Leçon 219 : Extremums : existence, caractérisation, recherche. Exemples et applications.

Développements :

Algorithme du gradient à pas optimal, Ellipsoïde de John-Loewner.

Bibliographie :

Rouvière, Gourdon, Pommellet, Rombaldi analyse réelle, OA, Bernis, Hirsch Lacombe, Allaire Kaber.

Rapport du jury 2017 :

Comme souvent en analyse, il peut être opportun d'illustrer dans cette leçon un exemple ou un raisonnement à l'aide d'un dessin. Il faut savoir faire la distinction entre propriétés locales (caractérisation d'un extremum) et globales (existence par compacité, par exemple). Dans le cas important des fonctions convexes, un minimum local est également global. Les applications de la minimisation des fonctions convexes sont nombreuses et elles peuvent illustrer cette leçon. L'étude des algorithmes de recherche d'extremums y a toute sa place : méthode de gradient, preuve de la convergence de la méthode de gradient à pas optimal, . . . Le cas particulier des fonctionnelles sur $\mathbb{R}^n$ de la forme $1/2 < Ax|x > - < b|x >$, où A est une matrice symétrique définie positive, ne devrait pas poser de difficultés. Les problèmes de minimisation sous contrainte amènent à faire le lien avec les extremums liés, la notion de multiplicateur de Lagrange et, là encore, des algorithmes peuvent être présentés et analysés. À ce sujet, une preuve géométrique des extrema liés sera fortement valorisée par rapport à une preuve algébrique, formelle et souvent mal maitrisée. Les candidats pourraient aussi être amenés à évoquer les problèmes de type moindres carrés, ou, dans un autre registre, le principe du maximum et ses applications.

Rapport du jury 2018 :

Comme souvent en analyse, il est tout à fait opportun d'illustrer dans cette leçon un exemple ou un raisonnement à l'aide d'un dessin. Il faut savoir faire la distinction entre propriétés locales (caractérisation d'un extremum) et globales (existence par compacité, par exemple). Dans le cas important des fonctions convexes, un minimum local est également global. Les applications de la minimisation des fonctions convexes sont nombreuses et elles peuvent illustrer cette leçon. L'étude des algorithmes de recherche d'extremums y a toute sa place : méthode du gradient et analyse de sa convergence, méthode à pas optimal,... Le cas particulier des fonctionnelles sur $\mathbb{R}^n$ de la forme $1/2(Ax|x)-(b|x)$ où A est une matrice symétrique définie positive, ne devrait pas poser de difficultés (la coercivité de la fonctionnelle pose problème à de nombreux candidats). Les problèmes de minimisation sous contrainte amènent à faire le lien avec les extrema liés et la notion de multiplicateur de Lagrange. À ce sujet, une preuve géométrique des extrema liés sera fortement valorisée par rapport à une preuve algébrique, formelle et souvent mal maîtrisée. Enfin, la question de la résolution de l'équation d'Euler-Lagrange peut donner l'opportunité de mentionner la méthode de Newton. Les candidats pourraient aussi être amenés à évoquer les problèmes de type moindres carrés, ou, dans un autre registre, le principe du maximum et ses applications.

Remarque 1. *Ici, on se donne $f : E \to \mathbb{R}$ avec E un $\mathbb{R}$-ev normé. On cherche à minimiser/maximiser f sur une partie C de E. Quitte à regarder $-f$, on peut parfois ne regarder que les minima.*

Remarque 2. *Cadre : E est un $\mathbb{R}$-evn de dimension finie, $C \subset E$, $f : E \to \mathbb{R}$, $a \in C$.*

1 Existence et unicité

1.1 Définitions

Définition 3 (Rouvière p370)**.** *Maximum local, global, strict (resp minimum).*

Remarque 4. *Extremum signifie maximum ou minimum.*

Exemple 5. *Une fonction admettant un minimum local non strict.* $(x, y) \mapsto |x|$.
Une fonction admettant un minimum global strict. Une norme.

Exemple 6. $\sin$ *admet un minimum local et global (non strict).*

1.2 Compacité

Proposition 7 (Gourdon p31)**.** *Une application continue sur un compact est bornée et atteint ses bornes.*

Application 8. *En dimension finie, toutes les normes sont équivalentes. (?)*

Application 9 (Gourdon p34)**.** *Si $d(f(x), f(y)) < d(x, y)$ où $f : E \to E$ avec E compact alors f admet un unique point fixe.*

Application 10 (Gourdon p33)**.** *[Pommellet p296] Pour C un compact de E et F un fermé de E disjoint de C, la fonction $x \in C \mapsto d(x, F)$ atteint son minimum $d(C, F)$. (La distance d'un point a à une partie fermée non vide d'un evn de dimension finie est atteinte.*

Contre exemple 11 (Gourdon p33)**.** *Contre exemple en dimension infinie.*
Prendre E l'ensemble des suites munies de la norme infinie, $F = \{X_n, n \in \mathbb{N}\}$ où X_n est la suite dont tous les termes sont nuls sauf le n-ième qui vaut $1+1/2n$, et $K = \{0\}$. Alors $d(K, F) = 1$ et pourtant $d(K, X_n) = 1+1/2n > 1$.

Application 12 (Pommellet p296)**.** *[OA p30] Polynôme de meilleure approximation.*

Proposition 13 (Pommellet p295)**.** *[Rouvière] Une fonction continue et coercive est minorée et atteint son minimum.*

Exemple 14. *Toute norme sur E est coercive.*

Exemple 15. $x \mapsto \frac{1}{2} < Ax, x > - < b, x >$ *est coercive.*

Contre exemple 16. *Une forme linéaire n'est pas coercive.*

Contre exemple 17 (Allaire p293)**.** *Contre exemple en dimension infinie : Prendre $f : x \in l^2 \mapsto (||x||r - 1)2 + \sum |x_n|/(n + 1)$.*

Proposition 18. *Pour K compact, F fermé, il existe x, y tels que $d(x, y) = d(K, F)$.*

1.3 Convexité

Proposition 19 (Romb p242)**.** *Si une fonction convexe sur I admet un minimum local en un point a de I alors ce minimum est global.*

Contre exemple 20. *Double puits.*

Remarque 21. *Par contre, pour une fonction convexe, il n'y a pas forcément existence d'un minimum ($x \mapsto x$) ni unicité quand il existe ($x \mapsto 0$).*

Proposition 22 (OA p30)**.** *Si f est une application strictement convexe sur un convexe C alors il existe au plus un point x de C minimisant f sur C.*

Contre exemple 23. *Mais pour une fonction strictement convexe, il n'y a pas forcément existence d'un minimum :* $\exp(x)$.

Exemple 24. $x \mapsto \frac{1}{2} < Ax, x > + < b, x >$ *admet un unique minimum.*

Proposition 25 (OA p26)**.** *Les ensembles de niveau d'une fonction convexe sont convexes.*

Application 26 (OA p30)**.** *L'ensemble des points réalisant le minimum est un convexe.*

Proposition 27 (Bernis)**.** *Ellipsoïde de John Loewner.*

Application 28. *Sous-groupes maximaux de $GL_n(\mathbb{R})$.*

1.4 Extremum sur un Hilbert

Théorème 29 (Hirsch p79)**.** *Théorème de projection sur un convexe fermé (minimise la distance) : Soit C un convexe fermé. Alors pour tout $x \in E$ il existe un unique $p(x) \in C$ tel que pour tout $y \in C$ on ait : $< x-y, p(x)-y > \leq 0$.*
Le projeté minimise aussi la distance : $||x - p(x)|| = d(x, C)$.

Proposition 30 (Hirsch p80)**.** *On a $|p(x) - p(y)| \leq |x - y|$. L'application de projection est 1-Lipschitzienne, donc continue sur E. (Inutile ?)*

Proposition 31 (Hirsch p80)**.** *Si C est un sev fermé, alors p est linéaire, donc linéaire continue. Détailler plus le théorème.*

Application 32. *Pour F un sev de E, $E = \overline{F} \oplus F^{\perp}$.*

Application 33 (Rouvière p384)**.** *Moindres carrés.*

Application 34 (FGN analyse 2 p22)**.** *Minimisation d'une intégrale.*

Application 35 (Testard p265)**.** *Théorème de Motzkin. Si $U \subset \mathbb{R}^n$ est un fermé. Alors on a : U est convexe si et seulement si $\forall x \in U^c, y \in U \mapsto ||y-x||$ atteint son minimum en un unique point de U.*

Théorème 36 (Hirsch p84)**.** *Théorème de Lax Milgram dans le cas d'une fonctionnelle continue.*

Application 37. *Une EDP.*

Proposition 38. *Optimisation dans un Hilbert : Soit H un espace de Hilbert, C une partie non-bornée de H, et $f : C \to \mathbb{R}$ une fonction continue, coercive, et convexe. Alors f atteint son minimum dans C.*

1.5 Extremum et propriété de la moyenne

Remarque 39. *On suppose pour cette partie que $E = \mathbb{C}$ et que f est holomorphe sur $\Omega \in \mathbb{C}$, avec Ω connexe.*

Définition 40 (OA p72)**.** *Propriété de la moyenne.*

Exemple 41 (OA p72)**.** *Les fonctions holomorphes vérifient la propriété de la moyenne.*

Proposition 42 (OA p72)**.** *Principe du maximum local.*

Proposition 43 (OA p72)**.** *Principe du maximum global.*

Contre exemple 44 (OA p73)**.** *Il peut exister des minima locaux : $z \mapsto a$ sur $D(0, 1)$.*

Application 45 (OA p80)**.** *Si f est holomorphe, $f(0) = 1$, et $|f(z)| \geq 2$ sur le cercle unité alors f s'annule sur le disque unité.*

Application 46 (Testard p231)**.** *[Pabion] Lemme de Schwarz.*

Application 47 (Pabion p79)**.** *Théorème de d'Alembert Gauss.*

2 Caractérisation et propriétés d'un extremum local

2.1 Conditions du premier ordre

Proposition 48 (OA p16)**.** *Condition nécessaire de minimalité locale.*

Définition 49 (OA p17)**.** *Point critique.*

Exemple 50 (Pommellet p298)**.** *U ouvert relativement compact de $\mathbb{R}^n$, $f \in C(\overline{U})$ nulle sur la frontière de U. Alors il existe $a \in U$ tel que $df(a) = 0$.*

Contre exemple 51 (OA p17)**.** *x^3 dont la dérivée et la dérivée seconde s'annulent en 0 mais qui n'a pas de minimum en 0.*

Contre exemple 52 (OA p17)**.** *Faux si pas ouvert : $x \in [0,1] \mapsto t$.*

Exemple 53. *Ex de détermination à la main d'un extremum (on teste tous les points critiques) Par exemple, $f(x,y) = x^2 + xy + y^2 - 3x - 6y$.*

Pour $f(x,y) = xy$, $(0,0)$ est un point critique mais n'est pas un extremum car $f(x,x) = x^2 > 0$ mais $f(x,-x) < 0$.

Application 54 (OA p17)**.** *Théorème de Rolle.*

Application 55 (OA p17)**.** *Théorème de Darboux.*

Remarque 56. *Ces deux théorèmes permettent de chercher les points critiques.*

Application 57 (Rouvière)**.** *La suite des points de Fermat. ?*

Proposition 58 (OA p30)**.** *Réciproque vraie dans un cas particulier : Si la fonction f est convexe sur I et différentiable en un point a de l'intérieur tel que $df(a) = 0$ alors f admet un minimum global en a.*

2.2 Conditions du second ordre

Proposition 59 (OA p18)**.** *Conditions nécessaires du second ordre.*

Remarque 60 (Pommellet p298)**.** *Séparer conditions nécessaires et suffisantes du second ordre.*

Contre exemple 61 (OA p18)**.** *$x \mapsto x^3$, $(x,y) \mapsto x^2 - y^3$, $(x,y) \mapsto x^2 + y^4$.*

Proposition 62 (Gourdon p317)**.** *[Pommellet p299] Notations de Monge et étude de $rt - s^2$ et r pour déterminer les extrema.*

Cas particulier de $n = 2$ Soit a un point critique de f.

1. Si $rt - s^2 > 0$ (i.e les deux vp de $grad(f(a))$ sont $\neq 0$ et de même signe

— Si $r + t > 0$ alors la matrice $grad(f(a))$ est définie positive et donc a est un minimum global

— Si $r + t < 0$ alors la matrice $grad(f(a))$ est définie négative et donc a est un maximum global

2. Si $rt - s^2 < 0$ (i.e les deux vp de $gradf(a)$ sont $\neq 0$ et de signe contraire. Alors a n'est un extremum local (si on note $\lambda_1 < 0$ et $\lambda_2 > 0$ les vps associés aux vecteurs propres h_1 et h_2. Alors $< gradf(a)h_1, h_1 >= \lambda_1 ||h_1||^2 < 0$ et donc la matrice $gradf(a)$ n'est pas positive, a n'est pas un minimum local. Et en utilisant h_2, on montre que a ne peut pas être un maximum local.

3. Si $rt - s^2 = 0$, on ne sait pas conclure.

Exemple 63 (Gourdon p317)**.** *$f(x,y) = x^4 + y^4 - 2(x-y)^2$.*

2.3 Optimisation sous contrainte

Remarque 64. *On s'intéresse au problème $inf_K f$ où K n'est pas ouvert (et donc on ne peut pas appliquer les théorèmes de localisation précédents)*

Théorème 65 (Gourdon)**.** *Théorème des extrema liés.*

Application 66 (Gourdon p321)**.** *$SO_n(\mathbb{R})$ est l'ensemble des éléments de $SL_n(\mathbb{R})$ de norme minimale.*

Application 67 (Gourdon p319)**.** *Inégalité arithmético-géométrique.*

Application 68 (OA p21)**.** *Diagonalisation des endomorphismes symétriques.*

Application 69 (Rouvière p410)**.** *Inégalité de Hadamard.*

Application 70 (Rouvière)**.** *Mise en boîte optimale. Obtenir un parallélépipède rectangle d'aire minimum et de volume donné*

Application 71. *Minimisation de la fonction $f(x,y) = x^2 - y$ pour des vecteurs de norme 1.*

3 Recherche numérique d'extrema

3.1 Méthode de Newton

Proposition 72 (Rouvière p152)**.** *[Allaire p348][OA p23] Méthode de Newton. Chercher dans le Allaire.*

Remarque 73 (Allaire p348)**.** *Pour trouver les points critiques, on applique la méthode de Newton à f'*

Remarque 74. *Inconvénient : système linéaire à résoudre à chaque étape, nécessité de connaître la Hessiene, d'être au départ proche de la solution. Cette méthode nécessite la résolution d'un système linéaire à chaque étape. En contrepartie, la convergence est quadratique.*

3.2 Méthodes de descentes

Remarque 75. *Soit U ouvert de $\mathbb{R}^p$, soit $J \in C^1(U, \mathbb{R})$ admettant un unique minimum en x.*

Remarque 76 (OA p22)**.** *On cherche à construire une suite convergeant vers le minimum de f. Pour cela, on fixe $x_0 \in U$ et un critère d'arrêt $\epsilon > 0$, et tant que $||x_n - x|| \geq \epsilon$, on calcule un terme supplémentaire $x_{n+1} = x_n - t_n grad(J(x_n))$, où $t_n > 0$ est appelé pas et $d_n \in \mathbb{R}^p$ est la direction de descente. On parle de méthode de gradient lorsque $d_n = gradf(x_n)$, à pas constant lorsque $t_n = a =$constante.*

Remarque 77. *On peut aussi choisir un pas maximisant la descente, en prenant pour tout $n \in \mathbb{N}$, $t_n = argmin(\mu > 0) J(x_n + \mu gradf(x_n))$. Cette méthode est appelée gradient à pas optimal.*

Proposition 78 (Oa p22)**.** *Quand est ce qu'ils convergent...*

3.3 Application : résolution de systèmes linéaires

Remarque 79 (Romb analyse matricielle p214)**.** *On considère le système linéaire $Ax = b$, d'inconnue $x \in \mathbb{R}^p$, où $A \in M_{p,n}(\mathbb{R})$, $b \in \mathbb{R}^n$ donnés.*

Proposition 80. *Si $A \in S_n^{++}$, pour $x_0 \in \mathbb{R}^n$, la suite définie par $x_{k+1} = x_k - \alpha gradf(x_k) = x_k + \alpha(b - Ax_k)$ converge vers la solution de $Ax = b$.*

Proposition 81. *Algorithme du gradient à pas optimal.*

Définition 82 (Allaire Kaber)**.** *Si le système est incompatible, le problème aux moindres carrés revient à chercher $x = argmin_{x \in \mathbb{R}^n} ||Ax - b||^2$.*

Proposition 83. *x est solution du problème aux moindres carrés, si et seulement si il vérifie $A^*Ax = b$.*

Théorème 84. *Le problème aux moindres carrés possède toujours une solution, et celle-ci est unique si et seulement si $ker(A) = \{0\}$.*