

Semaine 9, Optimisation sans contrainte

Objectif : $f \in C(\mathbb{R}^n, \mathbb{R})$. On s'intéresse aux 3 questions suivantes

1. Existence de $\bar{x} \in \mathbb{R}^n$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.
(On dit alors que $x \in \operatorname{argmin}_{\mathbb{R}^n} f$.)
2. Unicité de $\bar{x} \in \mathbb{R}^n$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.
(Si il y a existence et unicité, on dit alors que $x = \operatorname{argmin}_{\mathbb{R}^n} f$.)
3. Comment calculer ce point $\bar{x}$ (minimisant f).

Plus généralement, on peut se poser les mêmes questions en remplaçant $\mathbb{R}^n$ par E , espace vectoriel normé.

Point critique

On suppose que $f \in C^1(\mathbb{R}^n, \mathbb{R})$.

Soit $x \in \mathbb{R}^n$. Le point x est un point critique de f si $\nabla f(x) = 0$

Proposition

Soit $\bar{x} \in \mathbb{R}^n$.

$(f(\bar{x}) \leq f(x) \text{ pour tout } x \in \mathbb{R}^n) \implies \nabla f(\bar{x}) = 0$.

Pour trouver un point minimisant f on peut donc commencer par chercher un point annulant ∇f .

Noter que $\nabla f \in C(\mathbb{R}^n, \mathbb{R}^n)$. On est ainsi ramené à la recherche d'un point annulant une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^n$.

Attention : La réciproque est fausse !

Exemples : $f(x) = -x^2$, $f(x) = x^3$

Démonstration de la proposition

Hypothèse : $\bar{x} \in \mathbb{R}^n$, $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.

Soit $y \in \mathbb{R}^n$.

Pour $t \in \mathbb{R}$, on pose $\varphi(t) = f(\bar{x} + ty)$, de sorte que $\varphi \in C^1(\mathbb{R}, \mathbb{R})$.

L'hypothèse donne

$$\varphi(0) = f(\bar{x}) \leq \varphi(t) \text{ pour tout } t \in \mathbb{R}.$$

On en déduit $\varphi'(0) = 0$.

Or $\varphi'(t) = \sum_{i=1}^n \partial_i f(\bar{x} + ty) y_i = \nabla f(\bar{x} + ty) \cdot y$.

On a donc $\varphi'(0) = \nabla f(\bar{x}) \cdot y = 0$ pour tout $y \in \mathbb{R}^n$ et donc $\nabla f(\bar{x}) = 0$.

Convexité

Definition

Soit $f \in C(\mathbb{R}^n, \mathbb{R})$.

1. La fonction f est convexe si
$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) \text{ Pour tout } t \in]0, 1[,$$
$$x, y \in \mathbb{R}^n.$$
2. La fonction f est strictement convexe si
$$f(tx + (1 - t)y) < tf(x) + (1 - t)f(y) \text{ Pour tout } t \in]0, 1[,$$
$$x, y \in \mathbb{R}^n.$$

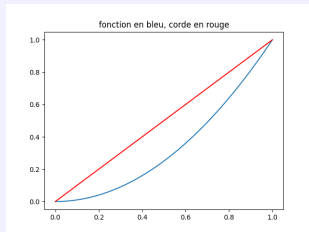


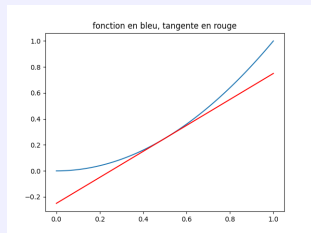
Figure: Fonction convexe

Première caractérisation de la convexité

Proposition

Soit $f \in C^1(\mathbb{R}^n, \mathbb{R})$.

1. La fonction f est convexe si et seulement si
$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x) \text{ Pour tout } x, y \in \mathbb{R}^n.$$
2. La fonction f est strictement convexe si et seulement si
$$f(y) > f(x) + \nabla f(x) \cdot (y - x) \text{ Pour tout } x, y \in \mathbb{R}^n, x \neq y.$$



La fonction f est au dessus de sa tangente.

Démonstration de la 1ere caractérisation ($\Rightarrow$)

On suppose f convexe. Soit $x, y \in \mathbb{R}^n$ et $t \in]0, 1[$. Par convexité

$$f(ty + (1 - t)x) \leq tf(y) + (1 - t)f(x),$$

et donc

$$\frac{f(x + t(y - x)) - f(x)}{t} \leq f(y) - f(x).$$

ou encore, en posant, $\varphi(t) = f(x + t(y - x))$,

$$\frac{\varphi(t) - \varphi(0)}{t} \leq f(y) - f(x).$$

Quand $t \rightarrow 0$, ceci donne $\varphi'(0) \leq f(y) - f(x)$. Enfin, $\varphi'(t) = \nabla f(x + t(y - x)) \cdot (y - x)$ et donc

$$\nabla f(x) \cdot (y - x) \leq f(y) - f(x).$$

Si f est strictement convexe, la même preuve donne

$f\left(\frac{x+y}{2}\right) \geq f(x) + \nabla f(x) \cdot \frac{y-x}{2}$. Puis, on utilise $\frac{1}{2}(f(x) + f(y)) > f\left(\frac{x+y}{2}\right)$ ($x \neq y$).

Démonstration de la 1ere caractérisation ($\Leftarrow$)

On suppose f au dessus de sa tangente en tout point. Soit $x, y \in \mathbb{R}^n$ et $t \in]0, 1[$. On pose $z = tx + (1 - t)y$. Comme f est au dessus de sa tangente au point z ,

$$f(x) \geq f(z) + \nabla f(z) \cdot (x - z),$$

$$f(y) \geq f(z) + \nabla f(z) \cdot (y - z).$$

On multiplie la 1er équation par t , la seconde par $(1 - t)$ et on additionne,

$$tf(x) + (1 - t)f(y) \geq f(z)(t + 1 - t) + \nabla f(z) \cdot (tx + (1 - t)y - z) = f(z).$$

Si f est strictement au dessus de sa tangente, les inégalités sont strictes et on obtient que f est strictement convexe

Deuxième caractérisation de la convexité

Proposition

Soit $f \in C^2(\mathbb{R}^n, \mathbb{R})$.

1. La fonction f est convexe si et seulement si $H_f(x)$ est symétrique positive pour tout $x \in \mathbb{R}^n$ ($H_f(x)$ positive signifie $H_f(x)y \cdot y \geq 0$ pour tout $y \in \mathbb{R}^n$).
2. La fonction f est strictement convexe si $H_f(x)$ est s.d.p pour tout $x \in \mathbb{R}^n$.

Attention : f strictement convexe n'implique pas $H_f(x)$ s.d.p pour tout $x \in \mathbb{R}^n$.

Exemple : $n = 1$, $f(x) = x^4$.

Lemme liminaire

$f \in C^2(\mathbb{R}^n, \mathbb{R})$, $x, y \in \mathbb{R}^n$. On va montrer

$$f(y) = f(x) + \nabla f(x) \cdot (y - x) + \int_0^1 H_f(x + t(y - x))(y - x) \cdot (y - x)(1 - t) dt$$

En effet, pour tout $t \in \mathbb{R}$, on pose $\varphi(t) = f(x + t(y - x))$ de sorte que $\varphi \in C^2(\mathbb{R}, \mathbb{R})$ et

$$\varphi'(t) = \sum_{i=1}^n \partial_i f(x + t(y - x))(y_i - x_i) = \nabla f(x + t(y - x)) \cdot (y - x),$$

$$\begin{aligned} \varphi''(t) &= \sum_{j=1}^n \sum_{i=1}^n \partial_j \partial_i f(x + t(y - x))(y_i - x_i)(y_j - x_j) \\ &= H_f(x + t(y - x))(y - x) \cdot (y - x). \end{aligned}$$

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt = \int_0^1 \varphi''(t)(1 - t) dt + \varphi'(0).$$

$$f(y) - f(x) = \nabla f(x) \cdot (y - x) + \int_0^1 H_f(x + t(y - x))(y - x) \cdot (y - x)(1 - t) dt.$$

Démonstration de la 2eme caractérisation ($\Rightarrow$)

On suppose f convexe. Soit $x, y \in \mathbb{R}^n$. La 1er caractérisation donne

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x),$$

et donc par le lemme linéaire

$$\int_0^1 H_f(x + t(y - x))(y - x) \cdot (y - x)(1 - t) dt \geq 0.$$

Avec $y = x + \varepsilon z$, on obtient

$$\int_0^1 H_f(x + t\varepsilon z) z \cdot z(1 - t) dt \geq 0.$$

Quand $\varepsilon \rightarrow 0$, $H_f(x + t\varepsilon z) \rightarrow H_f(x)$ (uniformément pour $t \in [0, 1]$) et donc

$$\frac{1}{2} H_f(x) z \cdot z \geq 0.$$

La matrice $H_f(x)$ est donc (symétrique) positive.

Démonstration de la 2eme caractérisation ($\Leftarrow$)

On suppose $H_f(x)$ est symétrique positive pour tout $x \in \mathbb{R}^n$. On va montrer que f est au dessus de sa tangente.

Soit $x, y \in \mathbb{R}^n$. Le lemme linéaire donne

$$f(y) - f(x) = \nabla f(x) \cdot (y - x) + \int_0^1 H_f(x + t(y - x))(y - x) \cdot (y - x)(1 - t) dt,$$

et donc, comme $H_f(x + t(y - x))$ est positive,

$$f(y) - f(x) \geq \nabla f(x) \cdot (y - x)$$

La fonction est donc au dessus de sa tangente.

Si $H_f(x)$ est s.d.p. pour tout x , l'inégalité devient stricte (pour $x \neq y$) et f est strictement au dessus de sa tangente, on obtient que f est strictement convexe

Caractérisation de $\bar{x}$ si f est convexe

Soit $f \in C^1(\mathbb{R}^n, \mathbb{R})$. On suppose f convexe. Soit $\bar{x} \in \mathbb{R}^n$. Alors

$$(f(\bar{x}) \leq f(x) \text{ pour tout } x \in \mathbb{R}^n) \iff \nabla f(\bar{x}) = 0.$$

Démonstration :

On a déjà vu le sens ($\Rightarrow$). La convexité de f est ici inutile.

Pour le sens ($\Leftarrow$), La convexité de f donne

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x}) \cdot (x - \bar{x}) \text{ pour tout } x \in \mathbb{R}^n,$$

et donc, comme $\nabla f(\bar{x}) = 0$, $f(x) \geq f(\bar{x})$ pour tout $x \in \mathbb{R}^n$.

Existence de $\bar{x}$

Theorem

$f \in C(\mathbb{R}^n, \mathbb{R})$. On suppose $f(x) \rightarrow +\infty$ quand $\|x\| \rightarrow +\infty$. Alors, il existe $\bar{x}$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.

Démonstration : Il existe $R > 0$ tel que

$$\|x\| > R \implies f(x) > f(0).$$

On pose $B_R = \{x \in \mathbb{R}^n, \|x\| \leq R\}$.

Comme B_R est compact et f est continue, il existe $\bar{x} \in B_R$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in B_R$.

En particulier, $f(\bar{x}) \leq f(0)$ et donc

$$f(\bar{x}) \leq f(x) \text{ pour tout } x \in \mathbb{R}^n.$$

Exemple de non unicité $f(x) = x^2(x - 1)^2$

Unicité de $\bar{x}$

Theorem

$f \in C(\mathbb{R}^n, \mathbb{R})$. On suppose f strictement convexe. Alors, il existe au plus un point $\bar{x}$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.

Démonstration : Soient $\bar{x}$ et $\bar{\bar{x}}$ tels que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$ et $f(\bar{\bar{x}}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.

On a donc $f(\bar{x}) = f(\bar{\bar{x}})$ et si $\bar{x} \neq \bar{\bar{x}}$, par strict convexité de f

$$f\left(\frac{\bar{x} + \bar{\bar{x}}}{2}\right) < \frac{1}{2}(f(\bar{x}) + f(\bar{\bar{x}})) = f(\bar{x}),$$

ce qui est impossible. Donc $\bar{x} = \bar{\bar{x}}$.

Exemple de non existence $f(x) = e^x$

Existence et unicité de $\bar{x}$

Theorem

$f \in C(\mathbb{R}^n, \mathbb{R})$. On suppose $f(x) \rightarrow +\infty$ quand $\|x\| \rightarrow +\infty$ et que f est strictement convexe. Alors, il existe un unique $\bar{x} \in \mathbb{R}^n$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$.

Généralisation : Le même théorème est vrai si $f \in C(E, \mathbb{R})$ et E est un espace de Hilbert (ou un espace de Banach réflexif)

Difficulté supplémentaire pour l'existence si $\dim(E) = +\infty$:

La boule $B_R = \{x \in E, \|x\| \leq R\}$ n'est plus compacte. Il faut utiliser la convexité de f et le fait que B_R est "séquentiellement compacte" pour une topologie dite "faible". Ce résultat est par exemple décrit dans ce cours : <https://www.i2m.univ-amu.fr/perso/thierry.gallouet/licence.d/CDO.pdf>

Exemple d'application (énoncé)

Proposition

Soit $f \in C^1(\mathbb{R}^n, \mathbb{R})$. On suppose qu'il existe $\alpha > 0$ tel que

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq \alpha \|x - y\|^2 \text{ pour tout } x, y \in \mathbb{R}^n.$$

Alors $f(x) \rightarrow +\infty$ quand $\|x\| \rightarrow +\infty$ et f est strictement convexe.

(et donc, il existe un unique $\bar{x} \in \mathbb{R}^n$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$ et il est caractérisé par $\nabla f(\bar{x}) = 0$.)

Exemple d'application (démonstration)

Soient $x, y \in \mathbb{R}^n$. En posant $\varphi(t) = f(x + t(y - x))$,

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt = \int_0^1 \nabla f(x + t(y - x)) \cdot (y - x) dt,$$

et donc

$$\begin{aligned} f(y) - f(x) - \nabla f(x) \cdot (y - x) &= \int_0^1 (\nabla f(x + t(y - x)) - \nabla f(x)) \cdot t(y - x) \frac{1}{t} dt \\ &\geq \int_0^1 \alpha t \|y - x\|^2 dt = \frac{\alpha}{2} \|y - x\|^2. \end{aligned}$$

Ceci montre que f est strictement convexe (1ere caractérisation)

Puis, en prenant $x = 0$,

$$\begin{aligned} f(y) &\geq f(0) - \nabla f(0) \cdot (y) + \frac{\alpha}{2} \|y\|^2 \\ &\geq f(0) - \|\nabla f(0)\| |y| + \frac{\alpha}{2} \|y\|^2 \rightarrow +\infty \text{ quand } \|y\| \rightarrow +\infty \end{aligned}$$

Fonctionnelle quadratique

Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $b \in \mathbb{R}^n$. On suppose que A est s.d.p.. On pose, pour $x \in \mathbb{R}^n$.

$$f(x) = \frac{1}{2}Ax \cdot x - b \cdot x.$$

Un calcul simple montre que $\nabla f(x) = Ax - b$

car $f(x+h) = f(x) + Ax \cdot h - b \cdot h + \frac{1}{2}Ah \cdot h$

Comme A est s.d.p, il existe $\alpha > 0$ tel que

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) = A(x - y) \cdot (x - y) \geq \alpha |x - y|^2.$$

Nous sommes dans le cadre de l'exemple précédent. Il existe un unique $\bar{x} \in \mathbb{R}^n$ tel que $f(\bar{x}) \leq f(x)$ pour tout $x \in \mathbb{R}^n$ et il est caractérisé par $\nabla f(\bar{x}) = 0$, c'est-à-dire

$$A\bar{x} = b$$

Chercher $\bar{x}$ minimisant f est équivalent à résoudre $A\bar{x} = b$