Chapitre 4

Optimisation des fonctions convexes

D6: Un sous-ensemble C de $\mathbb{R}^n$ est dit **convexe** si, pour tout $(a,b) \in C^2$, $[a,b] \subset C$ (c'est-à-dire pour tout $\lambda \in [0,1]$, $\lambda a + (1-\lambda)b \in C$).

D7 : Une fonction réelle f définie sur C convexe est dite **convexe** si, pour tout $(a,b) \in C^2$ et pour tout $\lambda \in [0,1]$, $f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b)$.

Propriétés :
1) Si C est un convexe de $\mathbb{R}^n$ et $(f_i)_{i\in I}$, une famille quelconque de fonctions convexes alors

a) $\sup\limits_{i\in I} f_i$ est convexe ;

b) si I est fini et si $(\lambda_i)_{i\in I}$ est une famille de réels <u>positifs</u>, alors $\sum\limits_{i\in I} \lambda_i f_i$ est convexe.

2) Si C est un convexe de $\mathbb{R}^n$, si f est une fonction convexe de C sur $\mathbb{R}$ et si φ est une fonction convexe <u>croissante</u> sur $\mathbb{R}$, alors $\varphi \circ f$ est une fonction convexe.

Preuve des propriétés :

1)a) $f_i(\lambda a + (1-\lambda)b) \leq \lambda f_i(a) + (1-\lambda) f_i(b) \leq \lambda(\sup f_i)(a) + (1-\lambda)(\sup f_i)(b)$. Ceci est vérifié pour tout $i \in I$ donc $(\sup f_i)(\lambda a + (1-\lambda)b) \leq \lambda(\sup f_i)(a) + (1-\lambda)(\sup f_i)(b)$.

1)b) $\lambda_i \geq 0$ donc $\lambda_i f_i(\lambda a + (1-\lambda)b) \leq \lambda_i \lambda f_i(a) + \lambda_i(1-\lambda) f_i(b)$. En faisant la somme, on a :

$$(\sum_i \lambda_i f_i)(\lambda a + (1-\lambda)b) \leq \lambda(\sum_i \lambda_i f_i)(a) + (1-\lambda)(\sum_i \lambda_i f_i)(b).$$

2) $f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b)$ et, en utilisant la croissance, puis la convexité de φ, on a :

$$\begin{aligned} \varphi\left(f(\lambda a + (1-\lambda)b)\right) &\leq \varphi\left(\lambda f(a) + (1-\lambda) f(b)\right) \\ &\leq \lambda \varphi(f(a)) + (1-\lambda)\varphi(f(b)). \end{aligned}$$

□

Caractérisation des fonctions convexes différentiables

TH5 : Soit Ω un ouvert convexe de $\mathbb{R}^n$ et f une fonction de Ω sur $\mathbb{R}$.
1) Si f est différentiable sur Ω, alors f est convexe si et seulement si, pour tout $(x,y) \in \Omega^2$,

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle.$$

2) Si f est de classe $\mathcal{C}^2$ sur Ω, alors f est convexe sur Ω si et seulement si, pour tout $x \in \Omega$, la matrice $\nabla^2 f(x)$ est positive.

Preuve du TH5 :

1) Si f est convexe, pour $\theta \in]0,1[$, on a :

$$f(x+\theta(y-x)) = f((1-\theta)x + \theta y) \leq (1-\theta)f(x) + \theta f(y)$$

d'où $f(x+\theta(y-x)) - f(x) \leq \theta(f(y) - f(x))$.
Or $f(x+\theta(y-x)) = f(x) + \theta d_x f(y-x) + \theta \|y-x\| \varepsilon(\theta(y-x))$.
En simplifiant par $\theta > 0$, on obtient :

$$d_x f(y-x) + \|y-x\|\varepsilon(\theta(y-x)) \leq f(y) - f(x)$$

et en faisant tendre θ vers 0, on a alors :

$$d_x f(y-x) \leq f(y) - f(x).$$

Réciproquement, si $f(b) \geq f(a) + d_a f(b-a)$ pour tout $(a,b) \in \Omega^2$, alors, avec $b = y$ et $a = x + \theta(y-x)$, puis avec $b = x$ et $a = x + \theta(y-x)$,

$$f(y) \geq f(x+\theta(y-x)) + d_{x+\theta(y-x)} f(y-x) \times (1-\theta)$$

$$f(x) \geq f(x+\theta(y-x)) + d_{x+\theta(y-x)} f(y-x) \times (-\theta).$$

On multiplie la première inégalité par θ, la deuxième par $(1-\theta)$ et on fait la somme :

$$\theta f(y) + (1-\theta) f(x) \geq (\theta + 1 - \theta) f(x + \theta(y-x))$$

soit $f((1-\theta)x + \theta y) \leq (1-\theta)f(x) + \theta f(y)$.

2) On applique la formule de Taylor-Mac-Laurin à l'ordre 2 en x :
il existe $\theta \in]0,1[$ tel que :

$$f(y) = f(x) + d_x f(y-x) + \frac{1}{2} d^2_{x+\theta(y-x)} f(y-x, y-x).$$

Or $d^2_{x+\theta(y-x)} f(y-x,y-x) \geq 0$ donc $f(y) \geq f(x) + d_x f(y-x)$ et f est convexe.

Réciproquement, si f est convexe, alors, pour tout $h \in \mathbb{R}^n$ et pour tout t tel que $x + th \in \Omega$,

$$f(x+th) \geq f(x) + d_x f(th).$$

Or $f(x+th) = f(x) + d_x f(th) + \frac{1}{2}t^2 d_x^2 f(h,h) + t^2\|h\|^2\varepsilon(th)$. D'où :

$$d_x^2 f(h,h) + 2\|h\|^2\varepsilon(th) \geq 0$$

et en faisant $t \to 0$, on obtient finalement $d_x^2 f(h,h) \geq 0$.

□

Minimum global d'une fonction convexe.

TH6 : Soit C un convexe de $\mathbb{R}^n$, f une fonction convexe de C sur $\mathbb{R}$ et $a \in C$, alors
1) un minimum local est un minimum global ;
2) si f est de classe $\mathcal{C}^1$ sur C et si C est ouvert, alors $a \in \mathrm{Arg}_C \min f$ si et seulement si $\nabla f(a) = 0$.

Preuve du TH6 :

1) $f(x+\theta(y-x)) \leq (1-\theta)f(x) + \theta f(y)$ donc $f(x+\theta(y-x)) - f(x) \leq \theta(f(y)-f(x))$.
Si x est un minimum local, alors, pour θ assez petit, on a $f(x+\theta(y-x)) - f(x) \geq 0$. Donc $f(y) - f(x) \geq 0$ et ceci pour y quelconque. Donc x est un minimum global.

2) Si $a \in \mathrm{Arg}_C \min f$, alors $f(y) \geq f(a)$ pour tout $y \in C$. Or C est ouvert, donc a est aussi minimum local et $\nabla f(a) = 0$.
Réciproquement, si $\nabla f(a) = 0$, *i.e.* $d_a f = 0$, comme $f(y) \geq f(a) + d_a f(y-a)$ pour tout $y \in C$, on a donc, pour tout $y \in C$, $f(y) \geq f(a)$ et a est un minimum global de f sur C.

□

Exercices

8. Soit $a \in \mathbb{R}^n$ tel que $\|a\| < 1$ et soit $f : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto (1+\|x\|^2)^{\frac{1}{2}} - \langle a, x\rangle$. Montrer que f est convexe et déterminer $\mathrm{Arg}_{\mathbb{R}^n} \min f$.

9. Soit $\varphi : \Omega \subset \mathbb{R}^n \to \mathbb{R}$. Pour tout $y \in \mathbb{R}^n$, on pose $\varphi^*(y) = \sup\limits_{x\in\Omega}(\langle y, x\rangle - \varphi(x))$.

a) Montrer que φ^* est convexe.

b) Soit $p \in]1, +\infty[$ et $\varphi(x) = \frac{\|x\|^p}{p}$. Montrer que φ est convexe ; déterminer $\varphi^*(y)$ et montrer que $\varphi^{**} = \varphi$. (*On utilisera q tel que* $\frac{1}{p} + \frac{1}{q} = 1$).

c) Soit $\varphi(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$ où A est une matrice symétrique définie positive. Montrer que φ est convexe ; déterminer $\varphi^*(y)$ et montrer que $\varphi^{**} = \varphi$.

$\boxed{\textbf{10.}}$ Soit $C \subset \mathbb{R}^n$ un convexe fermé non vide et $b \in \mathbb{R}^n$. Soit $\pi = \operatorname{Arg}_C \min N$ où $N(x) = \|x - b\|^2$.

a) Montrer que :
i) π est non vide ;
ii) si $p \in \pi$, pour tout $c \in C$, $\langle p - b, p - c\rangle \leq 0$.
(*On utilisera* $F(\lambda) = \|\lambda c + (1-\lambda)p - b\|^2$).
iii) π contient exactement 1 élément, noté $p(b)$.
iv) Si $\langle u - b, u - c\rangle \leq 0$ pour tout $c \in C$, alors $u = p(b)$.

b) Déduire de a) que, $b \notin C$ si et seulement si il existe $w \in \mathbb{R}^n$ tel que

$$\langle w, b\rangle < \inf_{c \in C} \langle w, c\rangle.$$