

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

Devs :

- Théorème des extrema liés (par la géométrie différentielle)
- Méthode de gradient à pas optimal
- Inégalité isopérimétrique

Références :

1. Gourdon, Analyse
2. Stein & Shakarchi, Fourier Analysis
3. Rouvière, Petit guide du calcul différentiel
4. Objectif Agrégation
5. Avez, Calcul différentiel
6. Hiriart-Urruty, Optimisation et analyse convexe

Définitions préliminaires

On se donne X une partie d'un espace vectoriel normé $(E, \|\cdot\|)$ sur $\mathbb{R}$, a un point de X , et $f: X \rightarrow \mathbb{R}$ une fonction numérique.

Définition 1.

- On dit que f admet un maximum global si $f(x) \leq f(a)$ pour tout $x \in X$.
- On dit que f admet un maximum local s'il existe un voisinage V de a tel que $f(x) \leq f(a)$ pour tout $x \in V \cap X$.

Définition 2.

On dit que f admet un maximum (global, respectivement local) strict en a si les inégalités des définitions précédentes sont strictes.

1 Existence et unicité

1.1 Extrema dans un espace compact. Dans un espace fermé.

Théorème 3. Soit X un espace topologique compact et $f: X \rightarrow \mathbb{R}$ une application continue. Alors f atteint ses bornes sur X : il existe $a \in X$ et $b \in X$ tels que $f(a) \leq f(x)$ et $f(x) \leq f(b)$ pour tout $x \in X$.

Exemple 4. Soit (X, d) un espace métrique, K une partie compacte de X et F une partie fermée. Si $K \cap F = \emptyset$, alors $d(K, F) > 0$, où $d(K, F) := \inf_{(x, y) \in K \times F} d(x, y)$.

Définition 5. On dit qu'une application $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est coercive sur un espace vectoriel normé $(E, \|\cdot\|)$ si $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$.

Exemple 6. Une norme est coercive. Une forme linéaire ne l'est pas.

Théorème 7. Soit $(E, \|\cdot\|)$ un espace vectoriel normé, $C \subset E$ un fermé non vide, et $f: C \rightarrow \mathbb{R}$ une application coercive et continue. Alors il existe $m \in S$ tel que $f(s) = \inf_{x \in S} f(x)$.

Application 8. (Théorème d'Alembert Gauss)

Soit $P \in \mathbb{C}[X]$ non constant. Alors P admet au moins une racine sur $\mathbb{C}$.

1.2 Extrema d'une fonctions convexe

On se donne un espace vectoriel normé $(E, \|\cdot\|)$.

Définition 9. On dit qu'une fonction $f: E \rightarrow \mathbb{R}$ est convexe sur une partie $C \subset E$ si

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

On dit que f est strictement convexe sur C si l'inégalité ci-dessus est stricte.

Théorème 10. Si E est de dimension finie, alors une fonction convexe sur E est continue. Le résultat devient faux en dimension infinie.

Théorème 11. (Caractérisation des fonctions convexes différentiables)

Soit U un ouvert convexe de $\mathbb{R}^n$ et $f: U \rightarrow \mathbb{R}$ une application différentiable sur U . Alors les propriétés suivantes sont équivalentes :

- f est convexe.
- Le graphe de f est au-dessus de ses tangentes

$$\forall x, y \in U \quad f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

- L'application $\nabla f: U \rightarrow \mathbb{R}^n$ vérifie

$$\forall x, y \in U \quad \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0.$$

Si de plus f est deux fois différentiable sur U , elle est convexe si et seulement si sa matrice Hessienne est positive : $\forall x, h \in U \quad \langle Hf(x) \cdot h, h \rangle \geq 0$, et elle est strictement convexe si et seulement si sa matrice Hessienne est définie positive.

Exemple 12. Soit $A \in \mathcal{S}_n$. Alors $f: x \mapsto \langle x, Ax \rangle$ est convexe sur $\mathbb{R}^n$ si et seulement si A est positive. Elle est strictement convexe si et seulement si $f \in \mathcal{S}_n^{++}$.

Proposition 13. Si $f: E \rightarrow \mathbb{R}$ est convexe, alors tout minimum local est global.

Théorème 14. Soit C un convexe non vide de E et $f: C \rightarrow \mathbb{R}$ une application strictement convexe. Alors il existe au plus un minimum de f sur C .

Corollaire 15. Soit $f: E \rightarrow \mathbb{R}$ une fonction convexe continue. Alors f admet un unique minimum sur E .

1.3 Minimum dans un espace de Hilbert

On se donne H un espace de Hilbert, et on note $\|\cdot\|$ la norme dérivant du produit scalaire $\langle \cdot, \cdot \rangle$ de H .

Théorème 16. (de projection sur un convexe fermé)

Soit Γ une partie fermée, convexe, non vide de H . Alors pour tout point $x \in H$, il existe un unique point $y \in \Gamma$ tel que

$$\|x - y\| = \inf_{\gamma \in \Gamma} \|x - \gamma\| = d(x, \Gamma).$$

Ce point est appelé projection de x sur Γ et est noté $p_\Gamma(x)$. Il est caractérisé par la propriété suivante :

$$\begin{cases} p_\Gamma(x) \in \Gamma \\ \forall \gamma \in \Gamma \quad \Re \langle x - p_\Gamma(x), \gamma - p_\Gamma(x) \rangle \leq 0 \end{cases}.$$

Corollaire 17. Sous les hypothèses du théorème 12, on a

$$\forall (x_1, x_2) \in E^2 \quad \|p_\Gamma(x_1) - p_\Gamma(x_2)\| \leq \|x_1 - x_2\|.$$

En particulier, la projection sur Γ est une application 1-lipschitzienne, donc continue.

Application 18. Soit F un sous-espace vectoriel de H (non nécessairement fermé). Pour

$x \in H$, le projeté orthogonal $p_F(x)$ de x sur F est caractérisé par :

$$\begin{cases} p_F(x) \in F \\ x - p_F(x) \in F^\perp \end{cases}$$

Par ailleurs, on a l'égalité :

$$H = \bar{F} \oplus (\bar{F})^\perp = \bar{F} \oplus F^\perp.$$

On en déduit un critère de densité pour un sous-espace vectoriel d'un espace de HILBERT :

$$F \text{ est dense dans } H \iff F^\perp = \{0\}.$$

Théorème 19. (de représentation de RIESZ).

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de HILBERT. Alors l'application δ définie par

$$\mathcal{L}: \begin{cases} H & \rightarrow H' \\ a & \mapsto \mathcal{L}_a = (x \mapsto \langle x, a \rangle) \end{cases}$$

est une bijection antilinéaire isométrique.

Remarque 20. Soit u un endomorphisme continu de H et $y \in H$. La forme linéaire $\langle u(\cdot), y \rangle$ est continue, donc d'après le théorème de représentation de RIESZ, il existe un unique $z \in H$ vérifiant

$$\forall x \in H \quad \langle u(x), y \rangle = \langle x, z \rangle.$$

Définition 21. On appelle adjoint de l'endomorphisme continu $u \in \mathcal{L}(H)$ l'unique endomorphisme continu u^* de $\mathcal{L}(H)$ vérifiant

$$\forall (x, y) \in H^2 \quad \langle u(x), y \rangle = \langle x, u^*(y) \rangle.$$

Théorème 22. (de RADON-NIKODYM)

Soit $(X, \mathcal{A})$ un espace mesurable et μ, ν deux mesures positives σ -finies sur $(X, \mathcal{A})$. Il y a équivalence entre :

$$1. \forall A \in \mathcal{A} \quad \mu(A) = 0 \implies \nu(A) = 0 \quad (\text{on note } \nu \ll \mu).$$

$$2. \exists f: (X, \mathcal{A}) \rightarrow (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)) \text{ intégrable telle que } \forall A \in \mathcal{A} \quad \nu(A) = \int_A f d\mu.$$

En outre, la fonction f est unique (à une égalité μ -presque partout près).

On note $f = \frac{d\nu}{d\mu}$ et on dit que f est la dérivée de RADON-NIKODYM, ou la densité de ν par rapport à μ .

1.4 Maximum d'une fonction holomorphe

Soit Ω un ouvert de $\mathbb{C}$, et $f: \Omega \rightarrow \mathbb{C}$ une application continue.

Définition 23. On dit que f est harmonique sur Ω si $\Delta f = 0$, où le laplacien Δf est défini par

$$\Delta f := \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

Définition 24. On dit que f possède la propriété de la moyenne si pour tout disque $D(a, r)$ tel que $D(a, r) \subset \Omega$ on a

$$f(a) = \frac{1}{2\pi} \int_0^{2\pi} f(a + re^{it}) dt.$$

Théorème 25. L'application f est harmonique sur Ω si et seulement si elle vérifie la propriété de la moyenne sur Ω . En particulier, une fonction holomorphe sur Ω vérifie la propriété de la moyenne sur Ω .

Théorème 26. (Principe du maximum local)

On suppose que f vérifie la propriété de la moyenne sur Ω . Si $|f|$ admet un minimum local en $a \in \Omega$, alors elle est constante sur un voisinage de a .

Théorème 27. (Principe du maximum global)

On suppose que Ω est connexe et borné, et que f vérifie la propriété de la moyenne sur Ω . Alors

$$\forall z \in \Omega \quad |f(z)| \leq \sup_{\zeta \in \partial\Omega} |f(\zeta)|.$$

De plus, s'il existe $z_0 \in \Omega$ tel que $|f(z_0)| = M$, alors f est constante sur Ω .

Application 28. (Lemme de Schwarz)

Soit $f: D(0, 1) \rightarrow \mathbb{C}$ holomorphe. On suppose que $f(0) = 0$ et que $|f(z)| \leq 1$ pour tout $z \in D(0, 1)$. Alors

$$\forall z \in D(0, 1) \quad |f(z)| \leq |z|.$$

De plus, s'il existe $z_0 \in D(0, 1)$ non nul tel que $f(z_0) = z_0$ ou si $|f'(0)| = 1$, alors il existe $a \in \mathbb{U}$ tel que $f(z) = az$ pour tout $z \in D(0, 1)$. Autrement dit, f est une homothétie.

2 Recherche et caractérisation

On considère un ouvert U de $\mathbb{R}^n$ et une application $f: U \rightarrow \mathbb{R}$.

2.1 Condition nécessaire d'ordre 1

Théorème 29. Si f est différentiable sur U et admet un extremum relatif en un point $a \in U$, alors $df_a = 0$. Un point a pour lequel $df_a = 0$ est appelé un point critique de f .

Remarque 30. La réciproque est fautive en général.

Théorème 31. (Rolle)

Soit $f: [a, b] \rightarrow \mathbb{R}$ continue, dérivable sur $]a, b[$, telle que $f(a) = f(b)$. Alors f' s'annule au moins en un point $c \in]a, b[$.

Théorème 32. (Darboux)

Soit I un intervalle de $\mathbb{R}$ et $f: I \rightarrow \mathbb{R}$ dérivable. Alors $f'(I)$ est un intervalle.

2.2 Conditions d'ordre 2

On suppose que f est de classe $\mathcal{C}^2$ sur $U \subset \mathbb{R}^n$.

Théorème 33. (Formule de Taylor-Young à l'ordre deux)

Soit $x \in \mathbb{R}^n$ et $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ tels que $[x, x+h] \subset U$. Alors

$$f(x+h) = f(x) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) + \frac{1}{2!} \left[\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) \right]^{[2]} + o(\|h\|^2),$$

avec

$$\left[\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) \right]^{[2]} := \sum_{i=1}^n h_i^2 \frac{\partial^2 f}{\partial x_i^2}(a) + 2 \sum_{1 \leq i < j \leq n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

Corollaire 34.

- Si f admet un minimum (resp un maximum) relatif en $a \in U$, alors la forme quadratique $Q(h) = \left[\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) \right]^{[2]} = \langle Hf(h), h \rangle$ est positive (resp. négative).
- Si $Q(h) = \left[\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) \right]^{[2]} = \langle Hf(h), h \rangle$ est définie positive (resp. définie négative), alors f admet un minimum (resp. un maximum) local en a .

Remarque 35. Si Q est seulement positive, l'existence d'un minimum relatif en a n'est pas assurée. Par exemple, $x \mapsto x^3$ en $a=0$ vérifie $Q=0$, et pourtant cette fonction n'a pas d'extremum relatif en zéro.

2.3 Optimisation sous contraintes

Définition 36. Soit $n \in \mathbb{N}^*$. On dit qu'une partie $M \subset \mathbb{R}^n$ est une sous-variété de $\mathbb{R}^n$ de dimension m en un point $x_0 \in M$ il existe un voisinage ouvert U de x_0 et un $\mathcal{C}^1$ -difféomorphisme $\varphi: U \rightarrow \varphi(U) \subset \mathbb{R}^n$ vérifiant $\varphi(x_0) = 0$ et $\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^m \times \{0\})$.

Proposition 37. On suppose qu'il existe $\varphi_1, \dots, \varphi_{n-m}$ différentiables sur un ouvert U contenant x_0 , à valeurs réelles, telles que $\varphi_1(x_0) = \dots = \varphi_{n-m}(x_0) = 0$ et que les formes linéaires $(D\varphi_i(x_0))_{0 \leq i \leq n-m}$ sont linéairement indépendantes.

Alors l'ensemble $M = \{x \in U : \varphi_1(x) = \dots = \varphi_{n-m}(x) = 0\}$ est une sous-variété en x_0 de dimension m .

Définition 38. Soit $M \subset \mathbb{R}^n$ une sous-variété et $x_0 \in M$. On appelle espace tangent en x_0 à M l'ensemble :

$$T_{x_0}(M) = \{v \in \mathbb{R}^n : \exists I \in \mathcal{I} \quad \exists \gamma \in D^1(I, M), \quad \gamma(0) = m \text{ et } \gamma'(0) = v\}$$

Où l'on a noté $\mathcal{I}$ l'ensemble des intervalles ouverts contenant 0, et $D^1(I, M)$ est l'ensemble des applications différentiables de I vers M , pour $I \in \mathcal{I}$.

Proposition 39. $T_{x_0}(M)$ est un sous-espace vectoriel de $\mathbb{R}^n$, de même dimension que M .

Théorème 40. Soit $\varphi_1, \dots, \varphi_{n-m}$ différentiables sur un ouvert U de $\mathbb{R}^n$ contenant x_0 , que $\varphi_1(x_0) = \dots = \varphi_{n-m}(x_0) = 0$ tel que les formes linéaires $(D\varphi_i(x_0))_{0 \leq i \leq n-m}$ sont linéairement indépendantes.

On note $M = \{x \in U : \varphi_1(x) = \dots = \varphi_{n-m}(x) = 0\}$ la sous-variété associée. Alors on a :

$$T_{x_0}(M) = \bigcap_{i=1}^{n-m} \text{Ker}(D\varphi_i(x_0))$$

Développement 1 :

Théorème 41. Théorème des extrema liés

Soit $f, g_1, \dots, g_k : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ des applications de classe C^1 , et $M = \{x \in U : g_1(x) = \dots = g_k(x) = 0\}$. On suppose que $f|_M$ admet un extremum local en $m \in M$, et que la famille $(Dg_i(m))_{0 \leq i \leq k}$ est libre. Alors il existe $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tels que :

$$Df(m) = \sum_{i=1}^k \lambda_i Dg_i(m)$$

Application 42. (Inégalité de Hadamard)

On considère le produit scalaire $\langle \cdot, \cdot \rangle$ usuel sur $\mathbb{R}^n$. Soit $x_1, \dots, x_n \in \mathbb{R}^n$, on a alors l'inégalité :

$$|\det(x_1, \dots, x_n)| \leq \|x_1\| \cdots \|x_n\|$$

Avec égalité si et seulement si les $(x_i)_{1 \leq i \leq n}$ forment une base orthogonale de E .

3 Problèmes d'optimisation

3.1 Méthode du gradient

L'objectif est de trouver une solution à $Ax = b$ via la recherche d'un minimum local de la fonction $x \mapsto \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle$ pour $A \in S_n^{++}(\mathbb{R})$.

Lemme 43. (Inégalité de Kantorovitch)

Soit $A \in S_n^{++}(\mathbb{R})$. On note λ_1 et λ_n la plus petite (respectivement la plus grande) valeur propre de A . Alors :

$$\begin{aligned} \forall x \in \mathbb{R}^n \quad \langle Ax, x \rangle \langle A^{-1}x, x \rangle &\leq \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \right)^2 \|x\|^4 \\ &= \frac{1}{4} \left(\sqrt{c(A)} + \sqrt{\frac{1}{c(A)}} \right)^2 \|x\|^4, \end{aligned}$$

où $c(A) \stackrel{\text{def}}{=} \frac{\lambda_1}{\lambda_n}$ désigne le contenu de A .

Développement 2 :

Théorème 44. Soit $A \in S_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche à minimiser $f : x \mapsto \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle$ quand x parcourt $\mathbb{R}^n$.

Il existe une unique solution $\bar{x}$ à ce problème, et elle est caractérisée par $\nabla f(\bar{x}) = 0$.

De plus, la suite définie par $\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N} \quad x_{k+1} = x_k + t_k d_k \end{cases}$ converge vers $\bar{x}$, où $t_k = -\nabla f(x_k)$ est l'unique réel minimisant la fonction $t \mapsto f(x_k + t d_k)$.

3.2 Une optimisation géométrique

On se donne $a < b$ et $c < d$ des réels.

Définition 45. 1. On appelle courbe paramétrée une application $\gamma : [a, b] \rightarrow \mathbb{R}^2$ vérifiant

- $\gamma \in C^1([a, b])$.
- $\forall s \in [a, b] \quad \gamma'(s) \neq 0$.

On note alors $\Gamma := \text{Im}(\gamma)$ et on dit que Γ est une courbe.

2. On dit qu'une courbe Γ est **simple** si elle ne s'intersecte pas, c'est-à-dire si

$$\forall (s_1, s_2) \in [a, b]^2 \quad \gamma(s_1) = \gamma(s_2) \implies s_1 = a \text{ et } s_2 = b$$

3. On dit qu'une courbe Γ est **fermée** si $\gamma(a) = \gamma(b)$.

Définition 46. On appelle longueur d'une courbe Γ paramétrée par $\gamma = (x, y)$ le réel

$$\begin{aligned} \ell(\Gamma) &:= \int_a^b |\gamma'(s)| ds \\ &= \int_a^b (x'(s)^2 + y'(s)^2)^{1/2} ds. \end{aligned}$$

Définition 47. On dit qu'une paramétrisation $\gamma : [a, b] \rightarrow \mathbb{R}^2$ d'une courbe Γ est une paramétrisation par longueur d'arc si $|\gamma'(s)| = 1$ pour tout $s \in [a, b]$.

Cela signifie en particulier que la longueur de Γ vaut exactement $b - a$. Aussi, à une translation près, une paramétrisation par longueur d'arc est définie sur $[0, \ell]$.

Théorème 48. Toute courbe fermée simple admet une paramétrisation par longueur d'arc.

Développement 3 :

Théorème 49. (Inégalité isopérimétrique)

Soit Γ une courbe fermée simple sur $\mathbb{R}^2$ de longueur ℓ , et $\mathcal{A}$ l'aire de la région enfermée par Γ . Alors on a

$$\mathcal{A} \leq \frac{\ell^2}{4\pi},$$

avec égalité si et seulement si Γ est un cercle.

