \mathrm{e}^{-\frac{1}{2}\xi^2}\left|\mathrm{e}^{n\psi(\xi/\sigma\sqrt{n})} - 1 - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\right| \mathrm{d}\xi + \mathrm{o}(1/n), \quad \text{où } \psi(\xi) := \log(\varphi(\xi)) + \frac{1}{2}\sigma^2\xi^2.$$

Notons qu'il est légitime de prendre un logarithme complexe puisque l'argument en question est dans le cercle de centre 1 et de rayon 1. On cherche donc à majorer ce que l'on a dans le module. Pour cela, on a à comparer une exponentielle et une quantité qui ressemble à son développement limité. On voudrait une inégalité qui permet de découpler l'exponentielle et son argument.

Lemme. Soient α et β deux complexes. Alors $|\mathrm{e}^\alpha - 1 - \beta| \leq (|\alpha-\beta| + \frac{1}{2}|\beta|^2)\mathrm{e}^{\max(|\alpha|,|\beta|)}$.

Démonstration du lemme. On écrit $|\mathrm{e}^\alpha - 1 - \beta| \leq |\mathrm{e}^\alpha - \mathrm{e}^\beta| + |\mathrm{e}^\beta - 1 - \beta|$. Le premier terme se majore comme suit :

$$\begin{aligned}
|\mathrm{e}^\alpha - \mathrm{e}^\beta| &\leq \sum_{k=1}^{\infty}\frac{|\alpha^k - \beta^k|}{k!} \\
&\leq \sum_{k=1}^{\infty}\frac{|\alpha-\beta|}{k!}\sum_{i=0}^{k-1}|\alpha^{k-1-i}\beta^i| \\
&\leq |\alpha-\beta|\sum_{k=1}^{\infty}\frac{\max(|\alpha|,|\beta|)^{k-1}}{(k-1)!} \\
&\leq |\alpha-\beta|\mathrm{e}^{\max(|\alpha|,|\beta|)}.
\end{aligned}$$

Et le deuxième de la façon suivante :

$$|\mathrm{e}^\beta - 1 - \beta| \leq \sum_{k=2}^{\infty}\frac{|\beta|^k}{k!} \leq \frac{|\beta|^2}{2}\sum_{k=2}^{\infty}\frac{|\beta|^{k-2}}{k!/2} \leq \frac{|\beta|^2}{2}\sum_{k=0}^{\infty}\frac{|\beta|^k}{(k+2)!/2} \leq \frac{|\beta|^2}{2}\mathrm{e}^{|\beta|},$$

puisque pour k entier, $\frac{(k+2)!}{2} \geq k!$.

Ceci étant fait, on peut revenir à la preuve du théorème ; le lemme donne (pour $|\xi| < \delta\sigma\sqrt{n}$) :

$$\left|\mathrm{e}^{n\psi(\xi/\sigma\sqrt{n})} - 1 - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\right| \leq \left(\left|n\psi(\xi/\sigma\sqrt{n}) - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\right| + \frac{\mu_3^2\xi^6}{72\,n\sigma^6}\right)\mathrm{e}^{\gamma},$$

avec γ un majorant des modules de nos α et β. Il nous reste donc à estimer $\psi(\xi) - \frac{\mu_3}{6}(\mathrm{i}\xi)^3$ au voisinage de 0. Or la fonction ψ est de classe $\mathcal{C}^3$ au voisinage de 0 (car φ l'est et log est une série entière). En composant les développements limités au voisinage de 0, on voit que $\psi(0) = \psi'(0) = \psi''(0)$ et que $\psi'''(0) = \mathrm{i}^3\mu_3$. Soit $\varepsilon > 0$, d'après la formule de Taylor-Young, il existe $\delta > 0$ tel que :

$$|\psi(\xi) - \frac{\mu_3}{6}(\mathrm{i}\xi)^3| \leq \varepsilon\sigma^3|\xi|^3 \quad \forall|\xi| < \delta.$$

Par ailleurs, on peut aussi calibrer δ assez petit pour que chacun des termes soit petit en module (on cherche notre γ) :

$$|\psi(\xi)| < \frac{1}{4}\sigma^2\xi^2, |\frac{\mu_3}{6}(\mathrm{i}\xi^3)| < \frac{1}{4}\sigma^2\xi^2 \quad \forall|\xi| < \delta.$$

On a alors, en prenant $\gamma = \frac{1}{4}\sigma^2\xi^2$:

$$\begin{aligned}
\left|\mathrm{e}^{n\psi(\xi/\sigma\sqrt{n})} - 1 - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\right| &\leq \left(\left|n\psi(\xi/\sigma\sqrt{n}) - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\right| + \frac{\mu_3^2\xi^6}{72\,n\sigma^6}\right)\mathrm{e}^{\frac{1}{4}\xi^2} \\
&\leq \left(n\left|\psi\left(\frac{\xi}{\sigma\sqrt{n}}\right) - \frac{\mu_3}{6}\left(\frac{\mathrm{i}\xi}{\sigma\sqrt{n}}\right)^3\right| + \frac{\mu_3^2\xi^6}{72\,n\sigma^6}\right)\mathrm{e}^{\frac{1}{4}\xi^2} \\
&\leq \left(\varepsilon\frac{|\xi|^3}{\sqrt{n}} + \frac{\mu_3^2\xi^6}{72\,n\sigma^6}\right)\mathrm{e}^{\frac{1}{4}\xi^2}.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\int_{|\xi|<\sigma\delta\sqrt{n}} \mathrm{e}^{-\frac{1}{2}\xi^2}\left|\mathrm{e}^{n\psi(\xi/\sigma\sqrt{n})} - 1 - \frac{\mu_3}{6\sigma^3\sqrt{n}}(\mathrm{i}\xi)^3\right|\mathrm{d}\xi &\leq \int_{|\xi|<\sigma\delta\sqrt{n}} \mathrm{e}^{-\frac{1}{4}\xi^2}\left(\varepsilon\frac{|\xi|^3}{\sqrt{n}} + \frac{\mu_3^2\xi^6}{72\,n\sigma^6}\right)\mathrm{d}\xi \\
&\leq \int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{4}\xi^2}\left(\varepsilon\frac{|\xi|^3}{\sqrt{n}} + \frac{\mu_3^2\xi^6}{72\,n\sigma^6}\right)\mathrm{d}\xi \\
&\leq \frac{\varepsilon}{\sqrt{n}}\int_{\mathbb{R}} |\xi|^3\mathrm{e}^{-\frac{1}{4}\xi^2}\,\mathrm{d}\xi + \frac{\mu_3^2}{72\,n\sigma^6}\int_{\mathbb{R}} |\xi|^6\mathrm{e}^{-\frac{1}{4}\xi^2}\,\mathrm{d}\xi \\
&\leq C\frac{\varepsilon}{\sqrt{n}}
\end{aligned}$$

où C est une constante et n assez grand. □

Référence : Feller vol. II

Chapitre 97

Théorème taubérien fort

On montre dans ce document le théorème taubérien fort, qui fournit une réciproque au lemme radial d'Abel.

Théorème. (théorème taubérien fort)
Soit $(a_n)_n \in \mathbb{C}^{\mathbb{N}}$ avec $a_n = \mathrm{O}(1/n)$ et

$$\sum_{n=0}^{\infty} a_n x^n \xrightarrow[x\to 1^-]{} \ell \in \mathbb{C}.$$

Alors la série $\sum_{n\in\mathbb{N}} a_n$ converge et sa somme vaut ℓ.

Démonstration. Quitte à changer a_0 on peut prendre $\ell = 0$. On voudrait écrire la chose suivante :

$$0 = \lim_{x\to 1^-} \sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} a_n \lim_{x\to 1^-} x^n = \sum_{n=0}^{+\infty} a_n,$$

mais aucun des théorèmes généraux ne permet de justifier l'interversion. On sait que pour une fonction polynomiale p nulle en 0, on peut passer à la limite dans $\sum_{n=0}^{\infty} a_n p(x^n)$ quand x tend vers 1. En effet, pour $k \in \mathbb{N}^*$,

$$\lim_{x\to 1^-} \sum_{n=0}^{+\infty} a_n (x^n)^k = \lim_{x\to 1^-} \sum_{n=0}^{+\infty} a_n (x^k)^n = \lim_{x\to 1^-} F(x^k) = 0$$

par composition de limites. On voudrait ce même résultat d'interversion pour une fonction constante égale à 1 au voisinage de 1. En effet, posons

$$\mathcal{F} = \left\{ \varphi : [0,1] \to \mathbb{R}, \sum_{n\in\mathbb{N}} a_n \varphi(x^n) \text{ converge pour tout } x < 1 \text{ et } \lim_{x\to 1^-} \sum_{n=0}^{+\infty} a_n \varphi(x^n) = 0 \right\}$$

et $g := \mathbf{1}_{[1/2,1]}$. Pour $x < 1$,

$$\sum_{n=0}^{\infty} a_n g(x^n) = \sum_{n=0}^{\frac{\ln(2)}{\ln(x)}} a_n.$$

Ainsi, si $g \in \mathcal{F}$, on a le résultat voulu. On a donc d'une part $x\mathbb{R}[x] \subset \mathcal{F}$ et on veut montrer d'autre part que g est dans $\mathcal{F}$. Cherchons donc à approximer g par des fonctions de $x\mathbb{R}[x]$. Pour

savoir comment et pour quelle norme, partons de ce que l'on veut. Si on a deux polynômes $p_1 \leq g \leq p_2$ nuls en 0 :

$$\begin{aligned}\left|\sum_{n=0}^{\infty} a_n g(x^n) - \sum_{n=0}^{\infty} a_n p_1(x^n)\right| &\leq \sum_{n=0}^{\infty} |a_n|(p_2(x^n) - p_1(x^n)) \\ &=: \sum_{n=0}^{\infty} |a_n| p_3(x^n).\end{aligned}$$

Pour que notre idée soit féconde, il faut pouvoir estimer ce genre de quantités lorsque x tend vers 1. Remarquons alors que pour $k \in \mathbb{N}$:

$$(1-x)\sum_{n=1}^{\infty} x^n x^{kn} = (1-x)\sum_{n=1}^{\infty} (x^{k+1})^n = (1-x)\frac{x^{k+1}}{1-x^{k+1}} \xrightarrow[x\to 1^-]{} \frac{1}{k+1} = \int_0^1 u^k \,\mathrm{d}u,$$

donc pour tout polynôme p,

$$(1-x)\sum_{n=1}^{+\infty} x^n p(x^n) \xrightarrow[x\to 1^-]{} \int_0^1 p.$$

On cherchera donc p_3 sous la forme $p_3(x) = x(1-x)q(x)$ avec $q \in \mathbb{R}[x]$, puisqu'alors on pourra appliquer notre hypothèse sur (a_n) pour avoir :

$$\begin{aligned}\sum_{n=0}^{+\infty} |a_n| p_3(x) &\leq \sum_{n=0}^{+\infty} \frac{M}{n} x^n(1-x^n)q(x^n) \\ &\leq M\sum_{n=0}^{+\infty} x^n q(x^n) \xrightarrow[x\to 1^-]{} M\int_0^1 q(t)\,\mathrm{d}t.\end{aligned}$$

Ainsi, on peut clarifier notre objectif ; il nous faut p_1 et p_2 deux polynômes qui encadrent g (comme précédemment) et tels que $\int_0^1 \frac{p_2(t)-p_1(t)}{t(1-t)}\,\mathrm{d}t$ est arbitrairement petit. On aura alors pour $\varepsilon > 0$ fixé :

$$\begin{aligned}\limsup_{x\to 1^-}\left|\sum_{n=0}^{\infty} a_n g(x^n)\right| &= \limsup_{x\to 1^-}\left|\sum_{n=0}^{\infty} a_n g(x^n) - \sum_{n=0}^{\infty} a_n p_1(x^n)\right| \\ &\leq M\int_0^1 \frac{p_2(t)-p_1(t)}{t(1-t)}\,\mathrm{d}t \\ &\leq \varepsilon,\end{aligned}$$

ce qui permettra de conclure. Le problème qui nous reste est alors uniquement un problème d'approximation de fonctions. On cherche p_1 et p_2 de la forme

$$p_i(x) = x + x(1-x)(r_i(x) + (-1)^i\varepsilon), \text{ avec } r_i \in \mathbb{R}[x].$$

Remarquons qu'on a rajouté le facteur x puisque si l'on veut approcher g, on peut s'attendre à ce que $p_i(1)$ soit proche de 1. On a alors

$$\int_0^1 \frac{p_2(t)-p_1(t)}{t(1-t)}\,\mathrm{d}t = \int_0^1 (r_2(t)-r_1(t))\,\mathrm{d}t + 2\varepsilon.$$

On cherche donc à avoir $\int_0^1 (r_2(t) - r_1(t))\,\mathrm{d}t$ petite, avec la condition

$$r_1(x) - \varepsilon \leq \frac{g(x) - x}{x(1-x)} =: h(x) \leq r_2(x) - \varepsilon.$$

Pour cela, on voudrait utiliser le théorème de Weierstrass mais on ne peut pas directement car la fonction h n'est pas continue en $\frac{1}{2}$. On approche donc h par s_1 et s_2 ($s_1 \leq h \leq s_2$) continues telles que $\int_0^1 (s_2 - s_1) \leq \varepsilon$. D'après le théorème de Weierstrass, on a alors r_1 et r_2 tels que $r_2(x) - s_2(x) \leq \varepsilon$ et $s_1(x) - r_1(x) \leq \varepsilon$ pour tout $x \in [0,1]$. On a alors les encadrements voulus, et

$$\int_0^1 (r_2 - r_1) = \int_0^1 (r_2 - s_2) + \int_0^1 (s_2 - s_1) + \int_0^1 (s_1 - r_1) \leq 3\varepsilon,$$

ce qui permet de conclure. □

Référence : Gourdon, *Analyse*

Chapitre 98

Inversion de Fourier pour la fonction caractéristique

La fonction caractéristique est un objet central en probabilités. Elle permet d'effectuer des calculs sur des fonctions régulières plutôt que sur des mesures qui sont plus difficiles à manipuler.

Définition. La *fonction caractéristique* d'une variable aléatoire réelle X est la transformée de Fourier de la mesure $\mathbb{P}_X$:

$$\varphi_X(t) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx} \,\mathrm{d}\,\mathbb{P}_X(x) = \mathbb{E}(\mathrm{e}^{\mathrm{i}tX}).$$

Proposition. (Inversion de Fourier pour les mesures positives)
Pour tous les réels $a < b$, et toute mesure μ positive finie, on a :

$$\frac{1}{2}\mu(\{a,b\}) + \mu(]a,b[) = \lim_{T\to+\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{\mathrm{e}^{-ita} - \mathrm{e}^{-itb}}{\mathrm{i}t}\varphi(t)\,\mathrm{d}t,$$

où φ est la fonction caractéristique associée à la mesure μ.

Ainsi, la fonction caractéristique caractérise la loi.

Démonstration. On note $I_T(x) = \frac{1}{2\pi}\int_{-T}^{T} \frac{\mathrm{e}^{-ita}-\mathrm{e}^{-itb}}{\mathrm{i}t}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t$. Pour $x \notin \{a,b\}$, on a, par changement de variables :

$$I_T(x) = \frac{1}{2\pi}\int_{-T(x-a)}^{T(x-a)} \frac{\sin(u)}{u}\,\mathrm{d}u - \frac{1}{2\pi}\int_{-T(x-b)}^{T(x-b)} \frac{\sin(u)}{u}\,\mathrm{d}u.$$

Mais comme l'intégrale du sinus cardinal est semi-convergente (son intégrale sur $\mathbb{R}$ vaut π), on voit que si $x < a$ ou $x > b$, les deux termes vont se compenser et notre limite sera nulle, et que si $x \in]a,b[$, ils vont s'ajouter et valoir à la limite $\frac{1}{2\pi}(\pi+\pi) = 1$.

Par ailleurs, si $x = a$, on peut faire le calcul à la main :

$$I_T(a) = \frac{1}{2\pi}\int_{-T}^{T} \frac{1-\mathrm{e}^{\mathrm{i}t(a-b)}}{\mathrm{i}t}\,\mathrm{d}t = \frac{1}{2\pi}\left[\int_{-T}^{T} \frac{1-\cos(t(b-a))}{\mathrm{i}t}\,\mathrm{d}t + \int_{-T}^{T} \frac{1-\sin(t(b-a))}{\mathrm{i}t}\,\mathrm{d}t\right].$$

Par parité, l'intégrale de gauche vaut zéro tandis que celle de droite va tendre vers l'intégrale sur

sinus cardinal sur $\mathbb{R}$. Somme toute, on trouve bien $\frac{1}{2}$. On fait de même pour b et on obtient :

$$\lim_{T\to\infty} I_T(x) = \frac{1}{2}\mathbf{1}_{\{a,b\}}(x) + \mathbf{1}_{]a,b[}(x).$$

Ensuite, passons au cas général. La fonction mesurable de deux variables $(x,t) \mapsto \frac{\mathrm{e}^{-\mathrm{i}ta}-\mathrm{e}^{-\mathrm{i}tb}}{\mathrm{i}t}\mathrm{e}^{\mathrm{i}tx}$ est bornée sur $[-T,T]\times\mathbb{R}$. Alors, elle est intégrable contre la mesure finie $\lambda_{[-T,T]}\otimes\mu$. Par le théorème de Fubini :

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-ita}-\mathrm{e}^{-itb}}{\mathrm{i}t}\varphi(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{\mathbb{R}} I_T(x)\,\mathrm{d}\mu(x).$$

Or $(T,x)\mapsto I_T(x)$ est bornée puisque le sinus cardinal est d'intégrale finie et borné entraîne μ-intégrable car μ est finie. Par le théorème de convergence dominée :

$$\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-ita}-\mathrm{e}^{-itb}}{\mathrm{i}t}\varphi(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{\mathbb{R}}\lim_{T\to\infty} I_T(x)\,\mathrm{d}\mu(x),$$

ce qui est la formule d'inversion voulue d'après la première étape.

Pour un singleton, on peut écrire, toujours avec le théorème de Fubini :

$$\frac{1}{2\pi}\int_{-T}^{T}\mathrm{e}^{-\mathrm{i}tb}\varphi(t)\,\mathrm{d}t = \mu(\{b\}) + \int_{\mathbb{R}\setminus\{b\}}\operatorname{sinc}(T(x-b))\,\mathrm{d}\mu(x),$$

et faire tendre T vers l'infini par convergence dominée pour obtenir :

$$\lim_{T\to+\infty}\frac{1}{2T}\int_{-T}^{T}\mathrm{e}^{-\mathrm{i}tb}\varphi(t)\,\mathrm{d}t = \mu(\{b\}).$$

Avec ceci, on voit que les lois de deux variables aléatoires réelles qui ont même fonction caractéristique sont égales sur les segments, qui forment un π-système. Ainsi, les lois de ces variables aléatoires sont égales par le lemme des classes monotones. □

Remarque. Avec la décomposition de Hahn pour les mesures signées, on doit pouvoir en déduire l'injectivité de la transformée de Fourier pour les mesures signées.

Références : Jean-Yves Ouvrard, *Probabilités 2*. Pour le lien avec les mesures signées, on peut se renseigner dans le Rudin (*Analyse réelle et complexe*).

Chapitre 99

Séries de variables aléatoires

Dans tout le document, on s'intéresse à des séries de la forme $\sum_n X_n$, où les (X_i) sont des variables aléatoires indépendantes et L^2. On définit par (S_n) la suite des sommes partielles, et on note $\sigma_i^2 = \mathrm{Var}(X_i)$. On garde ces notations dans tout le document.

Théorème. (inégalité maximale de Kolmogorov) : On suppose les (X_i) centrées. Alors pour $\varepsilon > 0$ et pour tout n :

$$\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\sum_{i=1}^{n} \sigma_i^2.$$

Démonstration. Soit $M_n = \max_{1\leq k\leq n} |S_k|$ pour tout n. On souhaite majorer $\mathbb{P}(M_n \geq \varepsilon)$, on définit E comme l'événement en question. On suppose E non vide, et on va le partitionner selon le premier k pour lequel on va dépasser ε :

$$E_k := (|S_k| \geq \varepsilon) \cap \left(\bigcap_{i=1}^{k-1} (|S_i| < \varepsilon)\right),$$

pour avoir ainsi $\mathbb{P}(E) = \sum_{k=1}^{n} \mathbb{P}(E_k)$. Sur E_k, $|S_k|$ est par définition supérieure à ε donc

$$\mathbb{E}(\mathbf{1}_{E_k} S_k^2) \geq \varepsilon^2\, \mathbb{P}(E_k)$$

ce que l'on peut réécrire

$$\mathbb{P}(E) \leq \frac{1}{\varepsilon^2}\sum_{k=1}^{n} \mathbb{E}(\mathbf{1}_{E_k} S_k^2).$$

Or en développant le carré dans $\mathbb{E}(\mathbf{1}_{E_k} S_n^2)$ à partir du k-ième terme, on montre que le double produit s'annule par indépendance et centrage et donc que $\mathbb{E}(\mathbf{1}_{E_k} S_n^2) \geq \mathbb{E}(\mathbf{1}_{E_k} S_k^2)$ pour $1 \leq k \leq n$. Finalement, on a :

$$\mathbb{P}(E) \leq \frac{1}{\varepsilon^2}\,\mathbb{E}(\mathbf{1}_E S_n^2) \leq \frac{1}{\varepsilon^2}\,\mathbb{E}(S_n^2),$$

et on conclut par indépendance. □

Proposition. Si les (X_i) sont centrées, et si la série des variances converge, alors la série $\sum_n X_n$ converge presque sûrement (et dans L^2).

Démonstration. Pour la convergence dans L^2, on montre que la suite des sommes partielles

est de Cauchy (pour la norme 2), ce qui est direct au vu des hypothèses.

Passons à la convergence presque sûre. Pour une suite (x_n) de réels, dont les sommes partielles sont notées (s_n), on sait que la série des (x_n) converge si et seulement si la suite (s_n) est de Cauchy, si et seulement si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall k \in \mathbb{N}, |s_{N+k} - s_N| \leq \varepsilon,$$

ce qui équivaut encore à

$$\inf_{N\in\mathbb{N}} \sup_{k\in\mathbb{N}} |s_{N+k} - s_N| = 0$$

On revient donc aux probabilités et on pose A_m la variable aléatoire $\sup_{k\in\mathbb{N}} |S_{m+k} - S_m|$ ainsi que $A = \inf_m A_m$. On cherche donc à calculer $\mathbb{P}(A = 0)$. Soit donc $\varepsilon > 0$,

$$\mathbb{P}(A_m > \varepsilon) = \lim_{n\to\infty} \mathbb{P}\left(\sup_{1\leq k\leq n} |S_{m+k} - S_m| > \varepsilon \right).$$

Mais d'après l'inégalité de Kolmogorov

$$\mathbb{P}\left(\sup_{1\leq k\leq n} |S_{m+k} - S_m| > \varepsilon \right) \leq \frac{1}{\varepsilon^2} \sum_{j=m+1}^{m+n} \sigma_j^2,$$

donc

$$\mathbb{P}(A_m > \varepsilon) \leq \frac{1}{\varepsilon^2} \sum_{j=m+1}^{\infty} \sigma_j^2.$$

Ainsi

$$\mathbb{P}(A > \varepsilon) = \mathbb{P}\left(\bigcap_{m\in\mathbb{N}} (A_m > \varepsilon) \right) \leq \inf_{m\in\mathbb{N}} \mathbb{P}(A_m > \varepsilon) = 0$$

car la série des variances converge. On a donc $\mathbb{P}(A > \varepsilon) = 0$ pour tout ε, donc $A = 0$ presque sûrement, donc la série converge presque sûrement. □

Théorème. (loi des grands nombres avec un moment d'ordre deux)
Si $\mathbb{E}(X_n)$ tend vers m et si $\sum_{j=1}^{+\infty} \frac{\sigma_j^2}{j^2} < \infty$ alors $\overline{X_n}$ converge presque sûrement et dans L^2 vers m.

Pour prouver ce résultat, on aura besoin du lemme de Cesàro ainsi que d'un autre lemme sur les suites.

Lemme. (de Kronecker)
Soit (x_n) une suite dont la série associée est convergente, et (b_n) une suite croissante de réels qui tend vers l'infini. Alors

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{j=1}^{n} b_j x_j = 0.$$

Démonstration. Notons S la somme des (x_j). On pose $S_n = \sum_{j=1}^{n} x_j - S$, qui est une suite

qui tend donc vers 0, avec $x_n = S_n - S_{n-1}$. Une transformation d'Abel donne alors :

$$\sum_{j=N}^{n} b_j x_j = \sum_{j=N}^{n} b_j(S_j - S_{j-1}) = S_n b_n - b_N S_{N-1} - \sum_{j=N}^{n-1} S_j(b_{j+1} - b_j).$$

Ainsi, dès que $b_n \neq 0$:

$$\frac{1}{b_n}\sum_{j=1}^{n} b_j x_j = \frac{1}{b_n}\sum_{j=1}^{N} b_j x_j - \frac{b_N S_{N-1}}{b_n} + S_n - \frac{1}{b_n}\left(\sum_{j=N}^{n-1} S_j(b_{j+1} - b_j)\right).$$

À N fixé, tout peut être rendu arbitrairement petit quand n est grand, ce qu'on voulait. □

Passons maintenant à la loi des grands nombres.

Démonstration. D'après le lemme de Cesàro, $\mathbb{E}(\overline{X_n})$ tend vers m quand n tend vers l'infini. Les variables aléatoires $Y_n := \frac{X_n - \mathbb{E}(X_n)}{n}$ sont donc indépendantes, centrées et de variance $\frac{1}{n^2}\sigma_n^2$, donc

$$\sum_{j=1}^{\infty} \sigma_j^2 = \sum_{j=1}^{\infty} \frac{1}{j^2}\sigma_j^2 < \infty.$$

D'après notre théorème de convergence sur les séries de variables aléatoires, la série aléatoire de terme général Y_n converge presque sûrement. On peut alors appliquer le lemme de Kronecker à (Y_n) et $(b_n) = (n)$, qui donne que $(\frac{1}{n}\sum_{j=1}^{n} jY_j)$ converge presque sûrement vers 0. Ceci, en écrivant ce que vaut Y_n, prouve la convergence presque sûre de $(\overline{X_n})$ vers m.

Pour la convergence en norme 2, on voit que

$$\left\| \frac{1}{n}\sum_{j=1}^{n}(X_j - \mathbb{E}(X_j)) \right\|^2 = \frac{1}{n^2}\sum_{j=1}^{n}\sigma_j^2,$$

qui tend vers 0 par le lemme de Kronecker. Enfin, on écrit

$$\overline{X_n} - m = \frac{1}{n}\sum_{j=1}^{n}(X_j - \mathbb{E}(X_j)) + \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j) - m,$$

et on conclut par inégalité triangulaire. □

Référence : Probabilités 2, Jean-Yves Ouvrard.

Chapitre 100

Formule sommatoire de Poisson

Dans ce développement, on montre la formule sommatoire de Poisson ainsi que plein d'applications de celle-ci. On note $\mathcal{S}(\mathbb{R})$ la classe de Schwartz et on prend pour conventions :

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)\mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}\xi \text{ et } c_n(g) = \frac{1}{2\pi}\int_0^{2\pi} f(x)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x,$$

si f est dans L^1 et g est L^1 sur $[0, 2\pi]$ et 2π-périodique.

Théorème. (Formule sommatoire de Poisson)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $x \in \mathbb{R}$,

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}(n)\mathrm{e}^{\mathrm{i}nx}.$$

Démonstration. Soit $f_n : x \mapsto f(x + 2\pi n)$. On s'intéresse à la série de fonctions $\sum_{n\in\mathbb{Z}} f_n$. Soit $[-K, K]$ un compact de $\mathbb{R}$. Pour $|n| > K$, une majoration simple donne :

$$\begin{aligned} |f(x + 2\pi n)| &\leq \frac{M}{|1 + (x + 2\pi n)^2|} \text{car } f \text{ est dans la classe de Schwartz} \\ &\leq \frac{M}{1 + (K + 2\pi|n|)^2}. \end{aligned}$$

Ceci montre que la série considérée converge normalement sur tout compact, donc que la série définit une fonction continue sur $\mathbb{R}$. En refaisant le même travail sur la série des dérivées, on obtient le caractère $\mathcal{C}^1$ de la somme, que l'on note F. Par un changement d'indice, on montre simplement que F est 2π-périodique.

On peut donc appliquer le théorème de Dirichlet à F, *i.e.* F est somme de sa série de Fourier.

Calculons les coefficients de Fourier de F. Soit $n \in \mathbb{Z}$:

$$\begin{aligned} c_n(F) &= \frac{1}{2\pi}\int_0^{2\pi} F(x)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x \\ &= \sum_{k\in\mathbb{Z}}\int_0^{2\pi} f(x+2k\pi)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x \text{ (on intervertit par convergence normale)} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} f(u)\mathrm{e}^{-\mathrm{i}nx}\,\mathrm{d}x \text{ (on recolle les intégrales par changement de variable)} \\ &= \frac{1}{2\pi}\widehat{f}(n). \end{aligned}$$

Somme toute, on a bien la formule sommatoire de Poisson. □

Remarque. On peut appliquer cette formule sous des hypothèses plus faibles. En effet, la seule chose dont on a vraiment besoin est le caractère $\mathcal{C}^1$ de la somme, c'est-à-dire la convergence normale de la série de fonctions, pour pouvoir appliquer le théorème de Dirichlet. On peut donc se restreindre aux hypothèses suivantes :

— $f \in \mathcal{C}^0 \cap \mathrm{L}^1$ et $|f| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1$

— $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| < \infty$ ou bien $|f'| \leq \frac{M}{(1+\cdot)^\alpha}, M > 0, \alpha > 1.$

Corollaire. (Formule d'inversion de Fourier)
Soit $f \in \mathcal{S}(\mathbb{R})$. Pour tout $x \in \mathbb{R}$:

$$f(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\xi x}\widehat{f}(\xi)\,\mathrm{d}\xi.$$

Démonstration. On considère, pour $t \in \mathbb{R}$, fonction de $\mathcal{S}(\mathbb{R})$ définie par :

$$f_t : x \mapsto f(x)\mathrm{e}^{\mathrm{i}tx}.$$

Appliquons la formule sommatoire de Poisson à f_t puis à sa transformée de Fourier : pour tout $y \in \mathbb{R}$

$$\sum_{n\in\mathbb{Z}} f_t(y+2\pi n) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{f}_t(n)\mathrm{e}^{\mathrm{i}ny} = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{\widehat{f}}_t(2\pi n - y).$$

En calculant les termes de la première et de la troisième somme :

$$\sum_{n\in\mathbb{Z}} f(y+2\pi n)\mathrm{e}^{2\mathrm{i}\pi nt} = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} \widehat{\widehat{f}}(2\pi n - y)\mathrm{e}^{-2\mathrm{i}\pi nt}.$$

Par convergence normale des séries sur $[0,1]$, on peut intégrer sur cet intervalle en intervertissant et ainsi on ne garde que le terme en $n = 0$ qui donne :

$$f(y) = \frac{1}{2\pi}\widehat{\widehat{f}}(-y),$$

ce qui est bien la formule d'inversion de Fourier. □

Remarque. Par densité de la classe de Schwartz dans L^1, on peut en déduire la formule d'inversion de Fourier pour une fonction L^1.

Une autre application de la formule sommatoire de Poisson est le théorème d'échantillonnage de Shannon.

Théorème. (Échantillonnage de Shannon)
Soit $f \in \mathcal{S}(\mathbb{R})$ telle que le support de $\widehat{f}$ soit inclus dans $[-F, F]$ avec $F > 0$ et $2F \leq 2\pi$, alors :

$$f = \sum_{k\in\mathbb{Z}} f(k)\,\mathrm{sinc}(\pi(\,\cdot - k)).$$

Démonstration. On commence par utiliser la formule de Poisson. Pour $\xi \in \mathbb{R}$, on a :

$$\begin{aligned}
\widehat{f}(\xi) &= \mathbf{1}_{[-\pi,\pi]}(\xi)\widehat{f}(\xi) \quad \text{car le support de } \widehat{f} \text{ est inclus dans } [-\pi,\pi] \\
&= \mathbf{1}_{[-\pi,\pi]}(\xi)\sum_{n\in\mathbb{Z}} \widehat{f}(\xi + 2\pi n) \\
&= \mathbf{1}_{[-\pi,\pi]}(\xi)\frac{1}{2\pi}\sum_{n\in\mathbb{Z}} f(-n)\mathrm{e}^{\mathrm{i}n\xi} \text{ d'après la formule de Poisson et par inversion de Fourier} \\
&= \mathbf{1}_{[-\pi,\pi]}(\xi)\frac{1}{2\pi}\sum_{n\in\mathbb{Z}} f(n)\mathrm{e}^{-\mathrm{i}n\xi}.
\end{aligned}$$

Par inversion de Fourier et d'après le calcul que l'on vient de faire :

$$f(x) = \frac{1}{2\pi}\int_0^{2\pi}\sum_{n\in\mathbb{Z}} f(n)\mathrm{e}^{\mathrm{i}(x-n)\xi}\mathrm{d}\xi.$$

Par convergence normale de la série (puisque l'on est dans la classe de Schwartz), on peut intervertir :

$$f(x) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} f(n)\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}(x-n)\xi}\,\mathrm{d}\xi.$$

On reconnaît la transformée de Fourier de la fonction porte et cela permet donc de conclure. □

Remarque. En termes physiques, cela signifie que si le spectre d'un signal est assez bien localisé, on peut reconstruire le signal à partir d'un nombre discret de points. Cela peut paraître surprenant à première vue, mais en fait ça peut se comprendre. En effet, on peut montrer qu'une fonction dont la transformée de Fourier est à support compact peut être prolongée en une fonction entière. En particulier, elle est analytique sur $\mathbb{R}$ et on peut donc la reconstruire par sa série de MacLaurin.

Application. (Fonction ϑ de Jacobi)
On pose $\vartheta : x \mapsto \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 x}$, pour $x > 0$. La fonction ϑ est solution de l'équation fonctionnelle suivante sur $\mathbb{R}_+^*$:

$$\vartheta(x) = \frac{1}{\sqrt{x}}\vartheta\left(\frac{1}{x}\right).$$

Démonstration. On considère la fonction $x \mapsto \mathrm{e}^{-\alpha x^2}$ pour $\alpha > 0$. Sa transformée de Fourier est $\xi \mapsto \sqrt{\frac{\pi}{\alpha}}\mathrm{e}^{-\frac{\xi^2}{4\alpha}}$. On applique la formule de Poisson à cette fonction, qui est dans la classe de

Schwartz, ce qui donne :

$$\sum_{n\in\mathbb{Z}} \mathrm{e}^{-\alpha(2\pi n)^2} = \frac{1}{2\pi}\sqrt{\frac{\pi}{\alpha}} \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\frac{n^2}{4\alpha}}.$$

En prenant $\alpha = \frac{x}{4\pi}$, on obtient le résultat voulu. □

Références : la formule de Poisson et l'échantillonnage peuvent se trouver dans Bernis et Bernis, *40 développements pour l'agrégation.* La partie sur le thêta de Jacobi vient de Zuily et Queffélec, *Analyse pour l'agrégation.*

Chapitre 101

Processus de Galton-Watson

On s'intéresse ici au processus de Galton-Watson, étudié historiquement pour comprendre l'extinction des patronymes dans la bourgeoisie européenne.

Soit P une loi sur $\mathbb{N}$, appelée *loi de reproduction*, et soit $Z_0 = 1$. On définit par récurrence :

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k},$$

où $(X_{n,k})_{n,k\in\mathbb{N}^*}$ est une suite doublement indicée i.i.d de loi P. La suite de variables aléatoires $(Z_n)_n$ est une chaîne de Markov sur $\mathbb{N}$, dont la matrice de transition $\mathcal{P}$ est donnée par :

$$\forall z \in \mathbb{N}, \mathcal{P}(z, \cdot) = P^{*z}.$$

Remarquons alors que 0 est un état absorbant pour cette chaîne de Markov ; on définit le *temps d'extinction*

$$T := \inf\{n \geq 0, Z_n = 0\} \in \mathbb{N} \cup \{\infty\}.$$

Pour que le modèle soit intéressant (ni déterministe ni géométrique), il faut supposer :

$$0 < p_0 \leq p_0 + p_1 < 1.$$

On note m et σ^2 les moyenne et variance de P, supposées finies. On peut montrer que si m est finie, l'espérance de Z_n vaut m^n, ce qui amène à distinguer trois cas :

— le cas *sous-critique*, qui correspond à $m < 1$

— le cas *critique*, qui correspond à $m = 1$

— le cas *sur-critique*, qui correspond à $m > 1$.

Proposition. Pour $n \geq 1$, $g_n := G_{Z_n} = g^{\circ n}$, si g est la série génératrice de la loi de reproduction P.

Démonstration. Pour $n \geq 1$ et $s \in [0, 1]$:

$$g_{n+1}(s) = \mathbb{E}(s^{Z_{n+1}}) = \mathbb{E}(\mathbb{E}(s^{Z_{n+1}}|Z_n)) = \mathbb{E}(\mathbb{E}(s^{X_{1,1}})^{Z_n}) = \mathbb{E}(g(s)^{Z_n}) = g_n(g(s)),$$

d'où le résultat par récurrence. □

Théorème. Presque sûrement,

$$\lim_{n\to\infty} Z_n \in \{0, +\infty\},$$

et donc $T < \infty$ si et seulement si Z_n tend vers 0 (est nulle à partir d'un certain rang).

Démonstration. On se passe ici de la théorie des chaînes de Markov. On part de

$$\mathbb{P}(\lim Z_n \notin \{0, \infty\}) = \mathbb{P}(\exists k \in \mathbb{N}^*, Z_n = k \text{ pour une infinité de } n) \leq \sum_{k=1}^{+\infty} \mathbb{P}(\limsup(Z_n = k)).$$

Puisque 0 est un état absorbant :

$$\begin{aligned}
\mathbb{P}(Z_n = k) &\leq \mathbb{P}(Z_n \neq 0) \\
&\leq \mathbb{P}(Z_n \neq 0 \cap Z_{n-1} \neq 0) \\
&\leq \mathbb{P}(Z_n \neq 0 | Z_{n-1} \neq 0)\, \mathbb{P}(Z_{n-1} \neq 0) \\
&\leq (1 - p_0)\, \mathbb{P}(Z_{n-1} \neq 0) \\
&\leq \cdots \\
&\leq (1 - p_0)^{n-1}.
\end{aligned}$$

Ce dernier terme étant sommable en n, on peut appliquer le lemme de Borel-Cantelli qui assure que $\mathbb{P}(\limsup(Z_n = k)) = 0$ donc $\mathbb{P}(\lim Z_n \notin \{0, \infty\}) = 0$. □

Théorème. La probabilité $\mathbb{P}(T < \infty)$ est un point fixe de g. Si $m \leq 1$ alors $\mathbb{P}(T < \infty) = 1$ et si $m > 1$ alors $\mathbb{P}(T < \infty)$ est l'unique point fixe de g dans $]0, 1[$.

Démonstration. Puisque $Z_n \in \mathbb{N}$:

$$\{T < \infty\} = \{\lim Z_n = 0\} = \bigcup_{n\geq 0} (Z_n = 0).$$

Ainsi, par union croissante :

$$\mathbb{P}(T < \infty) = \lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \lim_{n\to\infty} g^{\circ n}(0).$$

Par continuité de g, la limite est donc un point fixe de g.

Par ailleurs, puisque $p_0 + p_1 < 1$, la fonction g est strictement convexe. Ainsi sa dérivée sur $[0, 1]$ et atteint son maximum en 1, avec $g'(1^-) = m$. On considère alors la fonction $h : x \mapsto g(x) - x$, sa dérivée double est la même que celle de g donc h est strictement convexe. De plus $h(0) = p_0 > 0$ et $h(1) = 0$. Si $m \leq 1$, alors $h'(1) = m - 1 \leq 0$. Comme $h'(0) = p_1 - 1 \leq 0$ et que h' est croissante, h' est toujours négative donc h est décroissante, et puisque $h(1) = 0$, h est toujours positive, et même strictement positive par stricte convexité. Ainsi le seul point fixe de g est 1. Maintenant, si $m > 1$, la dérivée de h change de signe entre 0 et 1 donc s'annule sur $]0, 1[$ par le théorème des valeurs intermédiaires. Par stricte convexité de h, ce point où la dérivée s'annule est l'unique minimum de h sur $]0, 1[$, notons-le z. La fonction h est alors croissante entre z et 1

et ne s'y annule pas. Par stricte convexité et théorème des valeurs intermédiaires, h admet un unique zéro sur $]0, z[$ (et donc sur $]0, 1[$). C'est l'unique point fixe de g sur $]0, 1[$. □

Proposition. Dans le cas sous-critique, la convergence de $\mathbb{P}(Z_n = 0)$ vers 1 est géométrique.

Démonstration. Sur $[0, 1]$, $|g'| \leq m$ donc g est m-contractante, et on a le résultat par théorème de point fixe. □

Proposition. Si Z_1 admet une variance et si $m = 1$, alors

$$1 - q_n := \mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}.$$

Démonstration. On peut écrire une formule de Taylor pour g (avec $g'(1^-) = 1$) :

$$1 - g(s) = 1 - s - \frac{\sigma^2}{2}(1 - s)^2 + (1 - s)^2\alpha(s),$$

où α est continue sur un voisinage de 1 et tend vers 0 quand s tend vers 1^-. On a alors

$$\frac{1}{1 - g(s)} - \frac{1}{1 - s} = \frac{g(s) - s}{(1 - g(s))(1 - s)} = \frac{\frac{\sigma^2}{2} + \alpha(s)}{1 - \frac{\sigma^2}{2}(1 - s) - (1 - s)\alpha(s)},$$

cette dernière quantité pouvant s'écrire $\frac{\sigma^2}{2} + \beta(s)$ avec β qui vérifie les mêmes propriétés que α. En notant $q_n = \mathbb{P}(Z_n = 0)$, on a

$$\frac{1}{1 - q_{n+1}} - \frac{1}{1 - q_n} \sim \frac{\sigma^2}{2},$$

d'où le résultat par sommation des relations de comparaison. □

Proposition. Si Z_1 admet un moment d'ordre trois, et si l'on note $\rho = g^{(3)}(1)$, alors

$$1 - q_n = \frac{2}{\sigma^2 n} + \frac{2}{\sigma^2}\left(\frac{2\rho}{3\sigma^4} - 1\right)\frac{\log(n)}{n^2} + \mathrm{o}\left(\frac{\log(n)}{n^2}\right).$$

Démonstration. On part cette fois-ci d'un développement de Taylor à l'ordre trois :

$$g(s) = s + \frac{\sigma^2}{2}(s - 1)^2 + \frac{\rho}{6}(s - 1)^3 + (s - 1)^3\alpha(s).$$

Et on obtient de la même façon :

$$\frac{1}{1 - g(s)} - \frac{1}{1 - s} = \frac{\sigma^2}{2} + \frac{\sigma^2}{2}\left(\frac{\sigma^2}{2} - \frac{\rho}{3\sigma^2}\right)(1 - s) + \beta(s)(1 - s)$$

En évaluant en q_n puis en sommant et en utilisant l'équivalent précédemment obtenu, on trouve :

$$\frac{1}{1 - q_n} = \frac{n\sigma^2}{2} + \left(\frac{\sigma^2}{2} - \frac{\rho}{3\sigma^2}\right)\log(n) + \mathrm{o}(\log(n)).$$

En inversant cette asymptotique on trouve le résultat voulu. □

Référence : Chafaï et Malrieu, *Recueil de Modèles Aléatoires*. L'expert de ce sujet est et restera Jonathan Badin.

Chapitre 102

Nombre de descentes d'une permutation aléatoire

L'objectif est de donner une expression simple et plus pratique de la loi du nombre de descentes dans une permutation de $\mathfrak{S}_n$.

Définition. Une permutation $\sigma \in \mathfrak{S}_n$ présente une descente au rang $i \in \{1, \dots, n-1\}$ si $\sigma(i+1) < \sigma(i)$. On définit le nombre eulérien $A(n,k)$ comme le nombre de permutations de $\mathfrak{S}_n$ qui présentent k descentes. Sous la loi uniforme, si on note D_n le nombre de descentes d'une permutation de $\mathfrak{S}_n$, alors on a $n!\,\mathbb{P}(D_n = k) = A(n,k)$.

L'objectif de ce développement est de montrer le théorème suivant, dû à S. Tanny (1973).

Théorème. Soit $(U_1, \dots, U_n)$ un vecteur iid $(\mathcal{U}(0,1))^{\otimes n}$, et $S_n = U_1 + \ldots + U_n$. Alors $\lfloor S_n \rfloor$ et D_n ont même loi.

Avant de passer à la preuve du théorème à proprement parler, on passe par un lemme de probas.

Lemme. Soit $(V_1, \dots, V_n)$ un vecteur iid $(\mathcal{U}(0,1))^{\otimes n}$ et $U_k := \{V_1 + \ldots + V_k\}$ où $\{\cdot\}$ désigne la partie fractionnaire. Alors $(U_1, \dots, U_n)$ est un vecteur iid de loi $(\mathcal{U}(0,1))^{\otimes n}$.

Démonstration. Soit $a \in \mathbb{R}$ montrons que $\{a+U\} \sim \mathcal{U}(0,1)$ pour $U \sim \mathcal{U}(0,1)$. En effet pour une fonction h continue bornée de $]0,1[$,

$$\begin{aligned}
\mathbb{E}(h(\{a+U\})) &= \int_0^1 h(\{a+x\})\,\mathrm{d}x \\
&= \int_0^{1-\{a\}} h(\{a\}+x)\,\mathrm{d}x + \int_{1-\{a\}} h(\{a\}+x-1)\,\mathrm{d}x \\
&= \int_{\{a\}}^1 h(x)\,\mathrm{d}x + \int_0^{\{a\}} h(x)\,\mathrm{d}x \\
&= \int_0^1 h(x)\,\mathrm{d}x = \mathbb{E}(h(U)).
\end{aligned}$$

Pour la deuxième égalité on utilise la définition de la partie fractionnaire et on distingue donc les cas et dans la suivante on fait des changements de variable.

Une fois que l'on a cela, on peut écrire, pour $h = h_1 \otimes \ldots \otimes h_n$ une fonction continue bornée sur l'espace produit $([0,1]^n)$:

$$\begin{aligned}
\mathbb{E}(h(\{V_1 + \ldots + V_n\})) &= \int_{v_1,\ldots,v_n \in [0,1]^n} h(\{v_1\}, \ldots, \{v_1 + \ldots + v_n\}) \, \mathrm{d}v_1 \cdots \mathrm{d}v_n \\
&= \int_{v_1,\ldots,v_n \in [0,1]^n} h_1(\{v_1\}) \cdots h_n(\{v_1 + \ldots + v_n\}) \, \mathrm{d}v_1 \cdots \mathrm{d}v_n \\
&= \int_{v_1=0}^{1} h_1(v_1) \int_{v_2=0}^{1} h_2(\{v_1 + v_2\}) \cdots \int_{v_n=0}^{1} h_n(\{v_1 + \ldots, v_n\}) \, \mathrm{d}v_n \cdots \mathrm{d}v_2 \, \mathrm{d}v_1 \\
&= \int_0^1 h_1(v_1) \, \mathrm{d}v_1 \cdots \int_0^1 h_n(v_n) \, \mathrm{d}v_n \\
&= \mathbb{E}(h_1(V_n)) \cdots \mathbb{E}(h_n(V_n))
\end{aligned}$$

Pour la troisième égalité, on a utilisé Fubini et pour la quatrième le petit résultat du début. Ainsi, on obtient que le vecteur $(U_1, \ldots, U_n)$ est iid de loi $(\mathcal{U}(0,1))^{\otimes n}$ (en effet, les fonctions continues bornées sous la forme $h_1 \otimes \ldots \otimes h_n$ sont denses dans les fonctions continues bornées sur $[0,1]^n$, puisqu'elles approchent les fonctions $\mathbf{1}_{A_1} \otimes \ldots \otimes \mathbf{1}_{A_n}$). □

On passe alors à la preuve du théorème.

Démonstration. On suppose U construit comme dans le lemme à partir de $(V_1, \ldots, V_n) \sim (\mathcal{U}(0,1))^{\otimes n}$. On simule alors une permutation aléatoire comme suit. On sait qu'il existe une unique permutation σ de $\mathfrak{S}_n$ telle que $U_{\sigma(1)}(\omega) < \ldots < U_{\sigma(n)}(\omega)$ pour presque tout $\omega \in \Omega$. On définit alors $\sigma(\omega) := \sigma$. On peut montrer que $\sigma \sim \mathcal{U}(\mathfrak{S}_n)$, et on a alors une descente au rang i si et seulement si $U_i > U_{i+1}$.

On définit alors M_k comme le point d'affixe $\mathrm{e}^{2\mathrm{i}\pi U_k}$ pour $k \in [\![1, n]\!]$, et on considère la marche qui consiste à parcourir le cercle en allant de M_1 à M_2, puis de M_2 à M_3 etc dans le sens direct. Trouvons l'angle Θ_k que l'on parcourt en allant de M_{k-1} à M_k, c'est-à-dire le réel $\Theta_k \in [0, 2\pi[$ tel que $\mathrm{e}^{2\mathrm{i}\pi(U_k - U_{k-1})} = \mathrm{e}^{\mathrm{i}\Theta_k}$. On écrit pour cela :

$$\begin{aligned}
U_k - U_{k-1} &= \{V_1 + \ldots + V_k\} - \{V_1 + \ldots + V_{k-1}\} \\
&= V_k - \lfloor V_1 + \ldots + V_k \rfloor + \lfloor V_1 + \ldots + V_{k-1} \rfloor \\
&= V_k + m,
\end{aligned}$$

où m est un entier. Ainsi, on a $\Theta_k = 2\pi V_k$. On a fait ce travail puisqu'une descente a lieu si et seulement si on passe par 1 dans notre marche sur le cercle. Ainsi, le nombre total de descentes correspond au nombre de multiples entiers de 2π présents dans $\Theta_1 + \ldots + \Theta_n$, soit à :

$$D_n = \left\lfloor \frac{\Theta_1 + \ldots + \Theta_n}{2\pi} \right\rfloor = \lfloor V_1 + \ldots + V_n \rfloor \sim \lfloor S_n \rfloor,$$

d'où le résultat. □

On peut ajouter à ce résultat qui fait intervenir une probabilité une asymptotique, afin d'avoir un résultat qui est plus de l'ordre du dénombrement.

Proposition. Soit $\alpha \in [0,1]$, alors, lorsque n tend vers $+\infty$:

$$A(n, \alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}} \, \mathrm{e}^{-6(\alpha - 1/2)^2}.$$

Démonstration. On veut utiliser le théorème limite central. Pour cela, on se rappelle que $\mathbb{E}(S_n) = \frac{n}{2}$ et $\mathrm{Var}(S_n) = \frac{n}{12}$. Ainsi :

$$Z_n := \frac{\sqrt{12}(S_n - n/2)}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1).$$

De plus,

$$\begin{aligned}\mathbb{P}(\alpha n < S_n < \alpha n + 1) &= \mathbb{P}\left(\sqrt{12}(\alpha - 1/2) < \frac{\sqrt{12}(S_n - n/2)}{\sqrt{n}} < \sqrt{12}(\alpha - 1/2) + \frac{\sqrt{12}}{\sqrt{n}}\right) \\ &=: \mathbb{P}\left(\beta < Z_n < \beta + \frac{\sqrt{12}}{\sqrt{n}}\right).\end{aligned}$$

Pour F_X la fonction de répartition d'une variable aléatoire X, on a alors :

$$\begin{aligned}\mathbb{P}\left(\beta < Z_n < \beta + \frac{\sqrt{12}}{\sqrt{n}}\right) &= F_{Z_n}(\beta + \sqrt{12/n}) - F_{Z_n}(\beta)) \\ &= F_{Z_n}(\beta + \sqrt{12/n}) - \Phi(\beta + \sqrt{12/n}) + \\ &\Phi(\beta + \sqrt{12/n}) - \Phi(\beta) + \\ &\Phi(\beta) - F_{Z_n}(\beta) \\ &= \sqrt{12/n}\, f(\beta) + \mathrm{o}(1/\sqrt{n}).\end{aligned}$$

Ou le petit o provient d'une généralisation du théorème limite central pour les densités, et où f désigne la densité d'une loi normale centrée réduite. On a donc :

$$A(n, \alpha n) \underset{\alpha n \in \mathbb{N}}{\sim} n! \sqrt{\frac{12}{n}} \sqrt{\frac{6}{\pi}} \, \mathrm{e}^{-6(\alpha - 1/2)^2}.$$

□

Références : page Wikipédia *Loi uniforme continue*. Pour le TCL, on peut regarder William Feller, *An Introduction to Probability Theory and Its Applications, vol. II*, page 533. Merci à Jürgen Angst pour l'aide sur la preuve de l'asymptotique finale.

Chapitre 103

Espaces de Sobolev et problèmes aux limites

Dans ce document, on introduit l'espace de Sobolev H^1 dans l'optique de résoudre le problème de Dirichlet avec conditions aux bords nulles.

Définition. On définit, pour I un intervalle de $\mathbb{R}$ ouvert, l'*espace de Sobolev* $\mathrm{H}^1(I)$ par :

$$\mathrm{H}_0^1(I) = \left\{ u \in \mathrm{L}^2(I), \exists g \in \mathrm{L}^2(I), \int_I u\varphi' = -\int_I g\varphi \qquad \forall \varphi \in \mathcal{C}_\mathrm{c}^1(I) \right\}.$$

Remarque. On peut aussi prendre comme espace de fonctions test $\mathcal{D}(I)$. De plus, le g de la définition peut être noté u', puisqu'il coïncide avec la vraie dérivée de u si u est $\mathcal{C}^1$ sur I. De plus, par densité des fonctions test, cette dérivée est unique.

Proposition. L'espace $\mathrm{H}^1(I)$ est préhilbertien pour le produit scalaire :

$$\langle u, v \rangle = \langle u, v \rangle_{\mathrm{L}^2} + \langle u', v' \rangle_{\mathrm{L}^2}.$$

La norme associée est équivalente à la norme $\|x\|_{\mathrm{H}^1} = \|u\|_2 + \|u'\|_2$.

Proposition. C'est même un espace de Hilbert (séparable).

Démonstration. Soit (u_n) une suite de Cauchy $\mathrm{H}^1(I)$. Nécessairement, (u_n) et (u_n') sont des suites de Cauchy dans L^2. Ainsi, elles convergent respectivement vers u et g des fonctions de L^2. Dès lors, en faisant tendre n vers l'infini dans la relation

$$\int_I u_n\varphi' = -\int_I u_n'\varphi,$$

on obtient

$$\int_I u\varphi' = -\int_I g\varphi \quad \forall \varphi \in \mathcal{D}(I).$$

Ainsi, $u \in \mathrm{H}^1(I)$, $u' = g$ et $\|u_n - u\|_{\mathrm{H}^1}$ tend vers 0. □

Lemme. Soit $f \in \mathrm{L}^1_{\mathrm{loc}}(I)$ telle que $\int_I f\varphi' = 0 \forall \varphi \in \mathcal{D}(I)$. Alors f est constante (presque partout).

Démonstration. Soit $\psi \in \mathcal{C}_c^0(I)$ telle que $\int_I \psi = 1$. Pour tout $w \in \mathcal{C}_c^0(I)$, il existe $\varphi \in \mathcal{C}_c^1(I)$ telle que $\varphi' = w - (\int_I w)\psi$. En effet, $h := w - (\int_I w)\psi$ est continue à support compact inclus dans I. Puisque son intégrale sur I est nulle, elle admet une primitive à support compact. Ainsi, on peut réécrire notre relation sur f comme :

$$\forall w \in \mathcal{C}_c^0(I), \quad \int_I f \left[w - (\int_I w)\psi \right] = 0,$$

soit

$$\forall w \in \mathcal{C}_c^0(I), \quad \int_I \left[f - \left(\int_I f\psi \right) \right] w = 0.$$

Par densité, comme précédemment, $f = \int_I f\psi$ presque partout. □

Lemme. Soit $g \in \mathrm{L}^1_{\mathrm{loc}}(I)$. On fixe $y_0 \in I$ et on pose

$$v(x) = \int_{y_0}^{x} g(t)\,\mathrm{d}t, \qquad x \in I.$$

Alors v est continue sur I et pour toute fonction test φ :

$$\int_I v\varphi' = -\int_I g\varphi.$$

Démonstration. Elle est continue d'après les propriétés générales de l'intégrale et on peut faire le calcul qui suit par Fubini. □

Avec ces deux lemmes, on obtient un résultat fondamental sur l'espace de Sobolev.

Théorème. Soit $u \in \mathrm{H}^1(I)$. Alors u est égale presque partout à une fonction continue.

Démonstration. On fixe $y_0 \in I$ et on pose $\widetilde{u}(x) = \int_{y_0}^{x} u'(t)\,\mathrm{d}t$. D'après les lemmes précédents, $u - \widetilde{u} = C$ presque partout donc $\widetilde{u} + C$ convient. □

Théorème. (opérateur de prolongement)
Il existe un opérateur de prolongement $P : \mathrm{H}^1(I) \to \mathrm{H}^1(\mathbb{R})$ linéaire continu tel que pour tout u :

- $P(u_{|I}) = u$
- $\|P(u)\|_2 \leq C\|u\|_2$
- $\|P(u)\|_{\mathrm{H}^1(\mathbb{R})} \leq C\|u\|_{\mathrm{H}^1(I)}$,

avec C une constante qui ne dépend que de I.

Démonstration. On survole rapidement les idées de la preuve. Dans le cas $I = \mathbb{R}_+^*$, on prolonge par réflexion, et on montre que ça convient. Pour le cas $I =]0,1[$, on construit $\eta \in \mathcal{C}^1$, $0 \leq \eta \leq 1$, qui vaut 1 sur $]-\infty, \frac{1}{4}[$, et 0 pour $x > \frac{3}{4}$. On prolonge alors $u \in \mathrm{H}^1(I)$ par 0 sur $\mathbb{R}_+^*$ et on montre que $\eta u \in \mathrm{H}^1(\mathbb{R}_+^*)$, et qu'on peut dériver le produit comme on a l'habitude de la faire. Ainsi, si on écrit $u = \eta u + (1-\eta)u$, on peut prolonger chacune de ces fonctions sur le côté qui coince, puis prolonger par réflexion, puis prendre comme opérateur de prolongement la somme des prolongements de ηu et $(1 - \eta u)$. □

Théorème. Soit $u \in \mathcal{H}^1(I)$. Il existe une suite $(u_n) \in \mathcal{D}(\mathbb{R})^{\mathbb{N}}$ telle que $(u_{n|I})$ converge dans $\mathrm{H}^1(I)$ vers u.

Démonstration. Le théorème précédent permet de traiter uniquement le cas $I = \mathbb{R}$. On utilise deux lemmes de convolution et troncature, mentionnés ci-dessous dans démonstration :

— Si $\rho \in \mathrm{L}^1(\mathbb{R})$ et $v \in \mathrm{H}^1(\mathbb{R})$. Alors $(\rho * v) \in \mathrm{H}^1(\mathbb{R})$ et $(\rho * v)' = \rho * v'$.

— Si $\zeta \in \mathcal{D}(\mathbb{R})$ est telle que $0 \leq \zeta \leq 1, \zeta(x) = 1$ si $|x| \leq 1$ et $= 0$ si $|x| \geq 2$. Alors $\zeta(x/\cdot)f$ converge dans L^2 vers f ($f \in \mathrm{L}^2$).

Une fois que l'on a ces deux lemmes, on considère (ρ_n) une suite régularisante et $u_n := \zeta_n(\rho_n * u)$. On a alors

$$u_n - u = \zeta_n((\rho_n * u) - u) + (\zeta_n u - u)$$

d'une part et

$$u_n' - u' = \zeta_n'(\rho_n * u) + \zeta_n(\rho_n * u') - u'$$

d'autre part. On majore alors les normes pour obtenir ce qu'il faut. □

Définition. On note $\mathrm{H}^1_0(I)$ l'adhérence dans $\mathrm{H}^1(I)$ de $\mathcal{C}^1_{\mathrm{c}}$.

Remarque. C'est un espace de Hilbert séparable muni de la norme induite. De plus, $\mathcal{D}(I)$ est dense dans $\mathrm{H}^1_0(I)$. De plus, $\mathrm{H}^1(I) \cap \mathcal{C}^1_{\mathrm{c}}(I) \subset \mathrm{H}^1_0(I)$.

Théorème. Soit $u \in \mathrm{H}^1(I)$. Alors $u \in \mathrm{H}^1_0(I)$ si et seulement si $u = 0$ sur ∂I (la frontière de I).

Démonstration. Le sens direct est direct. Réciproquement, si u s'annule sur ∂I, on fixe $G \in \mathcal{C}^1(\mathbb{R})$ telle que $G(t) = 0$ si $|t| \leq 1$ et $G(t) = t$ si $|t| \geq 2$, avec $|G(t)| \leq |t|$ pour tout t. On pose alors $u_n = \frac{1}{n} G \circ (nu) \in \mathrm{H}^1(I)$. On vérifie alors que le support de u_n est inclus dans $\{|u| \geq \frac{1}{n}\}$. Ainsi u_n est à support compact donc dans H^1_0. Par convergence dominée, on vérifie que u_n converge dans H^1 vers u. □

Proposition. (inégalité de Poincaré)
On suppose que I est borné. Alors il existe C ne dépendant que de I tel que

$$\|u\|_{\mathrm{H}^1} \leq C\|u'\|_2 \qquad \forall u \in \mathrm{H}^1_0(I).$$

Démonstration. Si $I =]a, b[$, on écrit

$$|u(x)| = |u(x) - u(a)| \leq \int_a^x |u'| \leq \|u'\|_1 \leq C\|u'\|_2 \text{ par Cauchy-Schwarz,}$$

on conclut en mettant au carré, puis en intégrant, puis en prenant la racine carrée. □

Maintenant que l'on a tous ces outils, on veut résoudre le problème suivant :

$$-u'' + u = f$$

sur $I =]0, 1[$ avec les conditions $u(0) = u(1) = 0$, et f une fonction donnée.

Définition. Une *solution faible* de ce problème est une fonction $u \in \mathrm{H}^1_0(I)$ telle que

$$\int_I u'v' + \int_I uv = \int_I fv \qquad \forall v \in \mathrm{H}^1_0(I).$$

Proposition. Le problème précédent admet une unique solution faible si $f \in \mathrm{L}^2$.

Démonstration. Il s'agit juste d'appliquer le théorème de Riesz sur H^1_0 pour la forme linéaire $v \mapsto \int_I fv$. □

Remarque. Si $f \in \mathrm{L}^2(I)$, alors $u' \in \mathrm{H}^1$ et si $f \in \mathcal{C}^0$, alors $u \in \mathcal{C}^2$.

On peut aller plus loin en ajoutant des coefficients. On prend toujours $I =]0,1[$.

Théorème. (problème de Sturm-Liouville)
Soit $p \in \mathcal{C}^1(\overline{I})$ tel que $p \geq \alpha > 0$, $q \in \mathcal{C}^0(\overline{I})$ positive non nulle et $f \in \mathrm{L}^2(I)$. On s'intéresse au problème

$$-(pu')' + qu = f$$

avec les conditions aux bords $u(0) = u(1) = 0$.

Démonstration. On va adapter la preuve précédente, mais avec le produit scalaire $\langle u, v\rangle' = \langle pu', v'\rangle + \langle qu, v\rangle$. C'est bien un produit scalaire car p et q sont positives, avec $p \geq \alpha > 0$ sur $\overline{I}$. Par Poincaré d'une part et car p et q sont bornées d'autre part, la norme associée à ce nouveau produit scalaire est équivalente à celle de H^1. Ainsi, on obtient un espace de Hilbert et on peut appliquer Riesz à la forme linéaire continue $v \mapsto \int_I fv$. On obtient alors bien le résultat voulu. □

Référence : Haïm Brézis, *Cours d'analyse fonctionnelle.*

Chapitre 104

Optimisation dans un espace de Hilbert

Dans ce développement, on introduit la notion de convergence faible dans un espace de Hilbert, dans l'optique d'étudier un problème de minimisation qui ne se résout pas "simplement" à l'aide du théorème de Lax-Milgram.

Définition. Soit H un espace de Hilbert et $(x_n)_n$ une suite d'éléments de H. On dit que $(x_n)_n$ converge faiblement vers $x \in H$ si :

$$\forall y \in H, \langle x_n, y\rangle \to \langle x, y\rangle.$$

On note alors $x_n \rightharpoonup x$.

Théorème. (de Banach-Alaoglu)
Soit H un espace de Hilbert séparable et $(x_n)_n$ une suite bornée dans H. Alors il existe une extraction de $(x_n)_n$ qui converge faiblement dans H.

Remarque. En général, on n'a aucune chance qu'une suite bornée admette une sous-suite convergente dans H. En effet, le théorème de Riesz assure qu'une boule n'est jamais compacte en dimension infinie.

Démonstration. On note $(h_n)_n$ une suite dense de H et on note $M = \sup \|x_n\|$. Pour chaque k, la suite $(\langle x_n, h_k\rangle)_n$ est bornée. On peut donc en extraire une sous suite qui converge vers un $\varphi(x_k)$. Par extraction diagonale, on peut exhiber une extractrice qui va faire converger simultanément tous les $(\langle x_n, h_k\rangle)_n, k \in \mathbb{N}$. Notons ψ cette extractrice et montrons alors que pour un $y \in H$, la suite $(\langle x_{\psi(n)}, y\rangle)$ converge, *i.e.* qu'elle est de Cauchy. Pour p et q deux entiers, et $k \in \mathbb{N}$, on a :

$$|\langle x_{\psi(p)}, y\rangle - \langle x_{\psi(q)}, y\rangle| \leq |\langle x_{\psi(p)}, y - h_k\rangle| + |\langle x_{\psi(p)}, h_k\rangle - \langle x_{\psi(q)}, h_k\rangle| + |\langle x_{\psi(q)}, h_k - y\rangle|.$$

Pour $\varepsilon > 0$, en prenant h_k assez proche de y, on peut rendre les termes extrémaux plus petits que ε par Cauchy-Schwarz. Pour le terme du milieu, on peut le rendre plus petit que ε pour $p, q \to +\infty$ puisque la suite $(\langle x_{\psi(p)}, h_k\rangle)_n$ converge donc est de Cauchy. Somme toute, on a bien :

$$\forall y \in H, \langle x_{\psi(p)}, y\rangle \to \varphi(y).$$

Par linéarité de la limite et bilinéarité du produit scalaire, φ est une forme linéaire. Elle est

de plus continue (avec M comme module de continuité par Cauchy-Schwarz), donc on peut la représenter par un vecteur $x^* \in H$, ce qui achève la démonstration. □

Théorème. Soit H un espace de Hilbert et $J : H \to \mathbb{R}$ une fonction continue convexe et coercitive. Alors J atteint son minimum en un point x^*.

Démonstration. Soit (x_n) une suite minimisante, c'est-à-dire telle que $J(x_n) \to \inf_H J$. Cette suite est bornée, sinon cela contredirait l'hypothèse de coercitivité. Comme c'est une suite bornée, on peut appliquer le résultat du théorème précédent, et exhiber un $x^* \in H$ tel que (quitte à extraire) l'on ait $x_n \rightharpoonup x^*$.

Posons $\alpha > \inf_H J$ et $C_\alpha = J^{-1}(]-\infty, \alpha])$, qui est non vide par définition de la borne inférieure et fermé par continuité de J. On peut aussi voir simplement que la convexité de J entraîne celle de C_α. On peut donc définir p_α le projection sur C_α. Pour k assez grand, x_k appartient à C_α et on a alors, d'après le théorème de projection sur le convexe C_α :

$$\langle x^* - p_\alpha(x^*), x_k - p_\alpha(x^*)\rangle \leq 0.$$

Par continuité du produit scalaire, on obtient, pour k qui tend vers l'infini, $\|x^* - p_\alpha(x^*)\|^2 \leq 0$ soit $x^* \in C_\alpha$. Ceci valant pour tout α, on peut le faire tendre vers $\inf_H J$, ce qui donne $J(x^*) \leq \inf_H J$ et qui permet donc de conclure. □

Application. (Résolution d'un problème aux limites non linéaires au sens faible)
Soit $f \in \mathrm{L}^2(0,1)$ et $p > 0$. Il existe une unique fonction $u \in \mathrm{H}^1_0(0,1)$ tel que :

$$-u'' + u|u|^{p-1} = f,$$

où l'égalité est à comprendre au sens des distributions.

Démonstration. Qui dit au sens faible dit formulation faible. On considère la fonctionnelle J définie sur $\mathrm{H}^1_0(0,1)$ définie par :

$$J : u \longmapsto \int_0^1 \left(\frac{(u')^2}{2} + \frac{|u|^{p+1}}{p+1} - fu \right)$$

Pour alléger les notations, on note $\|\cdot\|$ la norme sur H^1_0 et $\|\cdot\|_2$ la norme L^2.

Soient $u, v \in \mathrm{H}^1_0(0,1)$:

$$\begin{aligned} J(u+v) &= \int_0^1 \left(\frac{(u'+v')^2}{2} + \frac{|u+v|^{p+1}}{p+1} - f(u+v) \right) \\ &J(u) + \int_0^1 \left(u'v' + \frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - fv \right) + \frac{1}{2}\|v'\|_2^2 \\ &= J(u) + \int_0^1 \left(u'v' + |u|^{p-1}uv - fv \right) + \int_0^1 \left(\frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv \right) + o(\|v\|). \end{aligned}$$

Montrons donc que $\varepsilon_u(v) := \int_0^1 \left(\frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv \right) = o(\|v\|)$.

Pour x, y deux *réels*, on pose $\varphi : t \mapsto \varepsilon_x(ty)$. C'est une fonction de t dérivable avec :

$$\varphi'(t) = y(p+1)\sigma(x+ty)|x+ty|^p - x|x|^{p-1}y,$$

où σ est la fonction signe, qui vaut par convention 0 en 0. Ainsi d'après l'inégalité des accroissements finis sur le segment $[0,1]$:

$$\left| \frac{|x+y|^{p+1}}{p+1} - \frac{|x|^p}{p+1} - |x|^{p-1}xy \right| \leq |y| \sup_{|t|\leq|y|} \{|\sigma(x+t)|x+t|^p| - \sigma(x)|x|^p|\}.$$

En rentrant cette inégalité dans l'intégrale, on a alors :

$$\begin{aligned}
\left| \int_0^1 \left(\frac{|u+v|^{p+1}}{p+1} - \frac{|u|^{p+1}}{p+1} - |u|^{p-1}uv \right) \right| &\leq \int_0^1 |v| \sup_{|t|\leq\|v\|_\infty} \{|\sigma(u+t)|u+t|^p| - \sigma(u)|u|^p|\} \\
&\leq \|v\|_2 \sqrt{\int_0^1 (\sup_{|t|\leq\|v\|_\infty} \{|\sigma(u+t)|u+t|^p| - \sigma(u)|u|^p|\})^2} \\
&\leq \|v\| \sqrt{\int_0^1 (\sup_{|t|\leq\|v\|_\infty} \{|\sigma(u+t)|u+t|^p| - \sigma(u)|u|^p|\})^2}.
\end{aligned}$$

Il nous reste à montrer que cette dernière intégrale tend vers 0 lorsque $\|v\|$ tend vers 0. Par continuité de la fonction $x \mapsto \sigma(x)|x|^p$, la borne supérieure dans l'intégrale converge simplement vers 0 quand $\|v\|$ tend vers 0. On peut alors supposer $\|v\| \leq 1$ dans perte de généralité et alors dominer l'intégrande par $((|u|+1)^p + |u|^p)^2 \in \mathrm{L}^1(0,1)$ car u admet un représentant continu. Par convergence dominée, on a la convergence vers 0. Ainsi, la différentielle de J en u vaut :

$$\mathrm{d}J_u(v) = \int_0^1 \big(u'v' + |u|^{p-1}uv - fv\big).$$

En effet, on vient de voir que l'on avait le développement à l'ordre un souhaité, il faut maintenant voir que c'est bien une forme linéaire continue. Le caractère linéaire est évident. Pour la continuité :

$$\begin{aligned}
|\mathrm{d}J_u(v)| &\leq \int_0^1 (|u'v'| + |u|^{p-1}uv + |fv|) \\
&\leq \sqrt{\int_0^1 |u'|^2}\|v'\|_2 + \sqrt{\int_0^1 |u|^{2p}}\|v\|_2 + \|f\|_2\|v\|_2 \\
&\leq C\|v\|,
\end{aligned}$$

puisque toutes les intégrales convergent (la première car $u' \in \mathrm{L}^2$, la deuxième car u admet un représentant continu).

Montrons maintenant que J est strictement convexe. L'intégrande est convexe en u ainsi il est facile de voir que J est convexe. Si on a égalité pour un $\lambda \in]0,1[$ de $J(\lambda u + (1-\lambda)v)$ et

$\lambda J(u) + (1-\lambda)J(v)$, alors on a :

$$\frac{1}{p+1}\int_0^1 (|\lambda u + (1-\lambda v)|^{p+1} - \lambda|u|^{p+1} - (1-\lambda)|v|^{p+1}) = \frac{1}{2}\int_0^1 (\lambda u'^2 - (\lambda u' + (1-\lambda)v')^2).$$

Par convexité du carré et de $|\cdot|^{p+1}$ sur $\mathbb{R}$, la première intégrande est négative presque partout et la deuxième positive presque partout. Nécessairement, les deux intégrandes sont nulles presque partout. En particulier, on a presque partout, et même partout par continuité :

$$|\lambda u + (1-\lambda)v|^{p+1} = \lambda|u|^{p+1} + (1-\lambda)|v|^{p+1}.$$

On a alors $u = v$ puisque $t \mapsto |t|^{p+1}$ est strictement convexe.

Enfin, J est coercitive puisque :

$$J(u) \geq \frac{1}{2}\int_0^1 u'^2 - \|u\|_2\|f\|_2 \geq C\|u\|^2 - \|u\|_2\|f\|_2 \xrightarrow[\|u\|\to\infty]{} +\infty,$$

où l'on a utilisé l'inégalité de Poincaré. D'après le théorème précédent, J admet donc un minimiseur u^*. En écrivant la condition nécessaire de minimalité, on retombe sur notre problème de départ au sens faible.

Pour l'unicité, il faut voir que si on a un autre u qui minimise, alors la différentielle de J en ce u s'annule sur l'espace des fonctions test. Par continuité de la différentielle et densité des fonctions test dans H_0^1, u est un point critique de J. Par stricte convexité de J, c'est forcément u^*. □

Références : la preuve de Banach-Alaoglu doit être dans Brezis, *Analyse Fonctionnelle*. L'application provient de la page web de Thomas Cavallazzi.

Cinquième partie

Couplages

Couplages par leçon

101. Groupe opérant sur un ensemble

— nombre de nilpotents sur un corps fini

— Quaternions

102. Nombres complexes de module un

— théorème de Dirichlet faible

— marche aléatoire sur le cercle

103. Conjugaison dans un groupe

— automorphismes de $\mathfrak{S}_n$

— Quaternions

104. Groupes finis

— automorphismes de $\mathfrak{S}_n$

— sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$

105. Groupe symétrique

— nombre de cycles d'une permutation aléatoire

— automorphismes de $\mathfrak{S}_n$

106. Groupe linéaire

— sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$

— nombre de nilpotents sur un corps fini

108. Parties génératrices d'un groupe

— Quaternions

— automorphismes de $\mathfrak{S}_n$

120. Anneaux $\mathbb{Z}/n\mathbb{Z}$

— théorème de Dirichlet faible

— Loi de réciprocité quadratique

121. Nombres premiers

— théorème de Dirichlet faible

— loi de réciprocité quadratique

122. Anneaux principaux

— Équation de Mordell

— forme normale de Smith

123. Corps finis

— nombre de polynômes irréductibles

— loi de réciprocité quadratique

125. Extensions de corps

— polygones constructibles

— nombre de polynômes irréductibles

127. Nombres remarquables

— polygones constructibles

— Équation de Mordell

141. Polynômes irréductibles

— nombre de polynômes irréductibles

— polygones constructibles

142. PGCD, PPCM

— forme normale de Smith

— décomposition de Dunford

144. Racines d'un polynôme

— nombres de polynômes irréductibles

— sous-groupes finis de $\mathrm{GL}(\mathbb{Z})$

148. Dimension

— polygones constructibles

— décomposition en valeurs singulières

149. Déterminant

— déterminant de Gram

— déterminant de Vandermonde et de Cauchy

150. Polynômes d'endomorphismes

— décomposition de Dunford

— nombre de nilpotents sur un corps fini

151. Sous-espaces stable

— décomposition de Dunford

— nombre de nilpotents sur un corps fini

152. Endomorphismes diagonalisables

— sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$

— décomposition en valeurs singulières

153. Valeurs propres, vecteurs propres

— sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$

— décomposition de Dunford

155. Exponentielle de matrices

— décomposition de Dunford et surjectivité

— décomposition polaire

156. Endomorphismes trigonalisables, nilpotents

— décomposition de Dunford

— nombre de nilpotents sur un corps fini

157. Matrices symétriques, hermitiennes

— sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$

— décomposition en valeurs singulières

158. Endomorphismes remarquables

— décomposition en valeurs singulières

— Quaternions

159. Dualité

— théorème de Krein-Milman

— décomposition en valeurs singulières

161. Application et isométries affines

— déterminant de Gram

— Quaternions

162. Systèmes linéaires

— forme normale de Smith

— décomposition en valeurs singulières

170. Formes quadratiques

— loi de réciprocité quadratique

— décomposition en valeurs singulières

171. Formes quadratiques réelles (IMPASSE)

181. Convexité

— théorème de Krein-Milman

— sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$

190. Combinatoire

— nombre de cycles d'une permutation aléatoire

— nombre de nilpotents sur un corps fini

191. Utilisation de techniques d'algèbre en géométrie

— décomposition en valeurs singulières

— théorème de Krein-Milman

201. Espaces de fonctions

— problème de Sturm-Liouville

— théorème de Banach-Steinhaus et séries de Fourier

203. Utilisation de la notion de compacité

— théorème d'Hadamard-Lévy

— optimisation dans un Hilbert

204. Connexité en analyse

— théorème d'Hadamard-Lévy

— théorème de Charkovski

205. Espaces complets

— optimisation dans un Hilbert

— théorème de Banach-Steinhaus et séries de Fourier

206. Dimension finie en analyse

— déterminant de Gram

— théorème de Hadamard-Lévy

208. Espaces vectoriels normés

— théorème de Banach-Steinhaus et séries de Fourier

— problème de Sturm-Liouville

209. Approximation par des fonctions régulières

— théorème de Fejér

— théorème taubérien fort

213. Espaces de Hilbert

— problème de Sturm-Liouville

— optimisation dans un Hilbert

214. théorème d'inversion locale

— lemme de Morse

— théorème de Hadamard-Lévy

215. Applications différentiables

— lemme de Morse

— théorème de Hadamard-Lévy

218. Formules de Taylor

— lemme de Morse

— théorème limite central

219. Extremums

— optimisation dans un Hilbert

— décomposition en valeurs singulières

220. Exemples d'EDO

— théorème de Hadamard-Lévy

— théorème de Lyapounov

221. Équations différentielles linéaires

— théorème de Lyapounov

— problème de Sturm-Liouville

223. Suites numériques

— séries de variables aléatoires

— processus de Galton-Watson

224. Développements asymptotiques

— théorème limite central

— processus de Galton-Watson

226. Suites récurrentes

— théorème de Charkovski

— processus de Galton-Watson

228. Continuité, dérivabilité

— théorème de Charkovski

— théorème limite central

229. Fonctions monotones, fonctions convexes

— optimisation dans un Hilbert

— processus de Galton-Watson

230. Séries numériques

— séries de variables aléatoires

— théorème taubérien fort

234. Espaces de fonctions Lebesgue-intégrables

— théorème de Fejér

— problème de Sturm-Liouville

235. Interversion de symboles

— formule sommatoire de Poisson

— théorème taubérien fort

236. Calcul d'intégrales

— formule des compléments

— méthode de Laplace

239. Intégrales à paramètres

— méthode de Laplace

— formule sommatoire de Poisson

241. Suites et séries de fonctions

— formule sommatoire de Poisson

— théorème taubérien fort

243. Séries entières

— indécomposabilité de la loi de Poisson

— théorème taubérien fort

245. Fonctions holomorphes

— formule des compléments

— indécomposabilité de la loi de Poisson

246. Séries de Fourier

— formule sommatoire Poisson

— théorème de Fejér

250. Transformation de Fourier

— formule sommatoire de Poisson

— théorème limite central

253. Convexité en analyse

— optimisation dans un Hilbert

— théorème de Krein-Milman

261. Loi d'une variable aléatoire

— nombres eulériens

— indécomposabilité de la loi de Poisson

262. Convergence d'une suite de variables aléatoires

— théorème limite central

— séries de variables aléatoires

264. Variables aléatoires discrètes

— indécomposabilité de la loi de Poisson

— processus de Galton-Watson

266. Indépendance en probabilités

— nombres eulériens

— séries de variables aléatoires indépendantes

Couplage par développement

Paramétrisation de $\mathrm{SO}_3(\mathbb{R})$ par les quaternions (101,103,108,158,161).

Dualité dans les groupes et marche aléatoire sur $\mathbb{Z}/n\mathbb{Z}$ (102).

Automorphismes de $\mathfrak{S}_n$ pour $n \neq 6$ (103,104,105,108).

Théorème de Dirichlet faible (102,120,121).

Sous-groupes finis de $\mathrm{GL}_n(\mathbb{Z})$ (104,144,152,153).

Nombre de cycles d'une permutation aléatoire (105,190).

Nombre d'endomorphismes nilpotents sur un corps fini (101,106,150,151,156,190).

Loi de réciprocité quadratique (120,121,123,170).

Équation de Mordell (122,127).

Forme normale de Smith (122,142,149,162).

Nombre de polynômes irréductibles sur un corps fini (123,125,141,144).

Polygones constructibles à la règle et au compas (125,127,141,148,).

Décomposition de Dunford et applications (142,150,151,153,155,156).

Déterminants de Vandermonde et de Cauchy (149).

Décomposition polaire par l'exponentielle matricielle (155).

Sous-groupes compacts de $\mathrm{GL}_n(\mathbb{R})$ (106,157,181).

Problème de Sturm-Liouville (201,208,213,221,234).

Théorème de Banach-Steinhaus et séries de Fourier (201,205,208).

Théorème d'Hadamard Lévy (203,204,206,214,215,220).

Théorème de Charkovski (204,226,228).

Optimisation dans un espace de Hilbert (203,205,213,219,229,253).

Théorème de Fejér (209,234,246).

Théorème taubérien fort (209,230,235,241,243).

Lemme de Morse (214,215,218).

Théorème limite central pour les densités (218,224,228,250,262).

Théorème de Lyapounov (220,221).

Processus de Galton-Watson (223,224,226,229,262,264).

Méthode de Laplace (236,239).

Formule sommatoire de Poisson (235,239,241,246,250).

Formule des compléments (236,245).

Indécomposabilité de la loi de Poisson (243,245,261,264).

Asymptotique des nombres eulériens par les probabilités (261,266).

Séries de variables aléatoires (223,230,262,266).

Décomposition en valeurs singulières (148,152,157,158,159,162,170,191,219).

Déterminant de Gram (149,161,206).

Théorème de Krein-Milman (159,181,191,253).