Optimisation : Préliminaires

Joon Kwon

Master 2 — MathSV

30 septembre 2021

Notations

Dans tout le cours :

- $d \geqslant 1$ un entier
- X un sous-ensemble de $\mathbb{R}^d$
- $f : X \to \mathbb{R}$ une fonction
- $\| \cdot \|$ la norme euclidienne

Soit M une matrice.

- On note $M^\top$ la transposée.
- Dans un contexte de calcul matriciel, un vecteur $x \in \mathbb{R}^d$ est identifié à un vecteur-colonne

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}.$$

- Le produit scalaire entre deux vecteurs $x, y \in \mathbb{R}^d$ se note :

$$\langle x, y \rangle \qquad \text{ou encore} \qquad x^\top y.$$

Minimums et minimiseurs

Definition

Soit $f : X \to \mathbb{R}$ et $x^* \in X$.

- $x^* \in X$ est un minimiseur (global) de f si :
$$f(x^*) = \min_{x \in X} f(x).$$
  $f(x^*)$ est alors un minimum global de f.
- x^* est un minimiseur local de f s'il existe un voisinage V de x tel que :
$$f(x^*) = \min_{x \in X \cap V} f(x).$$
  $f(x^*)$ est alors un minimum local de f.

Gradient et hessienne

Definition

- Si toutes les dérivées partielles de f en x existent, on note $\nabla f(x)$, et on appelle gradient de f en x, le vecteur :

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_d}(x) \end{pmatrix} \in \mathbb{R}^d.$$

- Si toutes les dérivées partielles d'ordre 2 de f en x existent, on note $\nabla^2 f(x)$, et on appelle matrice hessienne de f en x, la matrice :

$$\nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right) \in \mathbb{R}^{d \times d}.$$

Definition (Point critique)

Un point $x \in X$ est dit critique (ou stationnaire) si $\nabla f(x) = 0$.

Remarque

Si f est différentiable (resp. deux fois différentiable) en x, alors les développements de Tayor de f en x aux ordre 1 et 2 s'écrivent respectivement :

$$f(x') = f(x) + \nabla f(x)^\top (x' - x) + o(x' - x)$$

$$f(x') = f(x) + \nabla f(x)^\top (x' - x) + \frac{1}{2}(x' - x)^\top \nabla f^2(x)(x' - x) + o((x' - x)^2).$$

Théorème (de Schwarz)

Si f est deux fois dérivable en x, $\nabla^2 f(x)$ est une matrice symétrique.

Formes quadratiques et matrices symétriques

Definition

Soit A une matrice symétrique. La fonction $x \mapsto x^\top A x$ est appelée forme quadratique associée.

Definition (Matrice (semi-)définie positive)

Soit A une matrice réelle symétrique de taille $d \times d$.

- A est semi-définie positive si : $\forall x \in \mathbb{R}^d, \; x^\top A x \geqslant 0$.
  On note alors $A \succeq 0$.
- A est définie positive si $\forall x \in \mathbb{R}^d \setminus \{0\}, \; x^\top A x > 0$.
  On note alors $A \succ 0$.

Proposition

Soit A une matrice symétrique. A est semi-définie positive (resp. définie positive) si, et seulement ses valeurs propres sont positives (resp. strictement positives).

Convexité

Definition (Ensemble convexe)

Un ensemble $X \subset \mathbb{R}^d$ est convexe si pour tout $x, y \in X$, $[x, y] \subset X$.

Definition (Fonction convexe)

Soit $X \subset \mathbb{R}^d$ un ensemble convexe et $f : X \to \mathbb{R}$.

- f est convexe si pour tous $x, y \in X$ et $\lambda \in]0, 1[$:

$$f(\lambda x + (1 - \lambda)y) \leqslant \lambda f(x) + (1 - \lambda)f(y)$$

- f est strictement convexe si pour tous $x, y \in X$ tels que $x \neq y$ et $\lambda \in]0, 1[$,

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Caractérisations de la convexité

Proposition (Caractérisations d'ordre 1 de la convexité)

Soit $X \subset \mathbb{R}^d$ un ensemble ouvert convexe et $f : X \to \mathbb{R}$ différentiable.

- *f est convexe si, et seulement si :*

$$\forall x, y \in X, \quad f(y) \geqslant f(x) + \nabla f(x)^\top (y - x).$$

- *f est strictement convexe si, et seulement si :*

$$\forall x, y \in X, \quad x \neq y \quad \Longrightarrow \quad f(y) > f(x) + \nabla f(x)^\top (y - x).$$

Proposition (Caractérisations d'ordre 2 de la forte convexité)

Soit $X \subset \mathbb{R}^d$ un ensemble ouvert convexe et $f : X \to \mathbb{R}$ deux fois différentiable.

- *f est convexe si, et seulement si, $\nabla^2 f(x)$ est semi-définie positive pour tout $x \in X$.*
- *f est strictement convexe si, et seulement si, $\nabla^2 f(x)$ est définie positive pour tout $x \in X$.*

Forte convexité

Definition (Forte convexité)

Soit $K > 0$, $X \subset \mathbb{R}^d$ un ensemble convexe. $f : X \to \mathbb{R}$ est K-fortement convexe si pour tous $x, y \in X$ et $\lambda \in]0, 1[$:

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda) f(y) - \frac{K\lambda(1-\lambda)}{2} \|y - x\|^2 .$$

Proposition (Caractérisations d'ordre 1 de la forte convexité)

Soit $K > 0$, $X \subset \mathbb{R}^d$ un ensemble ouvert convexe et $f : X \to \mathbb{R}$ différentiable. f est K-fortement convexe si, et seulement si :

$$\forall x, y \in X, \quad f(y) \geqslant f(x) + \nabla f(x)^\top (y - x) + \frac{K}{2} \|y - x\|^2 .$$

Proposition (Caractérisations d'ordre 2 de la convexité)

Soit $K > 0$, $X \subset \mathbb{R}^d$ un ensemble ouvert convexe et $f : X \to \mathbb{R}$ deux fois différentiable. f est K-fortement convexe si, et seulement si :

$$\forall x \in X,\ \forall u \in \mathbb{R}^d, \quad u^\top \nabla^2 f(x) u \geqslant K \|u\|^2 .$$

Régularité (smoothness)

Definition (Régularité)

Soit $L > 0$, $\mathcal{X} \subset \mathbb{R}^d$ un ensemble ouvert et $f : X \to \mathbb{R}$ différentiable. f est L-régulière (smooth) si ∇f est L-lipschitzienne. Autrement dit :

$$\forall x, y \in X, \quad \|\nabla f(y) - \nabla f(x)\| \leqslant L \|y - x\| .$$

Proposition (Caractérisation d'ordre 1 de la régularité)

Soit $L > 0$, $\mathcal{X} \subset \mathbb{R}^d$ un ensemble ouvert et $f : X \to \mathbb{R}$ une différentiable. f est L-régulière (smooth) si, et seulement si :

$$\forall x, y \in X, \quad \left| f(y) - f(x) - \nabla f(x)^\top (y - x) \right| \leqslant \frac{L}{2} \|y - x\|^2 .$$