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

[6] 1. Notation. Soit E un espace vectoriel normé. On considère une partie $X \subset E$, une fonction $f\colon X \longrightarrow \mathbf{R}$ et un point $a \in X$.

2. Définition. Le point a est un *maximum global* de la fonction f si

$$\forall x \in X, \qquad f(x) \leqslant f(a).$$

C'en est un *maximum local* s'il existe un voisinage $V \subset X$ du point a tel que

$$\forall x \in V, \qquad f(x) \leqslant f(a).$$

On définit, de même, la notion de minimum local et global (en renversant les inégalités) et de maximum et minimum stricts (en mettant des inégalités strictes lorsque $x \neq a$). Un *extremum* est un minimum ou un maximum.

I. Étude globale : critère d'existence et d'unicité

I.1. Utilisation de la compacité

[4] 3. Proposition. Soient $X \subset E$ une partie compact et $f\colon X \longrightarrow \mathbf{R}$ une fonction continue. Alors elle est bornée et atteint ses bornes.

4. Exemple. Sur tout compact $K \subset \mathbf{C}$, la fonction $z \in \mathbf{C} \longmapsto |z|$ atteint ses bornes. Sur un espace vectoriel normé de dimension finie, toutes les normes sont équivalentes.

5. Application. La distance entre deux parties compactes K_1 et K_2 d'un espace métrique E est atteinte, c'est-à-dire qu'il existe deux points $x_1 \in K_1$ et $x_2 \in K_2$ tels que $d(x_1, x_2) = d(K_1, K_2)$. Le résultat reste vrai lorsque la partie K_1 est juste fermée.

6. Application. Soient E un espace métrique compact et $f\colon E \longrightarrow E$ une application telle que

$$\forall x, y \in E, \qquad x \neq y \quad \Longrightarrow \quad d(f(x), f(y)) < d(x, y).$$

Alors elle admet un unique point fixe et toute suite des itérées de la fonction f converge vers ce point fixe.

[1] 7. Définition. Soit $X \subset E$ une partie non bornée. Une fonction $f\colon X \longrightarrow \mathbf{R}$ est *coercive* si $f(x) \longrightarrow +\infty$ lorsque $x \longrightarrow +\infty$ et $x \in X$.

8. Proposition. On suppose que l'espace E est de dimension finie. Soient $X \subset E$ une partie fermée non bornée et $f\colon X \longrightarrow \mathbf{R}$ une fonction coercive continue. Alors elle admet un minimum global.

I.2. Utilisation de la convexité

[1] 9. Définition. Soit $C \subset E$ un convexe. Une fonction $f\colon C \longrightarrow \mathbf{R}$ est convexe si

$$\forall x, y \in C,\ \forall \lambda \in [0,1], \qquad f(\lambda x + (1-\lambda) y) \leqslant \lambda f(x) + (1-\lambda) f(y).$$

Lorsque l'inégalité est stricte avec $x \neq y$ et $\lambda \in \,]0,1[$, elle est *strictement convexe*.

10. Exemple. Les fonctions $x \in \mathbf{R} \longmapsto x^2$ et $x > 0 \longmapsto -\ln x$ sont strict. convexes.

11. Proposition. Soient $C \subset E$ un convexe et $f\colon C \longrightarrow \mathbf{R}$ une fonction strictement convexe. Alors elle admet au plus un minimum.

12. Contre-exemple. La seule convexité ne suffit pas à assurer au plus un minimum (le fonction nulle sur $\mathbf{R}$). La stricte convexité n'assure pas l'existence d'un minimum (la fonction exponentielle sur $\mathbf{R}$).

13. Application. Soient E un espace euclidien, $b \in E$ un vecteur et $u \in \mathscr{L}(E)$ un endomorphisme symétrique défini positif. Alors la fonction

$$f\colon \left| \begin{array}{l} E \longrightarrow \mathbf{R}, \\ x \longmapsto \frac{1}{2}\langle u(x), x\rangle - \langle b, x\rangle \end{array} \right.$$

admet un unique point minimum.

I.3. Résultats en analyse hilbertienne

14. Théorème *(de projection)*. Soient H un espace de Hilbert et $C \subset H$ un convexe [1]
fermé non vide. Pour tout $x \in H$, il existe un unique élément $p_C(x) \in C$ tel que

$$\|p_C(x) - x\| = \mathrm{d}(x, C) = \inf_{y \in C} \|y - x\|.$$

15. Contre-exemple. Toutes les hypothèses sont nécessaires.

- L'hypothèse hilbertienne est nécessaire. Dans l'espace $(\mathscr{C}^0([0,1]), \|\ \|_\infty)$, la distance $\mathrm{d}(1, C)$ avec $C \coloneqq \{f \in \mathscr{C}^0([0,1]) \mid 0 \leqslant f \leqslant 1, f(0) = 0\}$ est réalisée par les fonctions $f \in C$.
- L'hypothèse de complétude est nécessaire. En prenant $E \coloneqq \mathscr{C}^0([0,1]) \subset \mathrm{L}^2([0,1])$ avec $C \coloneqq (\mathbf{1}_{[0,1/2]})^\perp$ et $C_1 \coloneqq C \cap E$, la distance $\mathrm{d}(f_1, C_1)$ n'est pas atteinte pour toute fonction $f_1 \in E \setminus C_1$.
- L'hypothèse de convexité est nécessaire. Dans $\mathbf{R}^2$, l'origine admet une infinité de projetés sur la sphère unité $\mathbf{S}^1 \subset \mathbf{R}^2$.

16. Application *(moindres carrés)*. Étant donnés n points $(x_i, y_i) \in \mathbf{R}^2$ tels que les [6]
réels x_i ne soient pas tous égaux entre eux, il existe $\lambda, \mu \in \mathbf{R}$ qui minimisent la somme

$$\sum_{i=1}^{n} (\lambda x_i + \mu - y_i)^2.$$

17. Corollaire *(théorème du supplémentaire orthogonal)*. Soient H un espace de [1]
Hilbert et $F \subset H$ un sous-espace vectoriel. Alors $H = F \oplus F^\perp$.

18. Théorème *(Riesz)*. Soit H un espace de Hilbert. Alors l'application

$$\left| \begin{array}{l} H \longrightarrow H', \\ y \longmapsto \langle \cdot, y\rangle \end{array} \right.$$

est une isométrie surjective.

19. Théorème. Soit H un espace de Hilbert. Alors toute suite bornée $(x_n)_{n \in \mathbf{N}}$ [3]
de H admet une sous-suite convergeant faiblement, c'est-à-dire qu'il existe une extraction $\varphi\colon \mathbf{N} \longrightarrow \mathbf{N}$ et un vecteur $x \in H$ tels que

$$\forall y \in H, \qquad \langle x_{\varphi(n)}, y\rangle \longrightarrow \langle x, y\rangle.$$

20. Proposition. Soient H un espace de Hilbert et $C \subset H$ une partie convexe non bornée. Soit $J\colon C \longrightarrow \mathbf{R}$ une fonction convexe, continue et coercive. Alors cette dernière atteint sa borne inférieure. D1

21. Remarque. La proposition permet de généraliser l'application 13 à un espace [1]
de Hilbert H. Soient $b \in H$ un vecteur et $u \in \mathscr{L}(H)$ un endomorphisme symétrique

défini positif tel que l'application $x \in H \longmapsto \langle u(x), x\rangle$ soit coercive. Alors on retrouve la conclusion du point 13.

22. Corollaire *(théorème de Lax-Milgram)*. Soient H un espace de Hilbert et a une forme bilinéaire symétrique continue coercive sur H. Soit $\varphi \in H'$ une forme linéaire continue sur H. Alors il existe un unique élément $x \in H$ tel que

$$\forall y \in H, \qquad a(x,y) = \varphi(y).$$

De plus, cet élément x est caractérisé par l'égalité

$$\frac{1}{2}a(x,x) - \varphi(x) = \min_{y\in H}\left(\frac{1}{2}a(y,y) - \varphi(y)\right).$$

[2] 23. Remarque. Le théorème de Lax-Milgram est un outil permettant d'étudier certaines équations aux dérivées partielles. Soit $f\colon]0,1[\longrightarrow \mathbf{R}$ une fonction. On considère le problème de Dirichlet

$$\begin{aligned} -u'' + u &= f \quad \text{sur }]0,1[, \\ u(0) = u(1) &= 0. \end{aligned} \tag{1}$$

Lorsque $f \in \mathrm{L}^2(]0,1[)$, le problème (1) admet une unique solution faible dans $\mathrm{H}_0^1(]0,1[)$ et c'est la fonction minimisant la fonction

$$v \in \mathrm{H}_0^1(]0,1[) \longmapsto \frac{1}{2}\int_0^1 (v'^2 + v^2) - \int_0^1 fv \in \mathbf{R}.$$

I.4. Holomorphie et principe du maximum

[1] 24. Théorème *(principe du maximum global)*. Soit $\Omega \subset \mathbf{C}$ un ouvert borné. Soit $f\colon \overline{\Omega} \longrightarrow \mathbf{C}$ une fonction continue sur $\overline{\Omega}$ et holomorphe sur Ω. Alors la fonction $|f|$ atteint son maximum sur la frontière $\partial\Omega$ et, si son maximum est atteint sur Ω, alors elle est constante sur Ω.

25. Corollaire *(Liouville)*. Toute fonction entière bornée est constante.

26. Application *(théorème de d'Alembert-Gauss)*. Le corps des complexes est algébriquement clos.

II. Étude locale : critère d'existence par le calcul différentiel

27. Notation. Soient $U \subset E$ un ouvert et $f\colon U \longrightarrow \mathbf{R}$ une fonction.

II.1. Condition du premier ordre

[6] 28. Rappel. On suppose que la fonction f est différentiable en un point $a \in U$. Lorsque $h \longrightarrow 0$, on a

$$f(a+h) = f(a) + df(a)(h) + o(\|h\|).$$

29. Proposition. Soit $x^\star \in U$. On suppose que la fonction f admet un minimum et est différentiable en ce point $x^\star$. Alors $df(x^\star) = 0$.

30. Contre-exemple. La condition est loin d'être nécessaire puisque la fonction cube $x \in \mathbf{R} \longmapsto x^3$ voit sa dérivée s'annuler au point $x^\star = 0$ et, pourtant, ce point n'est pas un minimum.

[1] 31. Application. L'unique minimum de l'application f définie au point 13 est atteint au point $x^\star \in E$ vérifiant $u(x^\star) = b$.

32. Théorème *(Rolle)*. Soit $f\colon [a,b] \longrightarrow \mathbf{R}$ une fonction continue sur $[a,b]$ et déri- [4]
vable sur $]a,b[$ avec $f(a) = f(b)$. Alors il existe $c \in]a,b[$ tel que $f'(c) = 0$.

33. Proposition *(inégalité d'Euler)*. Soient $C \subset U$ une partie convexe et $f\colon U \longrightarrow \mathbf{R}$ [1]
une fonction qui admet un minimum local $x^\star \in C$ sur C et qui est différentiable en ce point $x^\star$. Alors

$$\forall y \in C, \qquad df(x^\star)(y - x^\star) \geqslant 0.$$

34. Proposition. Soient $C \subset E$ un convexe ouvert et $f\colon C \longrightarrow \mathbf{R}$ une fonction convexe différentiable. Alors tout point critique de f en est un minimum global.

35. Proposition *(point de Fermat)*. Soient A, B et C trois points non alignés du [6]
plan euclidien $\mathbf{R}^2$. On suppose que les trois angles du triangle ABC sont strictement inférieurs à $2\pi/3$. Alors la fonction D2

$$f\colon \left|\begin{aligned} &\mathbf{R}^2 \longrightarrow \mathbf{R}, \\ &M \longmapsto MA + MB + MC \end{aligned}\right.$$

admet un unique point minimum qui est dans l'intérieur strict du triangle ABC.

II.2. Condition du second ordre

36. Proposition. On suppose que la fonction f est deux fois différentiable en un [6]
point $a \in U$. Lorsque $h \longrightarrow 0$, on a

$$f(a+h) = f(a) + df(a)(h) + \frac{1}{2}\,d^2f(a)(h,h) + o(\|h\|^2).$$

37. Proposition. Soit $x^\star \in U$. On suppose que la fonction f est deux fois différentiable en ce point $x^\star$.

- Si le point $x^\star$ en est un minimum local, alors $df(x^\star) = 0$ et sa différentielle seconde $d^2f(x^\star)$ est une forme quadratique positive, c'est-à-dire
$$d^2f(x^\star)(h,h) \geqslant 0, \qquad \forall h \in E.$$
- Si $df(x^\star) = 0$ et sa différentielle seconde $d^2f(x^\star)$ est définie positive, alors le point $x^\star$ est un minimum local strict de la fonction f.

38. Contre-exemple. Les réciproques des deux points sont fausses. Pour le premier [1]
point, la fonction $(x,y) \in \mathbf{R}^2 \longmapsto x^2 - y^2$ admet un unique point critique qui est l'origine et, en ce point, sa hessienne est positive, mais l'origine n'est pas un minimum local. On considère le contre-exemple $(x,y) \in \mathbf{R}^2 \longmapsto x^2 + y^2$ pour le second point.

II.3. Extrema liés

39. Théorème. Soient $g_1, \dots, g_m\colon \mathbf{R}^n \longrightarrow \mathbf{R}$ des applications de classe $\mathscr{C}^1$. On pose [1]

$$C := \{x \in \mathbf{R}^n \mid g_1(x) = \dots = g_m(x) = 0\}.$$

Soit $U \subset E$ un ouvert vérifiant $U \supset C$. Soit $f\colon U \longrightarrow \mathbf{R}$ une application admettant un extremum local en un point $x^\star \in U$ et différentiable en ce point. On suppose que les différentielles $dg_i(x^\star)$ avec $i \in [\![1,m]\!]$ sont linéairement indépendantes. Alors il existe des réels $\lambda_1, \dots, \lambda_m \in \mathbf{R}$ tels que

$$df(x^\star) = \lambda_1\, dg_1(x^\star) + \dots + \lambda_m\, dg_m(x^\star).$$

Les réels λ_i sont appelés les *multiplicateurs de Lagrange*.

40. Contre-exemple. L'hypothèse d'indépendance linéaire est nécessaire : la fonc-

tion $f\colon (x,y) \in \mathbf{R}^2 \longmapsto x + y^2$ admet un unique minimum, qui est l'origine, sous la contrainte $g(x,y) \coloneqq x^3 - y^2 = 0$. Mais on a $dg(0,0) = 0$ et $df(0,0) \neq 0$.

41. Application *(théorème spectral en dimension finie).* Soient E un espace euclidien et $u \in \mathscr{L}(E)$ un endomorphisme symétrique. En considérant les fonctions

$$f\colon x \in E \longmapsto \langle u(x), x \rangle \qquad \text{et} \qquad g\colon x \in E \longmapsto \|x\|,$$

la fonction f atteint son minimum sur la sphère $\{g = 1\} \subset E$ en un point $x^\star \in E$ puisque c'est un compact. Le théorème 39 fournit alors une valeur propre $\lambda_1 \in \mathbf{R}$.

III. Algorithmes de recherche

42. Notation. On cherche des algorithmes permettant de chercher le minimum d'une fonction $f\colon \mathbf{R}^n \longrightarrow \mathbf{R}$ de classe $\mathscr{C}^1$.

III.1. Méthodes de gradient

[5] 43. Définition. Soit $(\rho_k)_{k\in\mathbf{N}}$ une suite réelle. La *méthode de gradient à pas variable* consiste en une suite réelle $(u_k)_{k\in\mathbf{N}}$ vérifiant

$$u_{k+1} = u_k - \rho_k \nabla f(u_k), \qquad k \in \mathbf{N}.$$

On dit que la méthode est à *pas fixe* lorsque la suite $(\rho_k)_{k\in\mathbf{N}}$ est constante.

44. Théorème. Soient $A \in \mathscr{S}_n^{++}(\mathbf{R})$ et $b \in \mathbf{R}^n$. On définit la fonction

$$f\colon \left| \begin{array}{l} \mathbf{R}^n \longrightarrow \mathbf{R}, \\ \quad x \longmapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle. \end{array} \right.$$

Soit $x^\star \in \mathbf{R}^n$ son unique minimum. Soit $(x_k)_{k\in\mathbf{N}}$ la suite réelle définie par

$$x_{k+1} = x_k - t_k \nabla f(x_k) \quad \text{avec} \quad t_k \coloneqq \operatorname*{arg\,min}_{t \geqslant 0} f(x_k - t\nabla f(x_k)), \qquad k \in \mathbf{N}.$$

Alors elle converge vers le point $x^\star$ et, plus précisément, en notant $\lambda_{\min}$ et $\lambda_{\max}$ les valeurs propres minimale et maximale de A, il existe une constante $C \geqslant 0$ telle que

$$\forall k \in \mathbf{N}, \qquad \|x_k - x^\star\| \leqslant C \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right)^k.$$

III.2. Méthode de Newton

[6] 45. Définition. Soit $f\colon]c,d[\longrightarrow \mathbf{R}$ une fonction de classe $\mathscr{C}^2$. On suppose qu'elle admet un unique zéro $a \in]c,d[$. La *méthode de Newton* consiste en, lorsqu'elle est bien définie, une suite réelle $(x_n)_{n\in\mathbf{N}}$ vérifiant

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n \in \mathbf{N}.$$

46. Théorème. Alors il existe deux constantes $C, \alpha > 0$ tel que, si $|x_0 - a| < \alpha$, alors la suite $(x_n)_{n\in\mathbf{N}}$ est bien définie, elle converge vers le point a et elle vérifie

$$\forall n \in \mathbf{N}, \qquad |x_{n+1} - a| \leqslant C(x_n - a)^2.$$

47. Corollaire. Lorsque la fonction f est strictement convexe et vérifie $f'(a) > 0$, la suite $(x_n)_{n\in\mathbf{N}}$ avec $x_0 > a$ décroît strictement et converge vers le point a.

48. Remarque. Soit $g\colon [c,d] \longrightarrow \mathbf{R}$ une fonction strictement convexe de classe $\mathscr{C}^3$. On peut ainsi approcher son minimum en appliquant la méthode de Newton à sa dérivée $f \coloneqq g'$. On peut également la généraliser à la dimension supérieure.

[1] Vincent Beck, Jérôme Malick et Gabriel Peyré. *Objectif Agrégation.* 2e édition. H&K, 2005.
[2] Haïm Brézis. *Analyse fonctionnelle.* 2e tirage. Masson, 1983.
[3] Philippe Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation.* 3e tirage. Masson, 1982.
[4] Xavier Gourdon. *Analyse.* 2e édition. Ellipses, 2008.
[5] Jean-Baptiste Hiriart-Urruty. *Optimisation et analyse convexe.* EDP Sciences, 2009.
[6] François Rouvière. *Petit guide de calcul différentiel.* Quatrième édition. Cassini, 2015.