219

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

Cadre : Soit K un espace métrique et $J : K \to \mathbb{R}$.

I) Existence des extremums

A] Cas des espaces compacts

[2]

Thm 1 : Si K est compact et J est continue, alors $J(K)$ est compact.

Cor 2 : Si K est compact et J continue, J est bornée et atteint ses bornes.

Ex 3 : Ceci n'est plus vrai si K n'est plus compact. Par exemple, Arctan est borné sur $K = \mathbb{R}$ mais n'atteint pas ses bornes.

On s'intéressera par la suite essentiellement à la recherche du min, quitte à changer J en $-J$.

[7]

Def 4 : Une suite $(u_n) \in K^{\mathbb{N}}$ est dite minimisante pour J si $\lim_{n \to +\infty} J(u_n) = \inf_{x \in K} J(x)$

Rem 5 : J admet toujours des suites minimisantes d'après la caractérisation séquentielle de la borne inférieure.

Thm 6 : On suppose que $K \subset \mathbb{R}^n$ est fermé non vide, et que J est continue telle que $\lim_{\|x\| \to +\infty} J(x) = +\infty$. Alors J admet un minimum sur K.

Def 7 : Si $K = \mathbb{R}^n$, J est dite coercive si $\lim_{\|x\| \to +\infty} J(x) = +\infty$

Cor 8 : Toute fonction continue coercive admet un minimum.

App 9 : Soit $A \in \mathcal{S}_n(\mathbb{R})$ symétrique définie positive. Alors la fonctionnelle quadratique $x \mapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$ admet un minimum sur $\mathbb{R}^n$ pour tout $b \in \mathbb{R}^n$.

[1]

Rem 10 : Ce résultat ne tient plus en dimension infinie, par exemple pour $K = \ell_2(\mathbb{N})$ et $J(x) = (\|x\|^2 - 1)^2 + \sum_{i=1} \frac{x_i^2}{i}$.

B] Cas où J est convexe

[1]

On suppose ici que K est convexe, c'est-à-dire $\forall x, y \in K, \forall \theta \in [0;1], \theta x + (1-\theta) y \in K$.

Def 11 : J est convexe si $\forall x, y \in K, \forall \theta \in [0;1]$, $J(\theta x + (1-\theta) y) \leq \theta J(x) + (1-\theta) J(y)$

Rem 12 : Ceci traduit le fait que J est en-dessous de ses cordes si $K \subset \mathbb{R}$ (Voir figure 1).

Prop 13 : Si J est convexe, tout minimum local est en fait global. De plus, l'ensemble des minimums de J est convexe.

Prop 14 : On suppose que K est un sous-espace vectoriel de $\mathbb{R}^n$ et que J est différentiable. Les assertions suivantes sont équivalentes :

① J est convexe

② $J(v) \geq J(u) + \langle J'(u), v-u\rangle \quad \forall v, u \in K$

③ $\langle J'(u) - J'(v), u-v\rangle \quad \forall v, u \in K$

Ex 15 : La fonctionnelle quadratique de l'application 9 est convexe.

Def 16 : J est dite strictement convexe si $\forall x \neq y \in K$, $\forall \theta \in]0;1[$, $J(\theta x + (1-\theta) y) < \theta J(x) + (1-\theta) J(y)$

Prop 17 : On retrouve les mêmes caractérisations que Prop 14 mais avec des inégalités strictes dès que $u \neq v$.

Prop 18 : Si J strictement convexe admet un minimum, alors il est unique.

[1] **App 19** : Le minimum de App9 est unique.

C] Cas d'un espace de Hilbert

Thm 20 : (de projection) Soient H un espace de Hilbert réel et $K \subset H$ un convexe fermé. Alors $\forall y \in H$, $\exists ! \, p_K(y) \in K \,/\, d(y,K) = \|y - p_K(y)\|$. $p_K(y)$ est appelé projeté de y sur K. [2]

App 21 : Si F est un sous-espace vectoriel fermé de H, $H = F \oplus F^{\perp}$.

App 22 : $F < H$ est dense si et seulement si $F^{\perp} = \{0\}$.

Thm 23 : $p(y)$ est caractérisé par la propriété suivante : $\forall z \in H$, $\langle z - p(y) \mid y - p(y) \rangle \leq 0$

II) Conditions d'optimalité et caractérisations

On souhaite trouver des conditions nécessaires, voir suffisantes, pour qu'un point minimise J.

A] Inéquation d'Euler, points critiques

Thm 24 : On suppose que K est convexe et J différentiable. Si $u \in K$ est minimum local de J, alors $\forall v \in K$, $\langle J'(u) \mid v - u \rangle \geq 0$ (inéquation d'Euler) [1]

Rem 25 : Si $u \in \mathring{K}$, on a en particulier $J'(u) = 0$ (équation d'Euler). Ceci est en particulier vérifié si $K = \mathbb{R}^n$.

Def 26 : $u \in K$ est dit point critique de J si J est différentiable en u et $J'(u) = 0$. [2]

Cor 27 : Si $K = \mathbb{R}^n$ et J différentiable, tout minimum local est point critique de J.

Rem 28 : La réciproque est fausse, par exemple pour $x \mapsto x^3$. On a cependant :

Prop 29 : Si K et J sont convexes, que J est différentiable en $u \in K$ vérifiant l'inéquation d'Euler, alors u est minimum local de J. [1]

Ex 30 : Si $A \in \mathcal{S}_n(\mathbb{R})$ est symétrique définie positive, le minimum sur $\mathbb{R}^n$ de $x \mapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle$ est solution de $Ax = b$.

B] Matrice Hessienne

On suppose ici que $K = \mathbb{R}^n$ et que J est C^2.

Thm 31 : (de Schwarz) $\forall i,j \in [\![1;n]\!]$, $\frac{\partial^2 J}{\partial x_i \partial x_j} = \frac{\partial^2 J}{\partial x_j \partial x_i}$

Def 32 : Soit $a \in K$. On définit la matrice Hessienne de J en a par : $Hess_a(J) = \left(\frac{\partial^2 J}{\partial x_i \partial x_j}(a)\right)_{1 \leq i,j \leq n}$

Rem 33 : C'est une matrice positive d'après le théorème de Schwarz.

Thm 34 : Si $a \in \mathbb{R}^n$ est minimum local de J, alors $Hess_a(J)$ est positive. [2]

Rem 35 : La réciproque est fausse, avec $x \mapsto x^3$. On a cependant :

Thm 36 : Soit $a \in \mathbb{R}^n$. On suppose que $Hess_a(J)$ est définie positive et que a est un point critique. Alors a est un minimum local strict de J.

App 37 : Si $n = 2$ et $Hess_a(J) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$, on peut, en utilisant le théorème spectral, déduire différents cas possibles à l'aide du déterminant :
- Si $rt - s^2 > 0$, si $r > 0$ (resp $r < 0$) a est minimum (resp maximum) local strict de J.
- Si $rt - s^2 < 0$, a n'est pas extremum local de J

C] Extrémas liés

[3] **Thm 38:** (DEV1)(des extrémas liés) Soient U ouvert de $\mathbb{R}^n$ et $f, g_1, \dots, g_k : U \to \mathbb{R}$ C^1. Soit $M = \{\omega \in U \mid \forall i, g_i(\omega) = 0\}$. Si $f_{|M}$ admet un extremum local en $m \in M$ et si $(dg_i x)_{1 \le i \le k}$ est libre pour tout $x \in M$, alors $\exists! (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ / $df_m = \sum_{i=1}^{k} \lambda_i dg_{i,m}$. Ces constantes sont appelées multiplicateurs de Lagrange.

App 39: (Théorème spectral) Tout endomorphisme réel symétrique est diagonalisable.

[2] **App 40:** (inégalité arithmético-géométrique) En posant $f : \mathbb{R}^n \to \mathbb{R}$, $(x_1, \dots, x_n) \mapsto \prod_{i=1}^{n} x_i$ et $M = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n \mid \sum_{i=1}^{n} x_i = S\}$, $S > 0$, on prouve que $\forall (x_1, \dots, x_n) \in \mathbb{R}_+^n$, $\left(\prod_{i=1}^{n} x_i\right)^{1/n} \le \frac{1}{n} \sum_{i=1}^{n} x_i$

III) Méthodes de recherche numériques.

A] Algorithme du gradient

Soit $K = \mathbb{R}^n$.

[1] **Alg 41:** Le principe d'un algorithme de descente est le suivant :
- On se donne $x_0 \in \mathbb{R}^n$.
- Pour $x_k \in \mathbb{R}^n$ donné, si $\nabla J(x_k) = 0$, c'est fini. Sinon, on choisit $d_k \in \mathbb{R}^n$ une direction de descente : $\langle \nabla J(x_k), d_k \rangle < 0$. On calcule $\alpha_k = \operatorname{argmin}_{\rho > 0} J(x_k + \rho d_k)$ et on pose $x_{k+1} = x_k + \alpha_k d_k$

Ex 42: L'algorithme de descente à pas constant consiste à choisir α_k constant.

Ex 43: L'algorithme de descente à pas optimal consiste à prendre $d_k = -\nabla J(x_k)$

[4] **Lem 44:** (DEV2)(Inégalité de Kantorovitch) Soit $A \in \mathcal{M}_n(\mathbb{R})$ symétrique définie positive. Alors :

$$\forall x \neq 0, \quad \frac{\|x\|^4}{\|x\|_A^2 \, \|x\|_{A^{-1}}^2} \geq 4 \frac{\lambda_{max} \lambda_{min}}{(\lambda_{max} + \lambda_{min})^2}.$$

Thm 45: Soient $x^* = A^{-1} b$ pour $b \in \mathbb{R}^n$ et (x_n) la suite générée par l'algorithme du gradient pour $J : x \mapsto \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$. Alors (x_n) converge vers x^* et :

$$\forall k \in \mathbb{N}, \quad \|x_k - x^*\| \le \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \left(\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} \right)^k \|x_0 - x^*\|$$

B] Méthode de Newton

[1] Soit $F : \mathbb{R}^n \to \mathbb{R}^n$. On souhaite résoudre $F(x) = 0$. Une façon de faire est de partir de $x_0 \in \mathbb{R}^n$ et de remplacer, à chaque étape, F par sa partie linéaire : $\forall n \in \mathbb{N}$, $x_{n+1} = x_n - F'(x_n)^{-1} F(x_n)$ (pourvu que ce soit possible)

Thm 46: Soit F de classe C^2 et $a \in \mathbb{R}^n$ tel que $F(a) = 0$ et $F'(a)$ est inversible. Alors il existe $x_0 \in \mathbb{R}^n$ assez proche de a tel que la suite (x_n) générée par la méthode de Newton converge quadratiquement vers a.

Rem 47: On peut utiliser cet algorithme pour résoudre numériquement l'équation d'Euler.

Références :

- Analyse numérique et optimisation, Allaire [1]
- Analyse, Gourdon [2]
- Calcul différentiel, Avez [3]
- Analyse pour l'agreg, Bernis [4]

Fig 1 :