

Rapport du jury

Cette leçon offre aux candidates et candidats une multitude d'approches possibles : utilisation de la topologie, du calcul différentiel, de la convexité (fonctions convexes, projection sur un convexe fermé et leurs multiples applications), de l'holomorphie. Les candidates et candidats peuvent proposer des problèmes d'optimisation sous contraintes, si possible autres que la preuve de l'inégalité arithmético-géométrique. À ce sujet, une bonne compréhension de la méthode des multiplicateurs de Lagrange requiert celle de la notion d'espace tangent, qui en donne une justification beaucoup plus claire que certains raisonnements purement matriciels. Les algorithmes de recherche d'extremums ont également leur place dans cette leçon (méthode de Newton, du gradient à pas optimal, problème des moindres carrés, etc). Les candidates et candidats solides peuvent s'intéresser aux diverses versions du principe du maximum (fonctions holomorphes ou harmoniques, équations aux dérivées partielles), au calcul des variations, ou réfléchir à l'unicité de la meilleure approximation dans divers espaces fonctionnels, à commencer par celle des fonctions continues sur un segment par des polynômes de degré au plus égal à d .

Références

- [Calcul différentiel - El Amrani](#)
- [Analyse de Fourier dans les espaces fonctionnels - M. El Amrani](#)
- [L'oral à l'agrégation de mathématiques - Isenmann et Pecatte](#)
- [Oraux X/ENS, algèbre 3](#)
- [Chabanol et Ruch, Probabilités et statistiques pour l'épreuve de modélisation](#)

Développements

- [Optimisation dans un Hilbert](#)
- [Théorème du min-max de Courant-Fisher](#)

Méta-plan

1. Extrema. Problèmes d'extrema

2. Calcul différentiel et recherche d'extrema

1. Points critiques
2. Conditions d'ordre 2
3. Optimisation sous contrainte

3. Utilisation de la convexité

1. Optimisation et fonctions convexes
2. Méthodes hilbertiennes

Plan détaillé

I - Extrema. Problèmes d'extrema [El Amrani calcul diff], [Chabanol-Ruch] [Oraux X/ENS]

Le but de cette section introductive est de présenter la notion d'extremum et de montrer que plusieurs problèmes se reformulent comme des problèmes d'extremum.

1. Définition : soit X un ensemble non vide. Soit $f : X \rightarrow \mathbb{R}$.

1. On dit que $x^* \in X$ est un maximum global de f si pour tout $x \in X$ $f(x) \leq f(x^*)$.
2. Si X est muni d'une structure métrique, on dit que x^* est un maximum local de f s'il existe U un voisinage de x^* tel que x^* est un maximum global de $f|_U$.

On définit de manière analogue la notion de minimum. Quand les inégalités sont strictes, on parle d'extremum strict.

2. Exemple : La fonction polynomial de degré 4 interpolant les points $(-2, 0)$, $(-1, -1)$, $(0, 0)$, $(1, -2)$ et $(2, 0)$ admet deux minima locaux, un seul minimum global, un maximum local et aucun maximum global. [Dessin en annexe]

3. Théorème : soit $f : K \rightarrow \mathbb{R}$ une fonction continue avec K un espace métrique compact. Alors f admet un minimum et un maximum global.

4. Application : soit I un intervalle compact et E un espace vectoriel de fonctions sur I contenant les fonctions polynomiales. Pour toute norme sur E , la fonction $P \mapsto \|f - P\|$ admet un minimum global sur $\mathbb{R}_n[X]$ pour toute fonction $f \in E$. Un tel minimum est dit "polynôme de meilleure approximation de degré n pour la norme $\|\cdot\|$ ".

5. Exemple : On considère la fonction de Heaviside, appartenant à $L^1([-1, 1])$. Pour toute constante $0 \leq a \leq 1$, la fonction constante égale à a est un polynôme de meilleure approximation de degré 0 de H pour la norme L^1 .
6. Exemple (régression linéaire des moindres carrés) : on considère une famille de points $(x_i, y_i) \in \mathbb{R}^k \times \mathbb{R}$. Une fonction de régression affine au sens des moindres carrés est une fonction de la forme $x \mapsto \langle x, \theta^* \rangle + b^*$ telle que $(\theta^*, b^*) \in \operatorname{Argmin} \sum (y_i - \langle x_i, \theta \rangle)^2$.
7. Définition: On considère un modèle statistique paramétrique $(\mathcal{X}, \mathfrak{F}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ avec $\Theta \subset \mathbb{R}^d$. On suppose qu'il existe une mesure de référence μ sur $\mathcal{X}$ telle que pour tout $\theta \in \Theta$, la loi $\mathbb{P}_\theta$ admette une densité f_θ contre μ . L'estimateur $x \mapsto \hat{\theta} \in \operatorname{Argmax} \{f(\theta; x)\}$ est appelé estimateur du maximum de vraisemblance de θ lorsqu'il est bien défini.
8. Exemple : dans le modèle uniforme $[0, \theta]^{\otimes n}$, $\hat{\theta}_{MV} = \max X_i$. C'est un estimateur fortement consistant tel que $n(\theta - \hat{\theta})$ converge en loi vers une loi exponentielle.
9. Définition (quotient de Rayleigh) : si u est un endomorphisme auto-adjoint, on appelle quotient de Rayleigh de u l'application $R_u : x \mapsto \frac{\langle u(x), x \rangle}{\|x\|^2}$.
10. Théorème (min-max de Courant-Fischer) : soit u un endomorphisme auto-adjoint d'un espace euclidien ou hermitien E de dimension d . Notons $\lambda_1(u) \leq \dots \leq \lambda_d(u)$ ses valeurs propres réelles. Alors les fonctions λ_i sont données par les formules ... (DVT 1)
11. Application : les fonctions λ_i sont 1-lipshitziennes sur l'ensemble des endomorphismes auto-adjoints. (DVT 1)

II - Calcul différentiel et recherche d'extrema [El Amrani Calcul diff]

On suppose pour cette partie connue les résultats généraux de calcul différentiel

Tous les espaces vectoriels considérés dans cette partie seront sur le corps des réels.

1 - Points critiques

1. Définition (point critique) : soit U un ouvert d'un espace de Banach E et $f : U \rightarrow \mathbb{R}$ une application différentiable. On dit que $a \in U$ est un point critique de f si $df(a) = 0$.

2. Exemples :

1. 0 est un point critique de $x \mapsto x^3$
2. Un point critique du quotient de Rayleigh d'un endomorphisme auto-adjoint u est nécessairement un vecteur propre de u
3. Proposition : si $f : U \rightarrow \mathbb{R}$ est différentiable et si $x^* \in U$ est un extremum local de f , alors x^* est un point critique de f .
4. Contre-exemples :
 1. 0 n'est pas un extremum local de $x \mapsto x^3$
 2. Il est indispensable que U soit ouvert. Si $M \subset U$ est un sous-ensemble de U , la question d'optimiser f sur M est qualifiée "d'optimisation sous contrainte" et ne peut être résolue simplement à l'aide des points critiques. Par exemple, avec $M = \{1, 2\}$, le maximum de $x \mapsto x^2$ sur M ne correspond pas à un point critique de f .
5. Application : si u est un endomorphisme auto-adjoint d'un espace euclidien, il admet $\dim(E)$ valeurs propres réelles et est diagonalisable en base orthonormée.

2 - Conditions d'ordre 2

1. Proposition : soit $f : U \rightarrow \mathbb{R}$ une fonction deux fois différentiable en x^* , avec U un ouvert d'un espace de Banach. Si x^* est un minimum local de f , alors la forme quadratique $d^2f(x^*)$ est positive.
2. Contre-exemple : on considère la fonction $f : x \mapsto -e^{-1/x^2}$, prolongée en 0 par continuité. Alors 0 est un point critique de f en lequel d^2f est nulle, donc positive, mais n'est pas un minimum de f .
3. Proposition : soit $f : U \rightarrow \mathbb{R}$ une fonction deux fois différentiable sur U . On suppose que $x^* \in U$ est un point critique de f et que $d^2f(x^*)$ est positive et coercive. Alors x^* est un minimum local strict de f .
4. Remarque : en dimension finie, il suffit que $d^2f(x^*)$ soit définie positive.
5. Exemple : la fonction $f : (x, y) \mapsto x^2 + y^2 - 2xy - y$ admet un minimum local en $(0, 0)$.

3 - Optimisation sous contraintes

| Danger avec les sous-variétés.

1. Définition (sous-variété de $\mathbb{R}^n$) : Un ensemble $M \subset \mathbb{R}^n$ est appelé sous-variété de classe C^k et de dimension d si pour tout $x \in M$, il existe U_x un voisinage ouvert de x dans $\mathbb{R}^n$ et $h : U_x \rightarrow \mathbb{R}^{n-d}$ une fonction de classe C^k telle que $dh(x)$ est surjective et $M \cap U_x = h^{-1}(\{0\})$.
2. Exemples :
 1. la sphère unité de $\mathbb{R}^n$ est une sous-variété de classe C^∞ et de dimension $n - 1$.
 2. Une conique non-vide et non-dégénérée du plan est une sous-variété de dimension 1 et de classe C^∞ .
3. Définition (plan tangent à une sous-variété)
4. Proposition : pour $x \in M$, h et U_x comme dans la définition de sous-variété, on a $T_x M = \text{Ker}(h) = \text{Vect}(\nabla h_i)^\perp$, où les h_i sont les composantes de h .
5. Théorème (extrema liés) : soit M une sous-variété de classe C^1 et de dimension d de $\mathbb{R}^n$. Soit f une fonction de classe C^1 à valeurs réelles définie sur un voisinage ouvert de M . Si $x^* \in M$ est un extremum local de $f|_M$, alors $\nabla f(x^*) \in (T_{x^*} M)^\perp$. [Faire un dessin en annexe]
6. Exemple : étude des extrema de $f : (x, y, z) \mapsto x + z$ sur l'intersection de la sphère unité de $\mathbb{R}^3$ et du plan d'équation $z - y = 0$.

III - Optimisation convexe

1 - Extrema d'une fonction convexe [El Amrani Calcul diff]

1. Définition (fonction (strictement) convexe / concave)
2. Proposition : soit f une fonction convexe. Si x^* est un extremum de f , c'est un minimum global. Si de plus f est strictement convexe, un tel minimum est unique.
3. Application : Soit $f \in C^0(I)$ avec I un intervalle compact. Pour tout n , il existe un unique polynôme de degré au plus n tel que $\|f - P\|_2 = d_2(f, \mathbb{R}_n[X])$.
4. Proposition : soit $f : C \rightarrow \mathbb{R}$ une fonction convexe sur C un convexe d'intérieur non vide. On suppose que f est différentiable sur l'intérieur de C et admet un point critique en $x^* \in \overset{\circ}{C}$. Alors x^* est un minimum global de f .
5. Application : Dans le modèle statistique de Poisson, l'estimateur du maximum de vraisemblance est défini et unique. Il correspond à l'unique point critique du logarithme de la vraisemblance.

6. Théorème : soit C un ouvert convexe de $\mathbb{R}^d$ et $f : C \rightarrow \mathbb{R}$ une fonction deux fois différentiable. Alors :
1. f est convexe sur C ssi $d^2 f$ est une forme quadratique positive
 2. si $d^2 f$ est définie positive sur C , f est strictement convexe.
7. Contre-exemple : $x \mapsto x^4$ est strictement convexe sur $\mathbb{R}$ bien que sa dérivée seconde s'annule en 0.

2 - Méthodes hilbertiennes [El Amrani Fourier, Isenmann-Pecatte]

On fixe $(E, \langle \cdot, \cdot \rangle)$ un $\mathbb{R}$ -espace préhilbertien.

1. Théorème (projection sur un convexe complet) : Soit C un convexe complet non-vidé de E . Pour tout $x \in E$, il existe un unique $p_C(x) \in C$ tel que $\|x - p_C(x)\| = d(x, C)$. De plus, $p_C(x)$ est l'unique élément de C tel que

$$\forall y \in C, \langle x - p_C(x), y - p_C(x) \rangle \leq 0$$

2. Exemple : si F est un sous-espace vectoriel de dimension finie de E admettant une base orthonormée $(e_1, \dots, e_n)$, on trouve

$$p_F(x) = \sum_{i=1}^n \langle x, e_i \rangle e_i$$

3. Application : On définit les polynômes de Legendre $(L_n)_{n \in \mathbb{N}}$ comme l'orthonormalisé de Schmidt de la base canonique de $\mathbb{R}[X]$ pour le produit scalaire

$$\langle P, Q \rangle := \int_{-1}^1 P(x)Q(x)dx$$

Pour tout fonction L^2 sur $[-1, 1]$, le polynôme de meilleur approximation de degré n de f pour la norme L^2 est donné explicitement par :

$$P(f) := \sum_{i=0}^n \left(\int_{-1}^1 L_i(x)f(x)dx \right) L_i$$

4. Définition (suite faible convergente) :

5. Proposition :

1. une suite fortement convergente est faiblement convergente.

2. Si $x_n \rightharpoonup x^*$, on a $\|x^*\| \leq \liminf \|x_n\|$

6. Théorème (Banach-Alaoglu) : toute suite bornée d'un espace de Hilbert H admet une sous-suite faiblement convergente (DVT 2)
7. Remarque : en dimension finie, ce théorème est équivalent à dire que les compacts sont les fermés bornés. Il s'agit donc d'une généralisation aux espaces de Hilbert du théorème de Bolzano-Weierstrass.
8. Application : Soit $J : H \rightarrow \mathbb{R}$ une fonction continue, convexe et coercive. J admet un minimum global sur H . (DVT 2)