

Cadre : Soient $(E, \|\cdot\|)$ un \mathbb{R} -espace vectoriel normé et $J : E \rightarrow \mathbb{R}$.

I Existence et unicité d'extrema

Définition 1. On dit que J admet un minimum (resp. maximum) local en $x_0 \in E$ s'il existe un voisinage V de x_0 tel que $\forall x \in V, J(x) \geq J(x_0)$ (resp. $\forall x \in V, J(x) \leq J(x_0)$). On dit que J admet un minimum (resp. maximum) global en $x_0 \in E$ si $\forall x \in E, J(x) \geq J(x_0)$ (resp. $\forall x \in E, J(x) \leq J(x_0)$). Un extremum de J est un maximum ou un minimum de J .

1) Compacité et fermeture

Proposition 2. Si K est un compact de E , et si J est continue, alors J est bornée sur K et atteint ses bornes.

Application 3. Si K et K' sont deux compacts de E , il existe $(x, x') \in K \times K'$ tel que $(x, x') = d(K, K')$.

Définition 4. On dit que J est coercive si $\lim_{\|x\| \rightarrow \infty} J(x) = +\infty$.

Théorème 5. Si $E = \mathbb{R}^n$, et si J est coercive et continue, alors J admet un minimum global.

Contre-exemple 6. Sur l'espace de Hilbert $\ell^2(\mathbb{N})$, la fonctionnelle $J(x) = (\|x\|^2 - 1) + \sum_{i=1}^{\infty} \frac{x_i^2}{i}$ est coercive mais n'admet pas de minimum.

2) Convexité

Définition 7. Soit C un convexe non vide de E . On dit que $J : C \rightarrow \mathbb{R}$ est convexe si, pour tous $a, b \in C$ et tout $\lambda \in [0, 1]$, on a :

$$J((1 - \lambda)a + \lambda b) \leq (1 - \lambda)J(a) + \lambda J(b)$$

On dit que J concave si $-J$ est convexe. Lorsque l'inégalité est stricte pour $a \neq b$ et $0 < \lambda < 1$, J est strictement convexe. Pour $\alpha > 0$, on dit que J est α -convexe si pour tous $a, b \in C$ distincts et tout $\lambda \in]0, 1[$, on a :

$$J((1 - \lambda)a + \lambda b) \leq (1 - \lambda)J(a) + \lambda J(b) - \frac{\alpha}{2} \|a - b\|^2 \lambda(1 - \lambda)$$

Théorème 8. On considère $J : C \rightarrow \mathbb{R}$.

- (i) Si J est convexe, tout minimum local est global.
- (ii) Si J est strictement convexe, J admet au plus un minimum global.
- (iii) Si J est α -convexe, J admet un unique minimum global.

3) Cas des espaces de Hilbert

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. Soit $K \subset E$ convexe fermé non vide.

Théorème 9. Pour tout $f \in H$, il existe un unique élément de K , noté $P_K(f)$, et appelé projection de f sur K , tel que :

$$\|P_K(f) - f\| = \inf_{v \in K} \|v - f\|$$

De plus, $P_K(f)$ est caractérisée par :

$$\forall v \in K, \operatorname{Re}(\langle f - P_K(f), v - P_K(f) \rangle) \leq 0$$

Corollaire 10. Soient M un sous-espace vectoriel fermé de H et $f \in H$. Alors $P_M(f)$ est caractérisé par :

$$P_M(f) \in M \quad \text{et} \quad \forall v \in M, \operatorname{Re}(\langle f - P_M(f), v \rangle) = 0$$

De plus, P_M est un opérateur linéaire.

Théorème 11 (Riesz-Fréchet). Soit $\varphi \in H'$. Alors :

$$\exists! f \in H, \forall v \in H, \langle \varphi, v \rangle = \langle f, v \rangle$$

Théorème 12 (Lax-Milgram). Soient H un espace de Hilbert, a une forme bilinéaire continue et coercive sur H , et $\ell \in H'$. Alors il existe un unique $u \in H$ tel que, pour tout $v \in H$, $a(u, v) = \ell(v)$. Si de plus a est symétrique, u réalise le minimum sur H de $v \mapsto \frac{1}{2}a(v, v) - \ell(v)$.

Application 13 (Dirichlet). Pour $f \in L^2$, on considère le problème :

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

Il existe une unique solution faible $u \in H_0^1([0, 1])$ à ce problème.

4) Holomorphie

On considère $E = \mathbb{C}$, $\Omega \subset \mathbb{C}$ un ouvert connexe non vide et $J \in \mathcal{H}(\Omega)$.

Théorème 14. Soient $z_0 \in \Omega$ et $r > 0$ tels que $B(z_0, r) \subset \Omega$. Alors $J(z_0) = \frac{1}{2\pi} \int_0^{2\pi} J(z_0 + re^{i\theta}) d\theta$.

Théorème 15 (Principe du maximum). Si $|J|$ atteint son maximum en un point de Ω , alors J est constante.

Application 16 (D'Alembert-Gauss). Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine dans \mathbb{C} .

II Caractérisation des extrema

1) Différentiabilité et points critiques

Soit U un ouvert de \mathbb{R}^n . On considère $J : U \rightarrow \mathbb{R}$.

Ordre 1

Définition 17. Soit $x_0 \in U$. On dit que x_0 est un point critique pour J si J est différentiable en x_0 et $dJ(x_0) = 0$.

Proposition 18. Si x_0 est un extremum local de J , et si J est différentiable en x_0 , alors x_0 est un point critique.

Remarque 19. Cette condition est nécessaire, mais pas suffisante : $x \mapsto x^3$ n'admet pas d'extremum en 0.

Théorème 20 (Rolle). Soit $J : [a, b] \rightarrow \mathbb{R}$ continue et dérivable sur $]a, b[$. Si $J(a) = J(b)$, alors il existe $c \in]a, b[$ tel que $J'(c) = 0$.

Corollaire 21 (Accroissements finis). Soit $J : [a, b] \rightarrow \mathbb{R}$ continue et dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que $J'(c)(b-a) = J(b) - J(a)$.

Ordre 2

Proposition 22. Soit $x_0 \in U$ un point critique de J . On suppose J de classe \mathcal{C}^2 en x_0 . Alors :

- (i) Si x_0 est un minimum (resp. maximum) local, alors $d^2J(x_0)$ est positive (resp. négative).
- (ii) Si $d^2J(x_0)$ est définie positive (resp. définie négative), alors x_0 est un minimum (resp. maximum) local.

Remarque 23. Encore une fois, ces conditions laissent un cas douteux : $x \mapsto x^3$ n'admet pas d'extremum en 0.

Exemple 24. Dans le cas où $n = 2$, on pose $A = \begin{pmatrix} r & s \\ s & t \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ la hessienne de J en x_0 . Par le théorème précédent, on a :

- (i) Si $rt - s^2 > 0$ et $r > 0$, J admet un minimum relatif en x_0 .
- (ii) Si $rt - s^2 > 0$ et $r < 0$, J admet un maximum relatif en x_0 .
- (iii) Si $rt - s^2 < 0$, J n'a pas d'extremum en x_0 .
- (iv) Si $rt - s^2 = 0$: cas douteux.

Exemple 25. Si $J(x, y) = x^4 + y^4 - 2(x - y)^2$, alors J a trois points critiques : $(0, 0)$ et $\pm(\sqrt{2}, -\sqrt{2})$. Il y a un minimum local en $\pm(\sqrt{2}, -\sqrt{2})$, mais on ne peut pas conclure en $(0, 0)$.

2) Fonctions convexes

Théorème 26. Soit $J : C \rightarrow \mathbb{R}$ différentiable. Il y a équivalence entre :

- (i) J est convexe sur C .
 - (ii) $\forall x, y \in C, \langle \nabla J(x) - \nabla J(y), x - y \rangle \geq 0$.
 - (iii) $\forall x, y \in C, J(x) \geq J(y) + \langle \nabla J(y), x - y \rangle$.
- Si J est deux fois différentiable, on a aussi : $\langle d^2J(x) \cdot y, y \rangle \geq 0$.

Théorème 27. Soit $J : C \rightarrow \mathbb{R}$ différentiable. Il y a équivalence entre :

- (i) J est α -convexe sur C .
 - (ii) $\forall x, y \in C, \langle \nabla J(x) - \nabla J(y), x - y \rangle \geq \alpha \|x - y\|^2$.
 - (iii) $\forall x, y \in C, J(x) \geq J(y) + \langle \nabla J(y), x - y \rangle + \frac{\alpha}{2} \|x - y\|^2$.
- Si J est deux fois différentiable, on a aussi : $\langle d^2J(x) \cdot y, y \rangle \geq \alpha \|y\|^2$.

Exemple 28. Si A est une matrice symétrique définie positive, alors la fonctionnelle quadratique $J : X \mapsto \langle AX, X \rangle - \langle B, X \rangle$ est λ_1 -convexe, où λ_1 est la plus petite valeur propre de A .

Théorème 29. Si $J : C \rightarrow \mathbb{R}$ est différentiable en $u \in C$ et admet un minimum local en u , alors $\langle \nabla J(u), v - u \rangle \geq 0$ pour tout $v \in C$.

Corollaire 30. Soit $J : C \rightarrow \mathbb{R}$ est convexe et différentiable en $u \in C$. Alors u est un extremum local si, et seulement si, $\nabla J(u) = 0$.

3) Optimisation sous contraintes

Définition 31. Soit M une sous-variété de \mathbb{R}^n . On dit qu'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a un extremum lié (ou relatif) en $a \in M$ s'il existe un voisinage U de a dans \mathbb{R}^n tel que $f(a)$ est un extremum de f sur $M \cap U$.

Théorème 32 (Extrema liés). Soit U un ouvert de \mathbb{R}^n . Soient g_1, \dots, g_k des fonctions de classe \mathcal{C}^1 de U dans \mathbb{R} telles que les formes linéaires $d_x g_1, \dots, d_x g_k$ sont linéairement indépendantes pour tout $x \in U$. Posons :

$$M = \{x \in U \mid \forall i \in \llbracket 1, k \rrbracket, g_i(x) = 0\}$$

Alors, si f a un extremum lié en $a \in M$, il existe $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tels que :

$$d_a f = \sum_{i=1}^k \lambda_k d_a g_i$$

Ces réels $\lambda_1, \dots, \lambda_k$ sont appelés multiplicateurs de Lagrange.

Application 33. Tout endomorphisme symétrique de E admet une valeur propre réelle.

III Optimisation numérique

1) Méthode de Newton

La méthode de Newton consiste à approcher une solution d'une équation $f(x) = 0$ en partant d'une approximation plus grossière. L'idée est de remplacer la courbe de f par sa tangente.

Théorème 34 (Méthode de Newton). Soient $a, b \in \mathbb{R}$ tels que $a < b$, et soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction de classe C^2 telle que $f(a) < 0 < f(b)$ et $f' > 0$ sur $[a, b]$. On considère la suite $(x_n)_{n \in \mathbb{N}}$ définie par :

$$x_0 \in [a, b] \quad \text{et} \quad \forall n \in \mathbb{N}, x_{n+1} = \phi(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$

La fonction f admet un unique zéro $\alpha \in]a, b[$, et on a :

(i) Il existe $\varepsilon > 0$ tel que, pour $x_0 \in I =]\alpha - \varepsilon, \alpha + \varepsilon[$, la suite $(x_n)_{n \in \mathbb{N}}$ converge quadratiquement vers α , et il existe $C > 0$ tel que :

$$\forall n \in \mathbb{N}, |x_{n+1} - \alpha| \leq C|x_n - \alpha|^2$$

(ii) Si de plus $f'' > 0$ sur $[\alpha, b]$, alors, pour $x \in]\alpha, b]$, la suite $(x_n)_{n \in \mathbb{N}}$ est strictement décroissante, et pour tout $n \in \mathbb{N}$ on a :

$$0 \leq x_{n+1} - \alpha \leq C(x_n - \alpha)^2 \quad \text{et} \quad x_{n+1} - \alpha \sim \frac{f''(\alpha)}{2f'(\alpha)}(x_n - \alpha)^2$$

2) Méthodes de gradient

Soit $J : \mathbb{R}^n \rightarrow \mathbb{R}$. On suppose J différentiable. On cherche, s'il existe, un élément $u \in \mathbb{R}^n$ tel que :

$$J(u) = \inf_{v \in \mathbb{R}^n} J(v)$$

Pour cela, on utilise les méthodes de gradient. On considère la suite :

$$u_0 \in \mathbb{R}^n \quad \text{et} \quad \forall k \in \mathbb{N}, u^{k+1} = u^k - \rho^k \nabla J(u^k)$$

Il existe plusieurs possibilités pour choisir les ρ^k , par exemple :

- (i) Gradient à pas fixe : $\rho^k = \rho$ une constante positive fixée.
- (ii) Gradient à pas optimal : ρ^k minimise $\rho \mapsto J(u^k - \rho \nabla J(u^k))$.

Théorème 35. Si J est α -convexe et différentiable, et que ∇J est L -lipschitzienne, alors la méthode de gradient à pas optimal converge vers l'unique minimum de J .

Application 36. Soient $A \in \mathcal{S}_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$. On considère la fonctionnelle quadratique $J : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par :

$$J(X) = \langle AX, X \rangle - \langle b, X \rangle + c$$

Cette fonctionnelle satisfait les conditions du théorème précédent. De plus, son minimum est atteint en $X_0 \in \mathbb{R}^n$ qui vérifie $\nabla J(X_0) = AX - b = 0$. On a donc une méthode itérative pour approcher la solution de $AX = b$.

Développements

- Projection sur un convexe fermé et théorème de Riesz (9,10,11) [Bre87]
- Extrema liés (32) [Ave83]
- Méthode de Newton (34) [Rou15]
- Algorithme de gradient à pas optimal (35) [Cia88]

Références

- [Gou08] X. Gourdon. *Les Maths en Tête : Analyse*. Ellipses
- [BMP05] V. Beck, J. Malick, et G. Peyré. *Objectif Agrégation*. H&K
- [Bre87] H. Brezis. *Analyse fonctionnelle*. Masson
- [Ave83] A. Avez. *Calcul différentiel*. Masson
- [Rou15] F. Rouvière. *Petit Guide de Calcul Différentiel*. Cassini
- [Cia88] P. Ciarlet. *Introduction à l'analyse numérique et à l'optimisation*. Masson

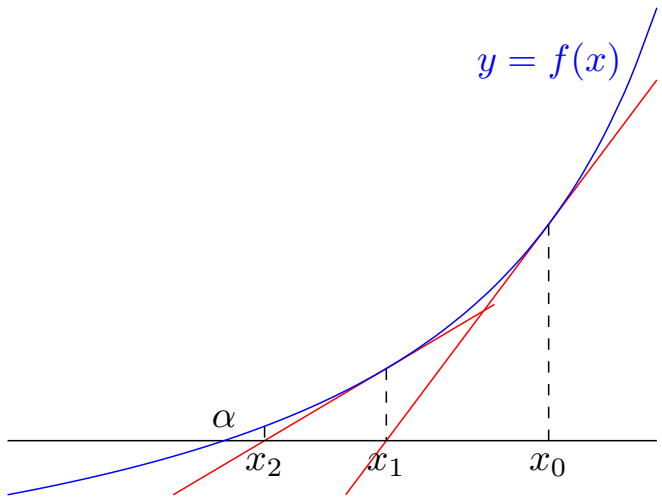


FIGURE 1 – Méthode de Newton