

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

Introduction : Les problèmes d'optimisation sont omniprésents : rotation des avions, tournées de véhicules, statistiques et machine learning... d'où l'intérêt de disposer de méthodes efficaces d'optimisation.

1 Existence et unicité

Définition 1. Soient (X, d) un espace métrique, $f : X \rightarrow \mathbb{R}, a \in X$. On dit que f admet un minimum global en a si $\forall x \in X, f(x) \geq f(a)$. Le minimum est dit strict si $\forall x \neq a, f(x) > f(a)$. On dit que a minimum local s'il existe V voisinage de a tel que pour tout $x \in V, f(x) \geq f(a)$. On dit que f a un maximum (global, strict, local) en a si $-f$ a un minimum (global, local, strict) en a . On dit que f a un extremum en a si elle a un minimum ou un maximum en a .

1.1 Utilisation de la compacité

Proposition 2. Supposons X compact. Alors si f continue, elle atteint un minimum et un maximum sur X .

Application 3. Si X compact et $f : X \rightarrow X$ vérifie $\forall x, y \in X, d(f(x), f(y)) < d(x, y)$, alors f possède un unique point fixe sur X .

Corollaire 4. Si E evn de dimension finie et si $f : E \rightarrow \mathbb{R}$ continue et coercive, ie. $f(x) \xrightarrow{\|x\| \rightarrow +\infty} +\infty$, alors f minorée et atteint son minimum sur E .

Exemple 5. Soit $n \in \mathbb{N}^*$, soit $f \in \mathcal{C}([0, 1], \mathbb{R})$.
Il existe $P \in \mathbb{R}_n[X]$ tel que $\|f - P\|_\infty = \inf_{Q \in \mathbb{R}_n[X]} \|f - Q\|_\infty$.

Application 6 (D'Alembert-Gauss). Soit $P \in \mathbb{C}[X]$ non constant. Alors P possède une racine dans \mathbb{C} .

1.2 Utilisation de la convexité

Définition 7. Soit C partie convexe de \mathbb{R}^n . $f : C \rightarrow \mathbb{R}$ est dite convexe si pour tous $x, y \in E, t \in [0, 1], f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$. Elle est dite strictement convexe si l'inégalité est stricte dès lors que $x \neq y$ et $t \in]0, 1[$.

Proposition 8. Si $f : C \rightarrow \mathbb{R}$ est strictement convexe, alors il existe au plus un point minimisant f sur C .

Exemple 9. $-\ln$ est strictement convexe sur \mathbb{R}_+^* .

Application 10. Soient $A, B \in S_n^{++}(\mathbb{R})$, soit $\alpha \in [0, 1]$. Alors $\det((1-\alpha)A + \alpha B) \geq \det(A)^{1-\alpha} \det(B)^\alpha$, avec inégalité stricte si $\alpha \in]0, 1[$ et $A \neq B$.

Application 11 (John-Loewner). Soit K compact d'intérieur non vide de \mathbb{R}^n . Alors il existe un unique ellipsoïde centré en 0 de volume minimal contenant K .

1.3 Cas particulier : fonctions holomorphes

Définition 12. Soient U ouvert de $\mathbb{C}, f \in \mathcal{C}(U, \mathbb{C})$. On dit que f vérifie la propriété de la moyenne si pour tous $a \in U, r > 0$ tels que $\overline{D}(a, r) \subset U$, on a $f(a) = \frac{1}{2\pi} \int_0^{2\pi} f(a + re^{it}) dt$.

Exemple 13. Les fonctions holomorphes vérifient la propriété de la moyenne.

Proposition 14. Si f vérifie la propriété de la moyenne et que $|f|$ a un maximum local en $a \in U$, alors f constante sur un voisinage de a . Si f non constante et $|f|$ admet un minimum local sur U , celui-ci est nul.

Théorème 15. Si U ouvert connexe et borné de \mathbb{C} , et f continue sur \overline{U} vérifiant la propriété de la moyenne sur U . Soit $M = \max_{z \in \partial U} |f(z)|$. Alors, pour tout $z \in U, |f(z)| \leq M$, et s'il y a égalité en un point, f est constante.

2 Localisation

2.1 Conditions du premier ordre

Proposition 16. Soient U ouvert de $\mathbb{R}^n, f : U \rightarrow \mathbb{R}, \bar{x} \in U$. Si f a un extremum local en \bar{x} et est différentiable en \bar{x} , alors $df(\bar{x}) = 0$.

Contre-exemple 17. $f : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto t^3$ vérifie $df(0) = 0$ mais 0 n'est pas un extremum local de f .

Application 18 (Rolle). Soit $f : [a, b] \rightarrow \mathbb{R}$ continue, dérivable sur $]a, b[$, vérifiant $f(a) = f(b)$. Alors il existe $c \in]a, b[, f'(c) = 0$.

Proposition 19. Soit U ouvert convexe de \mathbb{R}^n , soit $f : U \rightarrow \mathbb{R}$ convexe, différentiable en $\bar{x} \in U$. Si $df(\bar{x}) = 0$, alors \bar{x} minimum global de f sur U .

2.2 Conditions du second ordre

Théorème 20. Soit U ouvert de \mathbb{R}^n , $f \in \mathcal{C}^2(U, \mathbb{R})$, $\bar{x} \in U$.

- Si f a un minimum local en \bar{x} , alors $df(\bar{x}) = 0$ et $d^2f(\bar{x})$ est positive.
- Si $df(\bar{x}) = 0$ et $d^2f(\bar{x})$ définie positive, alors \bar{x} minimum local strict de f .

Contre-exemple 21. $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 - y^3$ vérifie $df(0, 0) = 0, d^2f(0, 0)$ positive, mais $(0, 0)$ n'est pas un minimum local.

Contre-exemple 22. $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^4$ admet un minimum local strict en $(0, 0)$ mais $d^2f(0, 0)$ non définie positive.

2.3 Optimisation sous contraintes

Théorème 23. Soit U ouvert de \mathbb{R}^n , soient $f, g_1, \dots, g_r \in \mathcal{C}^1(U, \mathbb{R})$, soit $\Gamma = \{x \in U, g_1(x) = \dots = g_r(x) = 0\}$. Si $f|_\Gamma$ possède un extremum relatif en $a \in \Gamma$, et si les formes linéaires $(dg_1(a), \dots, dg_r(a))$ forment une famille libre, alors il existe

$$\lambda_1, \dots, \lambda_r \in \mathbb{R} \text{ tels que } df(a) = \sum_{i=1}^r \lambda_i dg_i(a).$$

Application 24. $SO_n(\mathbb{R})$ est l'ensemble des éléments de $SL_n(\mathbb{R})$ de norme minimale.

Soient $k \leq n \in \mathbb{N}, A \in \mathcal{M}_{k,n}(\mathbb{R})$ de rang $k, b \in \mathbb{R}^k, c \in \mathbb{R}^n$.

On note $X_{ad} = \{x \in \mathbb{R}^n, x \geq 0, Ax - b \leq 0\}$. On cherche à déterminer $\bar{x} = \operatorname{argmin}_{x \in X_{ad}} \langle c, x \rangle$.

Proposition 25. X_{ad} est un polyèdre, dont les sommets (et donc points extrémaux) vérifient $Ax = b$. De plus, si le minimum est atteint, alors il l'est au moins sur un sommet de X_{ad} .

On détermine alors une solution exacte grâce à l'algorithme du simplexe : partant d'un sommet, on se déplace selon l'arête de plus forte décroissance jusqu'à ce que ce ne soit plus possible.

3 Optimisation numérique

3.1 Méthodes de descente

Soit U ouvert de \mathbb{R}^p , soit $J \in \mathcal{C}^1(U, \mathbb{R})$ admettant un unique minimum en \bar{x} .

Définition 26. Soit $\alpha > 0$. On dit que J est α -convexe si pour tous $u, v \in U, J\left(\frac{u+v}{2}\right) \leq \frac{1}{2}(J(u) + J(v)) - \alpha\|u - v\|^2$.

On cherche à construire une suite convergeant vers le minimum de f . Pour cela, on fixe $x_0 \in U$ et un critère d'arrêt $\varepsilon > 0$, et tant que $\|x_n - \bar{x}\| \geq \varepsilon$, on calcule un terme supplémentaire $x_{n+1} = x_n - t_n \nabla J(x_n)$, où $t_n > 0$ est appelé pas et $d_n \in \mathbb{R}^p$ est la direction de descente.

On parle de méthode de gradient lorsque $d_n = \nabla f(x_n)$, à pas constant lorsque $t_n = \mu = \text{constante}$.

Proposition 27. Si J est α -convexe et s'il existe $M > 0$ tel que dJ soit M -lipschitzienne, alors la méthode du gradient à pas constant converge lorsque $\mu \in]0, 2\alpha/M^2[$.

On peut aussi choisir un pas maximisant la descente, en prenant pour tout $n \in \mathbb{N}, t_n = \operatorname{argmin}_{\mu > 0} J(x_n + \mu \nabla f(x_n))$. Cette méthode est appelée gradient à pas optimal.

Proposition 28. Si J est α -convexe et df lipschitzienne sur toute partie bornée de U , alors la méthode de gradient à pas optimal converge.

3.2 Application : résolution de systèmes linéaires

On considère le système linéaire $Ax = b$, d'inconnue $x \in \mathbb{R}^p$, où $A \in \mathcal{M}_{p,n}(\mathbb{R}), b \in \mathbb{R}^n$ donnés.

Lemme 29. Si $A \in \mathcal{S}_n(\mathbb{R})$, on définit la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ par $f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle$. Alors :

- Si $A \in \mathcal{S}_n^{++}(\mathbb{R})$, f admet un unique minimum en x_0 qui est la solution de $Ax = b$.
- Si $A \in \mathcal{S}_n^+(\mathbb{R})$ et $b \in \operatorname{Im}(A)$, les points où f atteint son minimum sont exactement les solutions de $Ax = b$.
- Sinon, f n'est pas minorée.

Théorème 30. On suppose $A \in \mathcal{S}_n^{++}(\mathbb{R})$. Alors x_0 réalise le minimum de f ssi $\nabla f(x_0) = 0$. De plus, si $x \in \mathbb{R}^n$ tel que $\nabla f(x) \neq 0$, alors pour tout $\alpha \in]0, \frac{2}{\rho(A)}[$, $f(x - \alpha \nabla f(x)) < f(x)$.

Corollaire 31. Ainsi, pour $x_0 \in \mathbb{R}^n$, la suite définie par $x_{k+1} = x_k - \alpha \nabla f(x_k) = x_k + \alpha(b - Ax_k)$ converge vers la solution de $Ax = b$.

Proposition 32 (Pas optimal). La suite définie par $x_{k+1} = x_k + \alpha_k(b - Ax_k)$, où α_k minimise la fonction $g_k(\alpha) = f(x_k - \alpha \nabla f(x_k))$, converge vers la solution de $Ax = b$.

Définition 33. Si le système est incompatible, le problème aux moindres carrés revient à chercher $\bar{x} = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|_2$.

Proposition 34. x est solution du problème aux moindres carrés, si et seulement si il vérifie $A^*Ax = b$.

Théorème 35. Le problème aux moindres carrés possède toujours une solution, et celle-ci est unique ssi $\ker(A) = \{0\}$.

3.3 Méthode de Newton

Supposons f de classe \mathcal{C}^2 . L'idée de cette méthode est de chercher, à chaque étape, $\operatorname{argmin}_h f(x_n + h)$ mais $\operatorname{argmin} \left(\frac{1}{2} \langle Hf(x_n)h, h \rangle + \langle \nabla f(x_n), h \rangle + f(x_n) \right)$. Pour cela, on prend cette fois comme direction de descente $Hf(x_n)^{-1} \nabla f(x_n)$.

Remarque 36. Cette méthode nécessite la résolution d'un système linéaire à chaque étape. En contrepartie, la convergence est quadratique.

Développements

- Ellipsoïde de John-Loewner.
- Théorème des extrema liés.
- Méthode du gradient à pas optimal pour les systèmes linéaires.

Références

- [1] G. Allaire, ANALYSE NUMÉRIQUE ET OPTIMISATION, École Polytechnique.
- [2] V. Beck, J. Malick, G. Peyré, OBJECTIF AGRÉGATION, H&K.
- [3] X. Gourdon, LES MATHS EN TÊTE - ANALYSE, Ellipses.
- [4] A. Pommellet, COURS D'ANALYSE, Ellipses.
- [5] F. Rouvière, PETIT GUIDE DE CALCUL DIFFÉRENTIEL, Cassini.