

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

Références: Beck, Goursat, Rombaldi, Rouvière, Demailly, Tauvel
(Analyse réelle) (pour 1 résultat) (Ana Comp)
pour Reelle, TFS...

I - Existence et unicité des extremums

- 1) Compacité et coercivité
- 2) Convexité et optimisation
- 3) Le cadre hilbertien
- 4) Le cadre holomorphe

II - Optimisation de fonctions différentiables

- 1) Condition du premier ordre
- 2) Conditions du second ordre
- 3) Optimisation sous contraintes

III - Méthodes numériques d'optimisation

- 1) Méthode de Newton
- 2) Méthode du gradient

DEV 1: Projection sur un convexe fermé

DEV 2: Théorème des extrema liés

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

Soit (E, d) un espace métrique.

I - Existence et unicité d'extremums

1) Compacité et coercivité

DEF 1: Soient $X \subset E$, $a \in X$ et $f: X \rightarrow \mathbb{R}$. On dit que f admet en a un maximum local (resp. minimum local) lorsqu'il existe un voisinage V de a tel que :

$$\forall x \in V \cap X, f(x) \leq f(a) \text{ (resp. } f(x) \geq f(a))$$

DEF 2: On dit que f admet en a un maximum global (resp. minimum global) lorsque :

DEF 3: Lorsque les inégalités précédentes sont strictes, on parle de maximum ou de minimum strict. On désigne par extremum un maximum ou un minimum.

LEM 4: Un extremum global est local mais la réciproque est fautive : $x \mapsto x(x-1)(x+1)$ sur $\mathbb{R}$

PROP 5: On suppose E compact et (F, f) un espace métrique. Soit $f: E \rightarrow F$ continue. Alors, $f(E)$ est compact.

PROP 6: Soit $f: (E, d) \rightarrow \mathbb{R}$ continue avec (E, d) compact. Alors, f est bornée et atteint ses bornes :

$$\exists c, d \in E, f(c) = \inf_{x \in E} f(x) \text{ et } f(d) = \sup_{x \in E} f(x).$$

LEM 7: On utilise souvent une suite dite minimisante $(x_n)_{n \in \mathbb{N}}$ telle que $f(x_n) \xrightarrow{n \rightarrow \infty} \inf_{x \in E} f(x)$ qui existe par caractérisation de la borne inférieure.

LEM 8: La compacité de l'espace nous permet d'extraire une sous-suite convergente.

APPLI 9: Si $(E, \|\cdot\|)$ est un espace vectoriel normé, toutes les normes sont équivalentes sur E .

DEF 10: $(E, \|\cdot\|)$ é.v.m. $f: E \rightarrow \mathbb{R}$ est dite coercive lorsque $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$

EX 11: Une norme est coercive, une forme linéaire ne l'est pas.

PROP 12: Soit $f: E \rightarrow \mathbb{R}$ continue coercive. Alors, f est minorée et atteint son minimum global.

PROP 13: $(E, \|\cdot\|)$ é.v.m. et F s.v. de E de dimension finie. Alors $\forall y \in F, \exists x \in F, d(y, F) = d(x, y) = \|x - y\|$.

THM 14: Pour tout $n \in \mathbb{N}^*$ on note P_n l'espace des fonctions polynomiales de degré inférieur ou égal à n . $\forall n \in \mathbb{N}, \exists ! q_n \in P_n, \|f - q_n\| = d(f, P_n)$ où $f \in \mathcal{C}([a, b], \mathbb{R})$. q_n est appelé polynôme de meilleure approximation uniforme de f à l'ordre n .

2) Convexité et optimisation [RFC] $(E, \|\cdot\|)$ espace vectoriel normé.

DEF 15: Soit $C \subset E$. On dit que C est convexe lorsque

$$\forall (x, y) \in C, \forall t \in [0, 1], tx + (1-t)y \in C.$$

DEF 16: $f: E \rightarrow \mathbb{R}$ est convexe lorsque :

$$\forall (x, y) \in E, \forall t \in [0, 1], f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

On dit que f est strictement convexe lorsque l'inégalité est stricte.

PROP 17: On définit l'épigraph de f par $\mathcal{E}(f) = \{(x, y) \in E \times \mathbb{R} \mid y \geq f(x)\}$. f est convexe $\Leftrightarrow \mathcal{E}(f)$ est convexe.

EX 18: Une norme euclidienne est strictement convexe.

Une norme quelconque est convexe.

THM 19: Soient $U \subset \mathbb{R}^n$ ouvert convexe et $f: U \rightarrow \mathbb{R}$. Soit f différentiable sur U , LASSE :

- 1) f est convexe
- 2) $\forall x, y \in U, f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$ (le graphe est au-dessus des tangentes)
- 3) $\forall x, y \in U, \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$
- 4) $\forall x \in U, \langle \nabla f(x), \nabla f(x) \rangle \geq 0$

APPLI 20: Soit $A \in \mathcal{M}_n(\mathbb{R})$. $x \mapsto \langle Ax, x \rangle$ est convexe $\Leftrightarrow A$ est positive.

LEMME 21: Soit $C \subset E$ convexe non vide, $f: C \rightarrow \mathbb{R}$ strictement convexe sur C . Il existe au plus un point $\bar{x} \in C$ minimisant f sur C .

PROP 22: Si $f: C \rightarrow \mathbb{R}$ est continue et admet un minimum local, alors c'est un minimum global.

3) Le cadre hilbertien $[H, \langle \cdot, \cdot \rangle]$ Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert.

THM 23: Soit $C \subset H$ convexe fermée non vide. Alors, pour tout $x \in H$, il existe un unique $y \in C$ tel que $\|x - y\| = d(x, C)$. On appelle y le projeté de x sur C et on note $y = P_C(x)$. De plus, on a: $y = P_C(x) \iff y \in C$ et $\forall z \in C, \operatorname{Re}(\langle x - y, z - y \rangle) \leq 0$

PROP 24: $P_C: H \rightarrow C$ est 1-lipshitzienne.

PROP 25: Soit F un s.v.f. fermé de H . Alors P_F est linéaire et si $x \in H, P_C(x)$ est l'unique $y \in F$ tel que $x - y \in F^\perp$.

COR 26: On suppose $\dim(H) = m \in \mathbb{N}^*$. Soit $(x_1, \dots, x_p)$ linéaire dans $H, F = \operatorname{vect}(x_1, \dots, x_p)$. Soit $x \in H$. Alors $d(x, F)^2 = \frac{G(x_1, \dots, x_p, x)}{G(x_1, \dots, x_p)}$ où $G(x_1, \dots, x_p)$ est le déterminant de la matrice de Gram $(\langle x_i, x_j \rangle)_{1 \leq i, j \leq p}$.

PROP 27: Pour F s.v.f. de H , on a $H = F \oplus F^\perp$.

THM 28. (Riesz) L'application $f: H \rightarrow H^*$ est une isométrie linéaire surjective $y \mapsto f_y: H \rightarrow K, x \mapsto \langle x, y \rangle$.

THM 29. (Lax-Milgram) Soit $a: H \times H \rightarrow K$ bilinéaire telle que: $\exists C > 0, \exists \alpha > 0, \forall x, y \in F, |a(x, y)| \leq C \|x\| \|y\|$ et $\forall x \in F, a(x, x) \geq \alpha \|x\|^2$ (continue et coercive). Alors: $\exists ! u \in H, \forall y \in F, a(u, y) = L(y)$. Si a est symétrique si $\phi(x) = \frac{1}{2} a(x, x) - L(x)$.

Alors $\phi(u) = \min_{x \in H} \phi(x)$.

4) Le cadre holomorphe [QUE] ou [TAU] $U \subset \mathbb{C}^n$ ouvert non vide

PROP 30: Soit $f: U \rightarrow \mathbb{C}$ holomorphe. Alors f vérifie la propriété de la moyenne: $\forall a \in U, f(a) = \frac{1}{2\pi} \int_0^{2\pi} f(a + re^{it}) dt$

THM 31: (Principe du maximum) On suppose U connexe. Soit $f: U \rightarrow \mathbb{C}$ holomorphe. Si $z \mapsto |f(z)|$ admet un maximum local sur U , alors f est constante.

COR 32: On suppose U connexe borné. Soit $f: U \rightarrow \mathbb{C}$ continue, holomorphe sur U . Alors: $\sup_{z \in \bar{U}} |f(z)| = \sup_{z \in \partial U} |f(z)|$

LEMME 33. (Schwarz) Soit $f: D(0, 1) \rightarrow \mathbb{C}$ holomorphe. On suppose $f(0) = 0$ et $|f(z)| \leq 1$ pour tout $z \in D(0, 1)$. Alors: $\forall z \in D(0, 1), |f(z)| \leq |z|$. De plus, 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 $a \in \mathbb{C}$ tel que $f(z) = az$ pour tout $z \in D(0, 1)$, c'est-à-dire f est une homothétie.

THM 34: (Liouville) Si f est holomorphe sur $\mathbb{C}$ et bornée, alors, f est constante.

THM 35: (D'Alembert-Goursat) $\mathbb{C}$ est algébriquement clos.

II - Optimisation de fonctions différentiables

1) Condition du premier ordre [BFC] [Pau] [G-2]

Soit $U \subset \mathbb{R}^n$ un ouvert non vide.

DEF 36: Soit $x_0 \in U$. On dit que x_0 est un point critique de f lorsque $df(x_0) = 0$.

PROP 37: Soit $f: U \rightarrow \mathbb{R}$. Si f admet un extremum local en $x_0 \in U$, alors x_0 est un point critique de f .

REM 38: L'hypothèse U ouvert est essentielle:

$\exists x \in [0, 1] \mapsto x$ admet ~~un~~ maximum en 1 mais $f'(1) \neq 0$.

REM 39: La réciproque est fautive: $x \in \mathbb{R} \mapsto x^3$ a sa dérivée nulle en 0 mais 0 n'est pas un extremum local.

THM 40. (Rolle) Soit $f: [a, b] \rightarrow \mathbb{R}$ continue, et dérivable sur $]a, b[$ telle que $f(a) = f(b)$. Alors il existe $c \in]a, b[, f'(c) = 0$.

THM 41: (Accroissements finis) Sous les mêmes hypothèses, sans supposer $f(a) = f(b)$, on a: $\exists c \in]a, b[, f'(c) = \frac{f(b) - f(a)}{b - a}$

REM 42: Le théorème de Rolle est faux si f n'est pas à valeurs réelles: $t \mapsto e^{it}$ sur $[0, 2\pi]$.

COR 43: Soit f dérivable sur $I \subset \mathbb{R}$.

f est croissante sur $I \iff f' \geq 0$ sur I

f est décroissante sur $I \iff f' \leq 0$ sur I

f est constante sur $I \iff f' = 0$ sur I .

2) Conditions du second ordre [ROU] [RAU] Unicité de

PROP 44: Soit $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est deux fois différentiable et admet un minimum local (resp. maximum local) en a alors $D^2 f(a)$ est une forme quadratique positive (resp. négative).

PROP 45: Soit $f: U \rightarrow \mathbb{R}$ deux fois différentiable, $a \in U$ un point critique de f . Si $D^2 f(a)$ est une forme quadratique définie positive (resp. définie négative), alors f admet un minimum local (resp. maximum local) en a .

REM 46: Pour chercher les extrema de $f: U \rightarrow \mathbb{R}$ assez régulière, on commence par annuler son gradient puis on étudie en chaque point critique sa matrice hessienne $H_f(a)$. Comme celle-ci est symétrique, il suffit de regarder ses valeurs propres.

REM 47: La réciproque PROP 45 est fautive: $x \mapsto x^4$

EX 48: En dimension 2, soit $f: U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ et $a \in U$ tel que $df(a) = 0$. On pose $\alpha = \frac{\partial^2 f}{\partial x^2}(a)$, $\beta = \frac{\partial^2 f}{\partial x \partial y}(a)$, $\gamma = \frac{\partial^2 f}{\partial y^2}(a)$. Par le théorème, si $\alpha > 0$ et $\alpha\gamma - \beta^2 > 0$, f admet un minimum local en a , si $\alpha < 0$ et $\alpha\gamma - \beta^2 > 0$, f admet un maximum local en a . Si $\alpha > 0$ et $\alpha\gamma - \beta^2 < 0$, a est un point selle.

3) Optimisation sous contraintes [GAU] DEV 2/6

THM 49 (Extremum liés) Soient $f, g_1, \dots, g_k: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. Soit $M = \{x \in U \mid g_1(x) = \dots = g_k(x) = 0\}$. Si $f|_M$ admet un extremum local en $a \in M$ et si $dg_1(a), \dots, dg_k(a)$ sont linéairement indépendantes, alors:

$$\exists \lambda_1, \dots, \lambda_k \in \mathbb{R}, df(a) = \sum_{i=1}^k \lambda_i dg_i(a).$$

REMARQUE: Les coefficients $\lambda_1, \dots, \lambda_k$ sont appelés multiplicateurs de Lagrange.

LEMME 51: Soient $f, g_1, \dots, g_k$ des formes linéaires sur $\mathbb{R}^n$ telles que $(g_1, \dots, g_k)$ soit libre et $\bigcap_{i=1}^k \ker(g_i) \subset \ker(f)$. Alors: $\exists \lambda_1, \dots, \lambda_k, f = \sum_{i=1}^k \lambda_i g_i$.

III - Méthodes numériques d'optimisation

1) Méthode de Newton [ROU]

Soient $c < d$, $f: [c, d] \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$. On suppose $f(c) < 0 < f(d)$ et pour tout $x \in [c, d]$, $f'(x) > 0$. On considère la suite itérative:

$$x_{n+1} = F(x_n) \text{ et } F(x) = x - \frac{f(x)}{f'(x)}$$

PROP 52: f a un unique zéro $a \in [c, d]$. Pour tout $x \in [c, d]$, il existe $z \in [a, x]$, $F(x) - a = \frac{1}{2} \frac{f''(z)}{f'(z)} (x - a)^2$.

PROP 53: $\exists C > 0$, $\forall x \in [c, d]$, $|F(x) - a| \leq C|x - a|^2$. $\exists \delta > 0$, $I = [a - \delta, a + \delta]$ est stable par F . Pour tout $x_0 \in I$, (x_n) converge quadratiquement vers $a \in I$.

PROP 54: On suppose: $\forall x \in [c, d]$, $f''(x) > 0$. Alors $I = [a, d]$ est stable par F , pour $x_0 \in I$, (x_n) est strictement décroissante (ou constante) avec: $0 \leq x_n - a \leq C(x_0 - a)^2$ et $x_n - a \sim \frac{1}{2} \frac{f''(a)}{f'(a)} (x_0 - a)^2$.

REMARQUE: Cette méthode peut être utilisée pour approcher les zéros de f' qui sont les points critiques.

2) Méthode du gradient [BEC]

PROP 56: Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$. $\nabla f(a)$ donne la direction de la plus grande pente de f en a .

REM 57: Pour approcher la solution x^* d'un problème d'optimisation, on utilise une suite minimisante (x_n) .

Supposons que x_n ne soit pas assez proche de x^* , on cherche x_{n+1} de la façon suivante:

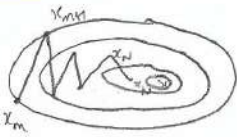
→ on choisit une direction de pente d_n

→ on choisit un pas de descente t_n

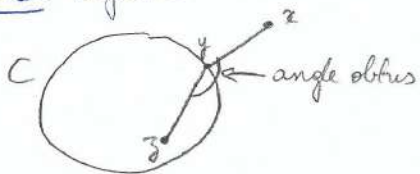
On pose $x_{n+1} = x_n + t_n d_n$. La méthode du gradient consiste à poser $d_n = -\nabla f(x_n)$.

Gradient à pas constant: $t_n = 1$
Gradient à pas optimal: t_n tel que $f(x_n + t_n d_n) = \min_{t \geq 0} f(x_n + t d_n)$

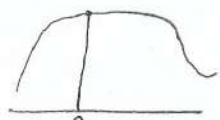
ANNEXE : Méthode du gradient



ANNEXE : Projection



ANNEXE :



maximum
global non strict



maximum
local strict