UFR
DE MATHÉMATIQUES
ET INFORMATIQUE

UNIVERSITÉ
PARIS
DESCARTES

Optimisation (MT2AM050)
Notes de cours

Master 1 Mathématiques, Modélisation, Apprentissage (MMA)

2020-2021

Quentin Denoyelle
Bureau 812-D
quentin.denoyelle@parisdescartes.fr

Table des matières

Introduction

Ce cours est une introduction aux problèmes d'optimisation. Le cours se focalise sur des problèmes d'optimisation sans contrainte pour les fonctions suffisamment différentiables en dimension finie. Après une introduction des différentes notions mathématiques nécessaires (rappels de calcul différentiel, conditions d'optimalité, convexité, etc.), une part importante est donnée à l'exposition des différents algorithmes classiques d'optimisation, l'étude théorique de leur convergence, ainsi que leur mise en œuvre pratique. Le langage Python sera utilisé en séance de Travaux Pratiques (TP).

L'auteur remercie Bruno Galerne qui est à l'origine de ce poly, ainsi que Joan Glaunès pour ses nombreux conseils.

Les principaux ouvrages de référence pour ce cours sont :

[ROUVIÈRE] François ROUVIÈRE, *Petit guide de calcul différentiel à l'usage de la license et de l'agrégation*, troisième édition, Cassini, 2009

[CIARLET] Philippe G. CIARLET, *Introduction à l'analyse numérique matricielle et à l'optimisation*, cinquième édition, Dunod, 1998

[BOYD & VANDENBERGHE] Stephen BOYD and Lieven VANDENBERGHE *Convex Optimization*, Cambridge University Press, 2004.
Ouvrage téléchargeable gratuitement ici :
`http://stanford.edu/~boyd/cvxbook/`

[ALLAIRE & KABER] Grégoire ALLAIRE et Sidi Mahmoud KABER, *Algèbre linéaire numérique*, Ellipses, 2002

La page web dédiée à ce cours est ici :
`https://qdenoyelle.github.io/M1_Optim/`

Chapitre 1

Rappels et compléments de calculs différentiels

La référence principale pour ce chapitre est le chapitre 2 de [ROUVIÈRE]. Voir aussi le chapitre A.4 de [BOYD & VANDENBERGHE] et le chapitre 7 de [CIARLET].

1.1 Cadre et notation

Dans ce cours on se placera toujours sur des espaces vectoriels normés de dimensions finis que l'on identifie à $\mathbb{R}^n$, munis de leur base canonique notée $(e_1, \dots, e_n)$, $n \geqslant 1$. Par convention, on identifiera très souvent les vecteurs de $\mathbb{R}^n$ à des vecteurs colonnes. On note $\langle \cdot, \cdot \rangle$ le produit scalaire canonique et $\|\cdot\|$ la norme euclidienne associée. On note $\mathcal{M}_{m,n}(\mathbb{R})$ l'ensemble des matrices de taille $m \times n$ à coefficients réelles et $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n,n}(\mathbb{R})$ l'ensemble des matrices carrées de taille $n \times n$. La transposée d'une matrice A est notée A^T. On a donc pour tous $x, y \in \mathbb{R}^n$, $\langle x, y \rangle = x^T y$ et par conséquent, pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$

$$\langle y, Ax \rangle = \left\langle A^T y, x \right\rangle .$$

Remarque (Notation de la transposée)**.** La notation A^T correspond plutôt à une convention anglo-saxonne. Elle a été choisie pour ce polycopié car elle est plus simple à taper en LaTeX. Toutefois les étudiantes et étudiants sont libres d'utiliser la notation classique ${}^tA$ pour leurs prises notes et leurs copies d'examen.

1.2 Différentielle et gradient

1.2.1 Applications linéaires et matrices associées

On désigne par $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ l'ensemble des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Vu la convention sur les vecteurs considérés comme vecteurs colonnes, on identifiera souvent un élément de $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ à une matrice rectangulaire de taille $m \times n$ correspondant à la matrice de l'application dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$: si $\varphi \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ alors pour tout $x \in \mathbb{R}^n$, $\varphi(x) = Ax$ avec A la matrice dont les colonnes sont les images par φ des vecteurs de la base canonique $(e_1, \dots, e_n)$ de la base canonique de $\mathbb{R}^n$

$$\varphi(x) = \varphi\left(\sum_{k=1}^{n} x_k e_k\right) = \sum_{k=1}^{n} x_k \varphi(e_k) = \begin{pmatrix}\varphi(e_1) & \cdots & \varphi(e_n)\end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax.$$

1.2.2 Différentielle

Dans tout ce chapitre, Ω désigne un *ensemble ouvert* de $\mathbb{R}^n$.

Définition 1.1. Soit $f : \Omega \to \mathbb{R}^m$. La fonction f est *différentiable au point* $x \in \Omega$ si il existe une application linéaire (continue) $l \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ telle que pour tout $h \in \mathbb{R}^n$ au voisinage de 0, on ait

$$f(x+h) = f(x) + l(h) + \|h\|\,\varepsilon(h)$$

avec $\lim_{h\to 0} \varepsilon(h) = 0$, *i.e.*, $\|h\|\,\varepsilon(\|h\|) = o(\|h\|)$. Dans ce cas l est unique et est appelée la différentielle de f au point x et notée $\mathrm{d}f(x)$.

La matrice de $\mathrm{d}f(x)$ (dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$) est appelée la matrice jacobienne de f au point x, notée $J_f(x)$.

On dit que f est *différentiable* sur Ω si f est différentiable en tout point de Ω.

On dit que f est *continûment différentiable*, notée $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$, si f est différentiable sur Ω et l'application $x \mapsto \mathrm{d}f(x)$ est continue sur Ω.

La fonction affine $f(y) = f(x) + \mathrm{d}f(x)(y - x)$ est l'approximation à l'ordre 1 de f au voisinage du point x.

Proposition 1.2 (Dérivées partielles)**.** Si $f = (f_1, \dots, f_m)$, avec $f_i : \mathbb{R}^n \mapsto \mathbb{R}$, est différentiable en $x \in \Omega$ alors f admet des dérivées partielles

$$\partial_{x_i} f(x) = \lim_{t\to 0} \frac{f(x + te_i) - f(x)}{t} = \mathrm{d}f(x)(e_i).$$

On a alors

$$\mathrm{d}f(x)(h) = \sum_{i=1}^{n} \partial_{x_i} f(x) h_i = J_f(x)h,$$

où $J_f(x) = (\partial_{x_i} f(x)) = (\partial_{x_i} f_j(x)) \in \mathcal{M}_{m,n}(\mathbb{R})$.

Exercice 1.3. Soit A une matrice de $\mathcal{M}_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$. Montrer que la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ définie par $f(x) = Ax + b$ est différentiable sur $\mathbb{R}^n$ et calculer sa différentielle en tout point $x \in \mathbb{R}^n$. En déduire que f est $\mathcal{C}^1$ (et même $\mathcal{C}^\infty$) sur $\mathbb{R}^n$.

Solution de l'exercice 1.3. Pour tous $x, h \in \mathbb{R}^n$,

$$f(x+h) = A(x+h) + b = f(x) + Ah$$

donc f est différentiable en x et $\mathrm{d}f(x) = A$. Comme $x \mapsto \mathrm{d}f(x)$ est constante, c'est une application $\mathcal{C}^\infty$ donc f est aussi $\mathcal{C}^\infty$.

1.2.3 Gradient

Dans ce cours on s'intéressera plus particulièrement à des fonctions à valeurs réelles, ce qui correspond au cas $m = 1$. Ce cas particulier important, nous amène à définir la notion fondamentale de gradient.

Définition 1.4 (Gradient d'une fonction)**.** Si $f : \Omega \to \mathbb{R}$ est différentiable en $x \in \Omega \subset \mathbb{R}^n$ alors $\mathrm{d}f(x)$ est une forme linéaire sur $\mathbb{R}^n$ et donc il existe un unique vecteur, noté $\nabla f(x) \in \mathbb{R}^n$, tel que

$$\forall h \in \mathbb{R}^n, \quad \mathrm{d}f(x)(h) = \langle \nabla f(x),\, h \rangle\,,$$

et appelé *gradient de* f au point x.

Vu autrement de manière équivalente, la matrice $J_f(x)$ est une matrice ligne de taille $1 \times n$ et la transposée de cette matrice est un vecteur de $\mathbb{R}^n$ appelé *gradient de* f au point x et noté $\nabla f(x)$. Pour tout $h \in \mathbb{R}^n$,

$$\mathrm{d}f(x)(h) = \nabla f(x)^T h = \langle \nabla f(x),\, h \rangle\,.$$

Ainsi, pour $f : \Omega \to \mathbb{R}$, f est différentiable si et seulement si il existe un vecteur $v \in \mathbb{R}^n$ tel que

$$f(x+h) = f(x) + \langle v,\, h \rangle + o(h).$$

et alors $v = \nabla f(x)$.

$\nabla f(x)$ s'interprète comme le vecteur de plus forte augmentation de f au voisinage de x. En particulier, $\nabla f(x)$ est orthogonal au ligne de niveaux de la fonction f.

Exercice 1.5. Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$. Soit $f : \mathbb{R}^n \to \mathbb{R}$ l'application $f : x \mapsto \frac{1}{2}\langle Ax,\, x \rangle$.

1. Montrer que f est différentiable sur $\mathbb{R}^n$ et déterminer $\nabla f(x)$ pour tout x.
2. Quelle est l'expression de ∇f si A est symétrique ?
3. Quel est le gradient de l'application $x \mapsto \frac{1}{2}\|x\|^2$?

Solution de l'exercice 1.5.

1. Pour tous $x, h \in \mathbb{R}^n$,

$$\begin{aligned} f(x+h) &= \frac{1}{2}\langle A(x+h),\, x+h \rangle \\ &= \frac{1}{2}\langle Ax,\, x \rangle + \frac{1}{2}\langle Ax,\, h \rangle + \frac{1}{2}\langle Ah,\, x \rangle + \frac{1}{2}\langle Ah,\, h \rangle \\ &= f(x) + \frac{1}{2}\left\langle (A + A^T)x,\, h \right\rangle + \frac{1}{2}\langle Ah,\, h \rangle\,. \end{aligned}$$

   Or par l'inégalité de Cauchy-Shwarz, $|\frac{1}{2}\langle Ah,\, h \rangle| \leqslant \frac{1}{2}\|Ah\|\|h\| \leqslant \frac{1}{2}\|A\|_{\mathcal{M}_n(\mathbb{R})}\|h\|^2$, donc c'est bien en o de $\|h\|$. D'où f est différentiable en tout point $x \in \mathbb{R}^n$ et

$$\nabla f(x) = \frac{1}{2}(A + A^T)x.$$

2. Si A est symétrique, alors $\nabla f(x) = \frac{1}{2}(A + A^T)x = Ax$.
3. C'est le cas particulier où $A = I_n$ qui est symétrique, donc $\nabla f(x) = x$.

1.2.4 Dérivation des fonctions composées

Théorème 1.6 (Dérivation des fonctions composées)**.** Soient $f : \mathbb{R}^n \to \mathbb{R}^m$ et $g : \mathbb{R}^m \to \mathbb{R}^p$ deux fonctions différentiables. Soit $h = g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ la fonction composée définie par $h(x) = g(f(x))$. Alors h est différentiable sur $\mathbb{R}^n$ et pour tout $x \in \mathbb{R}^n$,

$$\mathrm{d}h(x) = \mathrm{d}g(f(x)) \circ \mathrm{d}f(x).$$

où $\mathrm{d}g(f(x)) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$ et $\mathrm{d}f(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. En terme de matrice jacobienne, cela s'écrit

$$J_h(x) = J_g(f(x))J_f(x).$$

Remarque. On peut énoncer une version locale du résultat précédent car, comme le suggère la formule $\mathrm{d}h(x) = \mathrm{d}g(f(x))\mathrm{d}f(x)$, pour que h soit différentiable en x, il suffit que f soit différentiable en x et que g soit différentiable en $f(x)$.

Exemple 1.7. Déterminons le gradient de l'application $h : \mathbb{R}^n \to \mathbb{R}$ définie par

$$h(x) = g(Ax + b)$$

où A est une matrice de $\mathcal{M}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$ et $g : \mathbb{R}^m \to \mathbb{R}$ est une application différentiable. On a $h(x) = g \circ f(x)$ avec $f(x) = Ax + b$. Comme f est affine on a $\mathrm{d}f(x) = A$ en tout point x. On a donc d'après la règle de dérivation des fonctions composées

$$J_h(x) = J_g(f(x))J_f(x) = J_g(Ax + b)A.$$

Donc $\nabla h(x) = J_h(x)^T = A^T \mathrm{d}g(Ax + b)^T = A^T \nabla g(Ax + b)$.

1.3 Différentielle seconde et matrice hessienne

Définition 1.8 (Différentielle seconde)**.** Soit $f : \Omega \to \mathbb{R}^m$ une application différentiable sur Ω. On dira que f est deux fois différentiable en $a \in \Omega$ si l'application

$$df : x \mapsto \mathrm{d}f(x),$$

est différentiable en a. On note cette différentielle $\mathrm{d}^2 f(a) = \mathrm{d}(\mathrm{d}f)(a)$ et on l'appelle la différentielle seconde de f en a.

On dira que f est deux fois continûment différentiable sur Ω, noté $f \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$, si f est deux fois différentiable en tout point de Ω et $x \mapsto \mathrm{d}^2 f(x)$ est continue sur Ω.

Comme $\mathrm{d}f : \Omega \subset \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, on a ainsi $\mathrm{d}^2 f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ que l'on peut donc identifier à une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$.

Théorème 1.9 (Théorème de Schwarz)**.** Soit $f : \Omega \to \mathbb{R}^m$ une application deux fois différentiable en $a \in \Omega$. Alors $\mathrm{d}^2 f(a) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ est une application bilinéaire symétrique, c'est-à-dire

$$\forall (h, k) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \mathrm{d}^2 f(a)(h, k) = \mathrm{d}^2 f(a)(k, h).$$

En pratique dans ce cours, nous allons très souvent considérer le cas particulier $m = 1$. La différentielle seconde en un point est alors une *forme* bilinéaire symétrique que l'on peut caractérisé par sa matrice dans la base canonique de $\mathbb{R}^n$.

Définition 1.10 (Matrice hessienne)**.** Soit $f : \Omega \to \mathbb{R}$ une application deux fois différentiable en $a \in \Omega$. Alors la forme bilinéaire symétrique $\mathrm{d}^2 f(a)$ peut s'identifier à sa matrice dans la base canonique de $\mathbb{R}^n$, appelée *matrice hessienne de f en a* et notée

$$\nabla^2 f(a) = \left(\mathrm{d}^2 f(e_i, e_j) = \partial^2_{x_i,x_j} f(a)\right)_{1\leqslant i,j\leqslant n}.$$

On a alors par définition

$$\forall (h,k) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \mathrm{d}^2 f(x)(h,k) = \left\langle \nabla^2 f(x)h,\, k\right\rangle = k^T \nabla^2 f(x) h = h^T \nabla^2 f(x) k.$$

Pour ce cours on aura constamment besoin de calculer le gradient et la matrice hessienne de fonctionnelles $f : \Omega \to \mathbb{R}$ deux fois différentiable. En pratique on utilise la proposition suivante.

Proposition 1.11 (La matrice hessienne est la différentielle du gradient)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction différentiable sur Ω et deux fois différentiable au point $a \in \Omega$. Alors, la matrice hessienne $\nabla^2 f(a)$ de f au point a est la matrice jacobienne de l'application gradient $x \mapsto \nabla f(x)$ au point a. Dit autrement

$$\forall k \in \mathbb{R}^n, \quad \mathrm{d}(\nabla f)(a)(k) = \nabla^2 f(a) k.$$

Exercice 1.12. Démontrer la Proposition 1.11 ci-dessus.

Solution de l'exercice 1.12. Posons $g(x) = \nabla f(x)$ pour tout $x \in \Omega$. Alors par définition la différentielle de $g : \mathbb{R}^n \to \mathbb{R}^n$ est caractérisée par sa matrice jacobienne

$$J_g(a) = \left(\partial_{x_j} g_i(a)\right)_{1\leqslant i,j\leqslant n} = \left(\partial_{x_j}\left(\partial_{x_i} f\right)(a)\right)_{1\leqslant i,j\leqslant n} = \left(\partial^2_{x_i,x_j} f(x)\right)_{1\leqslant i,j\leqslant n} = \nabla^2 f(x).$$

Il est difficile de donner une règle de dérivation des fonctions composées pour l'ordre deux. Voici toutefois deux cas particuliers que l'on utilisera à plusieurs reprises pour ce cours et avec lesquels il faut être à l'aise.

Composition avec une fonction scalaire : On considère $f : \Omega \to \mathbb{R}$ une fonction deux fois différentiable et $g : \mathbb{R} \to \mathbb{R}$ une fonction deux fois dérivable. Alors, $h = g \circ f$ est deux fois différentiable et

$$\nabla^2 h(x) = g'(f(x)) \nabla^2 f(x) + g''(f(x)) \nabla f(x) \nabla f(x)^T.$$

Composition avec une fonction affine : Soit $g : \mathbb{R}^n \to \mathbb{R}$ définie par

$$g(x) = f(Ax + b)$$

où A est une matrice de $\mathcal{M}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$ et $f : \mathbb{R}^m \to \mathbb{R}$ est une application deux fois différentiable. Alors g est deux fois différentiable et

$$\nabla^2 g(x) = A^T \nabla^2 f(Ax + b) A,$$

formule qui s'obtient facilement en dérivant l'expression $\nabla g(x) = A^T \nabla f(Ax + b)$ montrée précédemment.

1.4 Formules de Taylor

Les formules de Taylor se généralisent aux fonctions de plusieurs variables. On se limite aux fonctions à valeurs réelles.

Théorème 1.13 (Formules de Taylor pour les fonctions une fois dérivable)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction.

(a) Définition de la différentielle = Formule de Taylor-Young à l'ordre 1 : Si f est différentiable en $x \in \Omega$, alors

$$f(x+h) = f(x) + \langle \nabla f(x),\, h\rangle + \|h\|\varepsilon(h) \quad \text{avec} \lim_{h\to 0} \varepsilon(h) = 0.$$

On considère maintenant un point h fixé tel que le segment $[x, x+h]$ soit inclus dans Ω.

(b) Inégalités des accroissements finis : Si f est continue sur Ω et différentiable sur $]x, x+h[$, alors

$$|f(x+h) - f(x)| \leqslant \sup_{y\in]x,x+h[} \|\nabla f(y)\|\|h\|.$$

(c) Formule de Taylor-Maclaurin : Si f est continue sur Ω et différentiable sur $]x, x+h[$, alors il existe $\theta \in]0, 1[$ tel que

$$f(x+h) = f(x) + \langle \nabla f(x+\theta h),\, h\rangle\,.$$

(d) Formule de Taylor avec reste intégral : Si $f \in \mathcal{C}^1(\Omega)$ alors

$$f(x+h) = f(x) + \int_0^1 \langle \nabla f(x+th),\, h\rangle\, dt.$$

Preuve. On applique les formules de Taylor à la fonction $\varphi(t) = f(x+th)$, $t \in [0, 1]$. □

Théorème 1.14 (Formules de Taylor pour les fonctions deux fois dérivable)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction.

(a) Formule de Taylor-Young à l'ordre 2 : Si f est différentiable dans Ω et deux fois différentiable en $x \in \Omega$, alors

$$f(x+h) = f(x) + \langle \nabla f(x),\, h\rangle + \frac{1}{2}\left\langle \nabla^2 f(x)h,\, h\right\rangle + \|h\|^2\varepsilon(h) \quad \text{avec} \lim_{h\to 0} \varepsilon(h) = 0.$$

On considère maintenant un point h fixé tel que le segment $[x, x+h]$ soit inclus dans Ω.

(b) Formule des accroissements finis généralisée : Si $f \in \mathcal{C}^1(\Omega)$ et f est deux fois différentiable sur $]x, x+h[$, alors

$$|f(x+h) - f(x) - \langle \nabla f(x),\, h\rangle\,| \leqslant \frac{1}{2} \sup_{y\in]x,x+h[} \|\nabla^2 f(y)\|_{\mathcal{M}_n(\mathbb{R})}\|h\|^2.$$

où $\|\cdot\|_{\mathcal{M}_n(\mathbb{R})}$ désigne la norme subordonnée des matrices pour la norme euclidienne.

(c) Formule de Taylor-Maclaurin : Si $f \in \mathcal{C}^1(\Omega)$ et f est deux fois différentiable sur $]x, x+h[$, alors il existe $\theta \in]0, 1[$ tel que

$$f(x+h) = f(x) + \langle \nabla f(x),\, h\rangle + \frac{1}{2}\left\langle \nabla^2 f(x+\theta h)h,\, h\right\rangle.$$

(d) Formule de Taylor avec reste intégral : Si $f \in \mathcal{C}^2(\Omega)$ alors

$$f(x+h) = f(x) + \langle \nabla f(x),\, h\rangle + \int_0^1 (1-t)\left\langle \nabla^2 f(x+th)h,\, h\right\rangle dt.$$

Chapitre 2

Problèmes d'optimisation : Existence et unicité des solutions

La référence principale pour ce chapitre est le chapitre 8 de [CIARLET].

2.1 Cadre et vocabulaire

On appelle *problème d'optimisation* tout problème de la forme

$$\text{Trouver } x^\star \text{ tel que } x^\star \in U \text{ et } \quad f(x^\star) = \min_{x \in U} f(x), \tag{2.1}$$

que l'on note également

$$x^\star \in \operatorname*{argmin}_{x \in U} f(x), \tag{2.2}$$

où U est un sous ensemble de $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction. Ces deux éléments sont des données du problème d'optimisation. On appelle

— f la *fonction objective* (ou encore *fonctionnelle*) du problème d'optimisation,

— l'élément $x^\star \in U$ une solution du problème d'optimisation.

On dira que

— le problème d'optimisation est *sans contraintes* si $U = \mathbb{R}^n$ et *sous contraintes* sinon,

— le problème est *convexe* si f et U sont convexes.

Il est souvent impossible en pratique de trouver explicitement une solution $x^\star$. Le but de l'optimisation est de proposer des algorithmes permettant d'approcher une solution $x^\star$ au sens où, partant d'un vecteur initial $x^{(0)}$ quelconque, on construit explicitement une suite $(x^{(k)})_{k \geqslant 0}$ convergeant vers une solution $x^\star$.

Dans ce cours on s'intéressera à résoudre des problèmes d'optimisation convexes, sans contraintes, et de dimension finie.

On établira dans ce chapitre des conditions d'existence et d'unicité des solutions de problèmes d'optimisation. Dans les chapitres suivants, on s'intéressera à l'élaboration d'algorithmes itératifs pour la résolution effectives de tels problèmes d'optimisation convexes, sans contraintes et de dimension finie.

Bien sûr, les méthodes développées dans ce cours permettent également de trouver les valeurs maximales de fonctions f. Pour cela il suffit de remplacer f par $-f$ puisque

$$\max_{x \in U} f(x) = -\min_{x \in U} -f(x).$$

Extremums des fonctions à valeurs réelles Soit $f : U \to \mathbb{R}$, où $U \subset \mathbb{R}^n$. On dit que la fonction f admet en un point $x \in U$ un *minimum local* (respectivement un *maximum local*) s'il existe un $\varepsilon > 0$ tel que pour tout $y \in U \cap B(x, \varepsilon)$, $f(y) \geqslant f(x)$ (resp. $f(y) \leqslant f(x)$). On dit que la fonction admet un *extremum local* en x si elle admet soit un minimum soit un maximum local en x.

Par abus de langage, on dira que x est un minimum local pour dire que la fonction f admet un minimum local en x.

On dit qu'un minimum local x est strict s'il existe un $\varepsilon > 0$ tel que pour tout $y \in U \cap B(x, \varepsilon)$, $y \neq x$, $f(y) > f(x)$. On définit de même la notion de maximum strict.

Enfin, on dit qu'un minimum x est *global* si pour tout $y \in U$, $f(y) \geqslant f(x)$. Si $W \subset U$, on dira qu'un minimum $x \in W$ est global sur W si pour tout $y \in W$, $f(y) \geqslant f(x)$. On définit de même la notion de maximum global.

2.2 Existence de solutions pour les fonctions coercives et continues

La première question concernant un problème d'optimisation est celle de l'existence d'une solution. Si on cherche à minimiser une fonction $f : U \subset \mathbb{R}^n \to \mathbb{R}$ continue sur U, alors il est bien connue que si U est compact (*i.e.* fermé et borné) la fonction f est bornée et atteint ses bornes sur U. Elle admet donc au moins un minimum global $x^\star \in U$. La notion de fonction coercive permet d'étendre ce type de raisonnement pour des fonctions définies sur des domaines non bornés.

Définition 2.3 (Fonctions coercives)**.** Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite *coercive* si

$$\lim_{\|x\| \to +\infty} f(x) = +\infty. \tag{2.4}$$

Théorème 2.5. Soient U une partie non vide fermée de $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue, coercive si l'ensemble U est non borné. Alors il existe au moins un élément $x^\star \in U$ tel que

$$f(x^\star) = \inf_{x \in U} f(x). \tag{2.6}$$

Preuve. Soit x_0 un point quelconque de U. Supposons U non borné, car sinon, comme mentionné précédemment, on a directement la conclusion par compacité. La coercivité de f entraîne qu'il existe un réel $r > 0$ tel que

$$\|x\| \geqslant r \Rightarrow f(x) > f(x_0). \tag{2.7}$$

Donc $\{x \in U; f(x) = \inf_{x \in U} f(x)\} \subset \{x \in U; f(x) \leqslant f(x_0)\} \subset U \cap \overline{\mathcal{B}(x_0, r)}$. Or on a toujours $\{x \in U; f(x) = \inf_{x \in U \cap \overline{\mathcal{B}(x_0, r)}} f(x)\} \subset \{x \in U; f(x) = \inf_{x \in U} f(x)\}$ d'où chercher l'infimum de f sur U se réduit à le chercher sur $U \cap \overline{\mathcal{B}(x_0, r)}$. Comme l'ensemble $U \cap \overline{B(0, r)}$ est fermé et borné et que f est continue, f est bornée et atteint ses bornes sur le compact $U \cap \overline{B(0, r)}$, ce qui assure donc l'existence d'un minimum (global) dans U (qui est inclus dans $U \cap \overline{B(0, r)}$). $\square$

2.3 Extremums locaux et dérivabilité

On va maintenant chercher à caractériser les minimums (ou maximums) locaux des fonctions différentiables. Dans toute la suite du chapitre Ω désigne un sous-ensemble ouvert de $\mathbb{R}^n$.

2.3.1 Extremums locaux et conditions d'ordre un

On commence par s'intéresser à une condition nécessaire d'extrémalité sur le gradient. Pour cela, on a besoin de la notion de point critique d'une fonction.

Définition 2.8 (Point critique)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction différentiable en $x \in \Omega$. On dira que x est un *point critique* de la fonction f si $\nabla f(x) = 0$.

Alors on a le théorème suivant.

Théorème 2.9 (Condition nécessaire d'extremum local)**.** Soit $f : \Omega \to \mathbb{R}$. Si la fonction f admet un extremum local en $x \in \Omega$ et si elle est différentiable en ce point, alors x est un point critique de f.

Démonstration. Soit $h \in \mathbb{R}^n$. Comme Ω est un ouvert, il existe $\varepsilon > 0$, tel que pour tout $t \in]-\varepsilon, \varepsilon[$, $x + th \in \Omega$. La fonction $\varphi : t \mapsto f(x + th)$ est donc définie sur $]-\varepsilon, \varepsilon[$. Elle est dérivable en $t = 0$, et d'après la formule de dérivation des fonctions composées, $\varphi'(0) = \mathrm{d}f(x)h = \langle \nabla f(x), h\rangle$. φ ayant un extremum en $t = 0$, on sait que $\varphi'(0) = 0$ (en utilisant le cas bien connu des fonctions réelles de la variable réelle). D'où $\langle \nabla f(x), h\rangle = 0$. Ceci est vrai pour tout $h \in \mathbb{R}^n$, donc on a bien $\nabla f(x) = 0$. □

Exercice 2.10.
1. Montrer que la conclusion du théorème est fausse si Ω n'est pas un ouvert.
2. Montrer que la réciproque du théorème est fausse : $\nabla f(x) = 0$ n'implique pas que x soit un extremum local.

Solution de l'exercice 2.10.
1. On considère par exemple pour $n = 1$, $f(x) = x$ sur le domaine fermé $\Omega = [0, 1]$ qui admet un minimum local en $x = 0$.
2. Par exemple, toujours sur $\mathbb{R}$, $f(x) = x^3$ a une dérivée nulle en 0 mais 0 n'est pas un extremum.

On remarque donc que s'intéresser uniquement aux points d'annulation du gradient n'est pas suffisant pour déterminer la nature des extremas locaux (ni même pour les identifier). Pour cela, il nous faut une information sur la courbure locale de la fonction autour des points critiques. Cette information est contenue dans la différentielle seconde.

2.3.2 Extremums locaux et conditions d'ordre deux

On s'intéresse maintenant aux conditions nécessaires et suffisantes faisant intervenir la matrice hessienne.

Théorème 2.11 (Condition nécessaire de minimum local pour la dérivée seconde)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction différentiable dans Ω. Si la fonction f admet un minimum local en un point $x \in \Omega$ et si f est deux fois différentiable en x, alors pour tout $h \in \mathbb{R}^n$,

$$\langle \nabla^2 f(x)h, h\rangle \geqslant 0 \quad \text{(ou encore } \mathrm{d}^2 f(x)(h, h) \geqslant 0\text{)},$$

autrement dit la matrice hessienne $\nabla^2 f(x)$ est positive.

Preuve. Soit $h \in \mathbb{R}^n$. Comme x est un minimum local, il existe $\varepsilon > 0$ tel que pour tout $t \in]-\varepsilon, \varepsilon[$, $x + th \in \Omega$ et $f(x+th) \geqslant f(x)$. La formule de Taylor-Young donne

$$f(x+th) = f(x) + t\langle \nabla f(x), h\rangle + \frac{t^2}{2}\langle \nabla^2 f(x)h, h\rangle + t^2\|h\|^2\varepsilon(th)$$

avec $\lim_{y\to 0} \varepsilon(y) = 0$. D'après le Théorème 2.9, on a également $\nabla f(x) = 0$. Ainsi,

$$0 \leqslant f(x+th) - f(x) = \frac{t^2}{2}\langle \nabla^2 f(x)h, h\rangle + t^2\|h\|^2\varepsilon(th)$$

En divisant par $\frac{t^2}{2}$ on en déduit que pour tout $t \neq 0$,

$$\left\langle \nabla^2 f(x)h,\, h\right\rangle + 2\left\|h\right\|^2 \varepsilon(th) \geqslant 0.$$

En faisant tendre t vers 0 on obtient bien que $\langle \nabla^2 f(x)h, h\rangle \geqslant 0$. □

Cependant il est possible que la matrice hessienne soit positive en un point critique sans que ce dernier soit un minimum local. On pensera par exemple à l'exemple de la selle de singe $f : (x, y) \in \mathbb{R}^2 \mapsto x^3 - 3xy^2$ en $(0, 0)$ (exemple traité en TD/TP). Il faut donc des conditions plus fortes sur la matrice hessienne. Commençons par une définition qui précise la notion de point critique.

Définition 2.12 (Point critique non dégénéré)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction deux fois différentiable en $x \in \Omega$ un point critique de f. Alors on dira que x est un point critique *dégénéré* si le discriminant de $\nabla^2 f(x)$, la matrice hessienne de f en x, est nulle. Autrement dit si $\nabla^2 f(x)$ admet au moins une valeur propre nulle. Sinon on dira que x est un point critique *non dégénéré*.

C'est donc le signe des valeurs propres de la hessienne en un point critique qui va déterminer sa nature.

Théorème 2.13 (Condition suffisante de minimum local pour la dérivée seconde)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction deux fois différentiable sur Ω et x un point critique de f non dégénéré. Si

$$\forall h \in \mathbb{R}^n \setminus \{0\}, \quad \langle \nabla^2 f(x)h, h\rangle > 0$$

c'est-à-dire si les valeurs propres de $\nabla^2 f(x)$ sont toutes strictement positives, ou dit autrement si $\nabla^2 f(x)$ est une matrice symétrique définie positive, alors la fonction f admet un minimum local *strict* en x.

Preuve. Comme $\nabla^2 f(x)$ est définie positive, il existe un nombre $\alpha > 0$ tel que

$$\forall h \in \mathbb{R}^n, \quad \langle \nabla^2 f(x)h, h\rangle \geqslant \alpha\|h\|^2$$

(en prenant par exemple $\alpha = \lambda_{\min}(\nabla^2 f(x))$ la plus petite valeur propre de $\nabla^2 f(x)$). D'après la formule de Taylor-Young,

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}\langle \nabla^2 f(x)h, h\rangle + \|h\|^2\varepsilon(h) \geqslant f(x) + (\frac{1}{2}\alpha - |\varepsilon(h)|)\|h\|^2$$

avec $\lim_{h\to 0} \varepsilon(h) = 0$. Soit $r > 0$ tel que pour tout $h \in B(0, r)$, $|\varepsilon(h)| < \frac{1}{2}\alpha$. Alors, pour tout $h \in B(0, r)$, $f(x+h) > f(h)$, donc x est bien un minimum strict. □

On notera que si $\nabla^2 f(x)$ est définie négative en un point critique x alors x un maximum local strict. Dans le cas où x est un point critique non dégénéré tel que $\nabla^2 f(x)$ admet des valeurs propres à la fois positives et négatives, on dira que x est un *point col*.

Exemple 2.14. La fonction $f : (x,y) \in \mathbb{R}^2 \mapsto x^2 - y^2$ admet un point critique non dégénéré en $(0,0)$ qui est un point col car les valeurs propres de $\nabla^2 f(0,0)$ sont -2 et 2. Au voisinage de $(0,0)$ la fonction ressemble à une selle de cheval puisque $t \mapsto f(t,0) = t^2$ et $t \mapsto f(0,t) = -t^2$.

Il est tout de même possible de démontrer qu'un point critique est un minimum local sous condition unique de positivité de la hessienne. Pour cela il faut imposer cette positivité sur tout un voisinage du point critique comme le fait remarquer la proposition suivante.

Proposition 2.15. Soit $f : \Omega \to \mathbb{R}$ une fonction $\mathcal{C}^1$ et deux fois différentiable sur Ω et x un point critique de f. S'il existe $r > 0$ tel que pour tout $y \in \mathcal{B}(x,r)$ on ait

$$\forall h \in \mathbb{R}^n, \quad \langle \nabla^2 f(y)h, h\rangle \geqslant 0,$$

c'est-à-dire la matrice hessienne de f est positive en tout point d'un voisinage de x, alors la fonction f admet en minimum local en x.

Preuve. Soit $x' \in \mathcal{B}(x,r)$. Alors, comme f est $\mathcal{C}^1$ et deux fois différentiable sur Ω, d'après la formule de Taylor-Maclaurin il existe $y \in]x, x'[$ tel que

$$\begin{aligned} f(x') &= f(x) + \langle \nabla f(x),\, x' - x\rangle + \frac{1}{2}\left\langle \nabla^2 f(y)(x'-x),\, x'-x\right\rangle, \\ &= f(x) + \underbrace{\frac{1}{2}\left\langle \nabla^2 f(y)(x'-x),\, x'-x\right\rangle}_{\geqslant 0 \text{ par hypothèse}}, \end{aligned}$$

donc $f(x') \geqslant f(x)$. D'où x est bien un minimum local de f. $\square$

2.4 Ensembles convexes

On rappelle qu'étant donnés deux vecteurs x et $y \in \mathbb{R}^n$, $[x,y]$ désigne le segment entre x et y, à savoir

$$[x,y] = \{tx + (1-t)y,\ t \in [0,1]\}.$$

Définition 2.16 (Ensembles convexes)**.** On dira qu'un ensemble $U \subset \mathbb{R}^n$ est *convexe* si

$$\forall x, y \in U, \quad [x,y] \subset U,$$

soit encore si

$$\forall x, y \in U, \forall t \in [0,1], \quad tx + (1-t)y \in U.$$

(autrement dit U contient tout segment rejoignant n'importe quel couple de ses points).

Exemple 2.17.

— Un sous-espace vectoriel est convexe

— Un hyperplan est convexe.

— La boule unité d'une norme est convexe.
— Toute intersection d'ensembles convexes est convexe.
— Un hyper-rectangle $[a_1, b_1] \times \cdots \times [a_n, b_n]$ est convexe. Plus généralement le produit cartésien $C = C_1 \times \cdots \times C_k$ d'ensembles convexes $C_1 \subset \mathbb{R}^{n_1}, \ldots, C_n \subset \mathbb{R}^{n_k}$ est un ensemble convexe de l'espace produit $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k}$.
— L'image d'un ensemble convexe par une application linéaire est convexe (voir exercice ci-dessous). En particulier, les translations, rotations, dilatations, projections d'ensembles convexes sont convexes.

Exercice 2.18. Soit $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire.

1. Montrer que si $U \subset \mathbb{R}^n$ est convexe alors l'image directe $W = \varphi(U) = \{\varphi(x),\ x \in U\}$ de U par φ est un ensemble convexe de $\mathbb{R}^m$.
2. Montrer que si $W \subset \mathbb{R}^m$ est un ensemble convexe alors l'image réciproque $U = \varphi^{-1}(W) = \{x \in \mathbb{R}^n,\ \varphi(x) \in W\}$ est un ensemble convexe de $\mathbb{R}^n$.

Solution de l'exercice 2.18.

1. Soient y_1 et $y_1 \in W$. Alors il existe x_1 et x_2 dans U tel que $y_1 = \varphi(x_1)$ et $y_2 = \varphi(x_2)$. Soit $t \in [0, 1]$. Alors, par linéarité,
$$tw_1 + (1-t)w_2 = t\varphi(x_1) + (1-t)\varphi(x_2) = \varphi(tx_1 + (1-t)x_2).$$
Or comme U est convexe, $tx_1 + (1-t)x_2 \in U$, et donc, $ty_1 + (1-t)y_2 \in W = \varphi(U)$. $W = \varphi(U)$ est bien un ensemble convexe.
2. Soient x_1 et $x_2 \in U = \varphi^{-1}(W)$ et $t \in [0, 1]$. Comme W est convexe
$$\varphi(tx_1 + (1-t)x_2) = t\varphi(x_1) + (1-t)\varphi(x_2) \in W.$$
Donc $tx_1 + (1-t)x_2 \in U = \varphi^{-1}(W)$, c'est bien un ensemble convexe.

Théorème 2.19 (Condition nécessaire de minimum local sur un ensemble convexe)**.** Soit $f : \Omega \to \mathbb{R}$ et U une partie convexe de Ω. Si la fonction f est différentiable en un point $x \in U$ et si elle admet en x un minimum local par rapport à l'ensemble U, alors
$$\forall y \in U, \quad \langle \nabla f(x), y - x\rangle \geqslant 0 \quad \text{(ou encore } \mathrm{d}f(x)(y-x) \geqslant 0).$$
En particulier si U est un sous-espace affine de $\mathbb{R}^n$ (c'est-à-dire $U = x + F$ avec F un sous-espace vectoriel de V), alors
$$\forall y \in U, \quad \langle \nabla f(x), y - x\rangle = 0 \quad \text{(ou encore } \mathrm{d}f(x)(y-x) = 0).$$

Preuve. Soit $y = x + h$ un point quelconque de l'ensemble U. U étant convexe, les points $x + th$, $t \in [0, 1]$, sont tous dans U. La dérivabilité de f en x permet d'écrire
$$f(x + th) - f(x) = t\langle \nabla f(x), h\rangle + t\|h\|\varepsilon(th),$$
avec $\lim_{t\to 0} \varepsilon(th) = 0$. Comme le membre de gauche est positif, on a nécessairement $\langle \nabla f(x), h\rangle = \langle \nabla f(x), y - x\rangle \geqslant 0$ (dans le cas contraire pour t assez petit le membre de droite serait < 0).

Les cas des sous-espaces affines $U = u + F$, on remarque que si $x + h \in U$ alors $x - h$ appartient également à U et donc on a la double inégalité $\langle \nabla f(x), h\rangle \geqslant 0$ et $\langle \nabla f(x), -h\rangle \geqslant 0$ et donc $\langle \nabla f(x), h\rangle = 0$. □

Remarque. L'interprétation géométrique du théorème précédent est très importante. Si u est un minimum local par rapport au convexe U tel que $\nabla f(u) \neq 0$, alors $\nabla f(u)$ est orienté vers l'intérieur du convexe. En effet, la condition $\langle \nabla f(u), v - u \rangle \geqslant 0$ signifie que l'angle formé par les vecteurs $\nabla f(u)$ et $v - u$ est un angle aigu. Dans le cas d'un espace affine $U = u + F$, cela revient à une condition d'orthogonalité $\nabla f(u) \in F^{\perp}$. Dans ce cours on ne considérera pas de problème de minimisation sous contrainte convexe. En revanche, on sera amené à constamment minimiser des fonctionnelles sur des sous-espaces affines, et en premier lieu des droites.

2.5 Fonctions convexes

2.5.1 Définition et exemples

Définition 2.20 (Fonctions convexes)**.** Soit $U \subset \mathbb{R}^n$ un ensemble convexe. Soit $f : U \to \mathbb{R}$ une fonction à valeurs réelles.

— f est *convexe* si

$$\forall x, y \in U,\ \forall t \in [0,1], \quad f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y).$$

— f est *strictement convexe* si

$$\forall x, y \in U,\ x \neq y,\ \forall t \in]0,1[, \quad f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

Une fonction f est (strictement) *concave* si son opposée $x \mapsto -f(x)$ est (strictement) convexe.

Remarque. On peut également restreindre t à $]0,1[$ pour la définition de la convexité.

Voici quelques exemples de fonctions convexes :

— Sur $\mathbb{R}$, la fonction $x \mapsto x^2$ est strictement convexe.
— Sur $\mathbb{R}$, la fonction $x \mapsto |x|$ est convexe mais pas strictement convexe.
— De même, sur $\mathbb{R}^n$, la fonction $x \mapsto \|x\|^2$ est strictement convexe, la fonction $x \mapsto \|x\|_1$ est convexe mais pas strictement convexe.
— Le sup d'une famille quelconque de fonctions convexes est convexe.
— La composée d'une fonction affine et d'une fonction convexe est convexe (voir ci-dessous).

Exercice 2.21. Soient $A \in M_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$. Montrer que si $f : \mathbb{R}^m \to \mathbb{R}$ est convexe alors la fonction $g : \mathbb{R}^n \to \mathbb{R}$ définie par $g(x) = f(Ax + b)$ est également convexe.

Solution de l'exercice 2.21. Soient $x, y \in \mathbb{R}^n$ et $t \in [0,1]$. On a, par linéarité puis convexité,

$$\begin{aligned} g(tx + (1-t)y) &= f(t(Ax+b) + (1-t)(Ay+b)), \\ &\leqslant tf(Ax+b) + (1-t)f(Ay+b), \\ &= tg(x) + (1-t)g(y). \end{aligned}$$

Donc g est bien convexe.

Une fonction convexe a automatiquement une certaine régularité, comme le montre le théorème suivant.

Théorème 2.22 (Continuité des fonctions convexes)**.** Soit $U \subset \mathbb{R}^n$ un ensemble convexe d'intérieur non vide et $f : U \to \mathbb{R}$ une fonction convexe sur U. Alors f est continue sur l'intérieur de U.

On admet la preuve de ce résultat qui se démontre en revenant au cas des fonctions de la variable réelle en restreignant f dans chaque direction. On notera qu'une fonction convexe peut être discontinue au bord de son domaine (par valeur supérieure).

Définition 2.23. Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$. On appelle sous-ensemble de niveau α de f l'ensemble

$$C_\alpha = \{x \in U,\ f(x) \leqslant \alpha\}.$$

Proposition 2.24. Soit $U \subset \mathbb{R}^n$ un ensemble convexe et $f : U \to \mathbb{R}$ une fonction convexe sur U. Alors pour tout $\alpha \in \mathbb{R}$, l'ensemble C_α est convexe. En particulier, l'ensemble des minimums globaux de f est un ensemble convexe (qui peut-être vide).

Preuve. Soit $\alpha \in \mathbb{R}$. Soient x_1 et x_2 dans C_α et $t \in [0, 1]$. Alors, comme f est convexe,

$$f(tx_1 + (1-t)x_2) \leqslant tf(x_1) + (1-t)f(x_2) \leqslant t\alpha + (1-t)\alpha = \alpha$$

donc $tx_1 + (1-t)x_2 \in C_\alpha$ ce qui prouve bien que C_α est convexe. L'ensemble des minimums globaux de f n'est autre que le sous-ensemble de niveau $\alpha^* = \min_{x\in U} f(x)$ de f, et il est donc bien convexe. $\square$

Remarque. La réciproque de la proposition précédente est fausse. Il existe des fonctions non convexes dont tous les sous-ensembles de niveaux sont convexes, comme par exemple $x \mapsto -e^{-x}$ sur $\mathbb{R}$. Un autre exemple : Sur $\mathbb{R}$, tous les ensembles de niveaux de la fonction $\mathbb{1}_{[0,+\infty[}$ sont convexes mais cette fonction n'est pas convexe. En effet,

$$C_\alpha = \{x \in \mathbb{R},\ \mathbb{1}_{[0,+\infty[}(x) \leqslant \alpha\} = \begin{cases} \emptyset & \text{si } \alpha < 0, \\ ]-\infty, 0[& \text{si } \alpha \in [0, 1[, \\ \mathbb{R} & \text{si } \alpha > 1, \end{cases}$$

qui sont bien tous des ensembles convexes. En revanche la fonction n'est pas convexe car par exemple

$$1 = f(0) > \frac{1}{2}f(1) + \frac{1}{2}f(-1) = \frac{1}{2}.$$

2.5.2 Caractérisation des fonctions convexes différentiables et deux fois différentiables

Avant de considérer l'influence de la convexité sur l'existence et l'unicité de minimums, nous donnons des caractérisations de la notion de convexité pour les fonctions différentiables et deux fois différentiables.

Le théorème ci-dessous exprime le fait qu'une fonction différentiable est convexe si et seulement si son graphe est au-dessus de chacun des ses plans tangents.

Théorème 2.25 (Convexité et différentiabilité)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction différentiable dans l'ouvert Ω et soit $U \subset \Omega$ un sous-ensemble convexe.

(a) La fonction f est convexe sur U si et seulement si pour tout $x, y \in U$,

$$f(y) \geqslant f(x) + \langle \nabla f(x), y - x\rangle \quad \text{(ou encore } f(y) \geqslant f(x) + \mathrm{d}f(x)(y-x)\text{)}.$$

(b) La fonction f est strictement convexe sur U si et seulement si pour tout $x, y \in U$, $x \neq y$,

$$f(y) > f(x) + \langle \nabla f(x), y - x\rangle \quad \text{(ou encore } f(y) > f(x) + \mathrm{d}f(x)(y-x)\text{)}.$$

Preuve. (a) : $\Rightarrow$: Soient x, y deux points distincts de U et $t \in]0,1[$. Comme f est convexe,

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

et on a donc

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

En passant à la limite $t \to 0$ on a

$$\langle \nabla f(x), y - x\rangle = \lim_{t\to 0} \frac{f(x+t(y-x)) - f(x)}{t} \leqslant f(y) - f(x).$$

$\Leftarrow$: Réciproquement supposons que pour tout $x, y \in U$,

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

Soient x, y deux points distincts de U et $t \in]0,1[$. En appliquant l'inégalité aux deux couples $(tx + (1-t)y, y)$ et $(tx + (1-t)y, x)$ on a

$$f(y) \geqslant f(tx + (1-t)y) + \langle \nabla f(tx + (1-t)y), t(y-x)\rangle,$$

et

$$f(x) \geqslant f(tx + (1-t)y) + \langle \nabla f(tx + (1-t)y), (1-t)(x-y)\rangle.$$

En multipliant par $(1-t)$ et t ces deux inégalités, on obtient en les sommant

$$tf(x) + (1-t)f(y) \geqslant f(tx + (1-t)y),$$

donc f est bien convexe.

(b) : La preuve de l'implication indirecte est identique en remplaçant les inégalités larges par des inégalités strictes. En revanche pour l'implication directe, le passage à la limite change les inégalités strictes en inégalités larges, donc on ne peut pas conclure aussi rapidement. Pour cela on se donne cette fois-ci deux poids $0 < t < \omega < 1$. Alors, comme f est strictement convexe et $(1-t)x + ty \in [x, (1-\omega)x + \omega y]$ avec

$$(1-t)x + ty = \frac{\omega - t}{\omega} x + \frac{t}{\omega}((1-\omega)x + \omega y)$$

$$f((1-t)x + ty) < \frac{\omega - t}{\omega} f(x) + \frac{t}{\omega} f((1-\omega)x + \omega y).$$

D'où,

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

En passant à la limite $t \to 0$ on a alors

$$\langle \nabla f(x), y - x\rangle \leqslant \frac{f(x + \omega(y-x)) - f(x)}{\omega} < f(y) - f(x).$$

$\square$

Théorème 2.26 (Convexité et dérivabilité seconde)**.** Soit $f : \Omega \to \mathbb{R}$ une fonction $\mathcal{C}^1(\Omega)$ et deux fois différentiable et soit $U \subset \Omega$ un sous-ensemble convexe.

(a) La fonction f est convexe sur U si et seulement si pour tout $x, y \in U$,

$$\langle \nabla^2 f(x)(y-x), y-x \rangle \geqslant 0$$

(b) Si pour tout $x, y \in U$, $x \neq y$,

$$\langle \nabla^2 f(x)(y-x), y-x \rangle > 0$$

alors f est strictement convexe sur U.

En particulier, si $\Omega = U$ est un ouvert convexe, alors

(a) f est convexe sur Ω si et seulement si pour tout $x \in \Omega$ la matrice hessienne $\nabla^2 f(x)$ est positive.

(b) Si pour tout $x \in \Omega$ la matrice hessienne $\nabla^2 f(x)$ est définie positive, alors f est strictement convexe sur Ω.

Preuve. Soient x et $y = x + h$ deux points distincts de U. Alors, comme f est deux fois différentiable, d'après la formule de Taylor-Maclaurin il existe $z \in]x, x+h[$ tel que

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle \nabla^2 f(z) h, h \rangle.$$

Mais $z \in]x, x+h[$, donc il existe $t \in]0,1[$ tel que $z = tx + (1-t)y$, soit $z - x = (1-t)(y-x) = (1-t)h$. Ainsi

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \frac{1}{(1-t)^2} \langle \nabla^2 f(z)(z-x), (z-x) \rangle.$$

Si par hypothèse $\langle \nabla^2 f(z)(z-x), (z-x) \rangle$ est positif (resp. strictement positif) on déduit du théorème de caractérisation des fonctions convexes différentiables que f est convexe (resp. strictement convexe).

Il reste à montrer que si f est convexe alors pour tout $x, y \in U$, $\langle \nabla^2 f(x)(y-x), y-x \rangle \geqslant 0$. Soient $x, y = x + h \in U$. En appliquant la formule de Taylor-Young en x pour l'accroissement th (avec $t \in [0,1]$),

$$f(x+th) = f(x) + t \langle \nabla f(x), h \rangle + \frac{t^2}{2} \langle \nabla^2 f(x) h, h \rangle + t^2 \|h\|^2 \varepsilon(th)$$

avec $\lim_{t \to 0} \varepsilon(th) = 0$. Donc

$$0 \leqslant f(x+th) - f(x) - t \langle \nabla f(x), h \rangle = \frac{t^2}{2} \left(\langle \nabla^2 f(x) h, h \rangle + 2 \|h\|^2 \varepsilon(th) \right)$$

et on en déduit que $\langle \nabla^2 f(x) h, h \rangle \geqslant 0$ avec le raisonnement habituel.

Le cas où $U = \Omega$ est une conséquence directe. On peut aussi le démontrer rapidement en étudiant la différence entre f et ses approximations au premier ordre. En effet, pour $x \in \Omega$,

$$g(y) = f(y) - f(x) - \langle \nabla f(x), y - x \rangle$$

est une fonction convexe (en tant que somme de fonctions convexes), deux fois différentiable et telle que $\nabla^2 f(y) = \nabla^2 g(y)$. Comme $g(y) \geqslant 0$ et que $g(x) = 0$, x est un minimum global de f et donc nécessairement pour tout $h \in \mathbb{R}^n$, $\langle \nabla^2 f(x) h, h \rangle \geqslant 0$. $\square$

2.5.3 Problèmes d'optimisation convexes

On rappelle qu'un problème d'optimisation

$$\text{Trouver } x^\star \text{ tel que } x^\star \in U \text{ et } \quad f(x^\star) = \min_{x \in U} f(x),$$

est dit *convexe* si U et f sont convexes. Vis-à-vis de l'optimisation, la convexité joue un rôle crucial puisqu'elle permet d'assurer qu'un minimum local est en fait un minimum global, comme précisé par le résultat suivant.

Théorème 2.27 (Minimum de fonctions convexes)**.** Soit $U \subset \mathbb{R}^n$ un ensemble convexe.

(a) Si une fonction convexe $f : U \to \mathbb{R}$ admet un minimum local en un point x, elle y admet en fait un minimum global sur U.

(b) Une fonction $f : U \to \mathbb{R}$ strictement convexe admet au plus un minimum local qui est en fait un minimum global strict.

(c) Soit $f : \Omega \to \mathbb{R}$ une fonction convexe différentiable définie sur un ouvert convexe $\Omega \subset \mathbb{R}^n$. Alors un point $x \in \Omega$ est un minimum global de f si et seulement si x est un point critique.

(d) Soit $f : \Omega \to \mathbb{R}$ une fonction différentiable sur un ouvert Ω contenant U et telle que f est convexe sur U. Alors $x \in U$ est un minimum de f sur U si et seulement si pour tout $y \in U$,

$$\langle \nabla f(x), y - x \rangle \geqslant 0 \quad (\text{ou encore } \mathrm{d}f(x)(y-x) \geqslant 0.$$

En particulier, si $U = x + F$ est un sous-espace affine, alors $x \in U$ est un minimum de f sur U si et seulement si pour tout $y \in U$,

$$\langle \nabla f(x), y - x \rangle = 0 \quad (\text{ou encore } \mathrm{d}f(x)(y-x) = 0.$$

Preuve. (a) Soit y un point quelconque de U. Comme précédemment, la convexité entraîne que

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

pour tout $t \in]0, 1[$. Comme x est un minimum local, il existe un t_0 assez petit tel que $f(x + t_0(y-x)) - f(x) \geqslant 0$. Mais alors, $f(y) - f(x) \geqslant 0$, donc x est bien un minimum global.

(b) Si f est strictement convexe et que x est un minimum local de f, alors pour $y \neq x$ le raisonnement précédent donne l'existence d'un $t_0 > 0$ tel que

$$f(y) - f(x) > \frac{f(x + t_0(y-x)) - f(x)}{t_0} \geqslant 0.$$

Donc $y \neq x$ implique $f(y) > f(x)$. x est donc bien un minimum strict qui est global et unique.

(c) On sait que $\nabla f(x) = 0$ est une condition nécessaire pour être un minimum global. Montrons que c'est une condition suffisante si f est convexe. D'après le Théorème 2.25, si $x \in \Omega$ est tel que $\nabla f(x) = 0$, alors pour tout $y \in \Omega$, $f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle = f(x)$, donc x est bien un minimum global.

(d) C'est le même raisonnement. La condition est nécessaire d'après le Théorème 2.19, et si elle est vérifiée, alors d'après le Théorème 2.25 $f(y) \geqslant f(x) + \langle \nabla f(x), y - x \rangle \geqslant f(x)$ donc x est bien un minimum global sur U. □

Remarque.
- Une fonction non strictement convexe peut admettre plusieurs minimums locaux. Cependant, comme on l'a vu, l'ensemble des minimums globaux forme un ensemble convexe.
- Le théorème précédent est fondamental pour la suite de ce cours. Sauf exception, en pratique on ne s'intéressera qu'à des problèmes d'optimisation convexes.

2.6 Etude des fonctionnelles quadratiques

Définition 2.28. On appelle *fonctionnelle quadratique* toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ de la forme

$$f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c,$$

où $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice carré symétrique, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$.

Remarque. Le fait d'imposer A symétrique dans la définition précédente n'est pas restrictif car on peut facilement se ramener à ce cas. En effet si A est simplement une matrice carré, on a pour tout $x \in \mathbb{R}^n$, $\langle Ax,\, x\rangle = \left\langle x,\, A^T x\right\rangle = \frac{1}{2}\langle Ax,\, x\rangle + \frac{1}{2}\left\langle x,\, A^T x\right\rangle$ d'où

$$f(x) = \frac{1}{2}\left\langle \frac{1}{2}(A + A^T)x,\, x\right\rangle - \langle b,\, x\rangle + c,$$

où cette fois $\tilde{A} = \frac{1}{2}(A + A^T)$ est symétrique.

La proposition suivante résume les propriétés des fonctionnelles quadratiques.

Proposition 2.29 (Propriétés des fonctionnelles quadratiques)**.** Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonctionnelle quadratique, donc de la forme $f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c$ avec A symétrique. Alors,

a) f est $\mathcal{C}^2$ sur $\mathbb{R}^n$ (et même $\mathcal{C}^\infty$).

b) Pour tout $x \in \mathbb{R}^n$,

$$\nabla f(x) = Ax - b \quad \text{et} \quad \nabla^2 f(x) = A.$$

c) f est convexe si et seulement si A est positive.

d) f est strictement convexe si et seulement si A est définie positive.

e) $\inf\limits_{x\in\mathbb{R}^n} f(x)$ est fini si et seulement si A est positive et le système linéaire $Ax = b$ admet (au moins) une solution. Dans ce cas, l'ensemble des solutions de $Ax = b$ est l'ensemble des minimums globaux de f.

Ainsi résoudre le problème d'optimisation associé à f revient à résoudre le système linéaire $Ax = b$.

Démonstration. a) f est polynomiale en les coefficients de $x \in \mathbb{R}^n$ donc f est $\mathcal{C}^2$ sur $\mathbb{R}^n$ (et même $\mathcal{C}^\infty$).

b) On écrit d'abord la définition de la différentiabilité pour f en $x \in \mathbb{R}^n$, pour tout $h \in \mathbb{R}^n$ on a

$$\begin{aligned} f(x+h) &= f(x) + \frac{1}{2}\langle Ax,\, h\rangle + \frac{1}{2}\langle Ah,\, x\rangle - \langle b,\, h\rangle + \frac{1}{2}\langle Ah,\, h\rangle\,, \\ &= f(x) + \langle Ax - b,\, h\rangle + o(\|h\|_2), \end{aligned}$$

où on a utilisé que A est symétrique. Donc $\nabla f(x) = Ax - b$. Enfin, la différentielle de ∇f en x est la matrice A donc $\nabla^2 f(x) = A$.

c) Comme f est $\mathcal{C}^2$ sur $\mathbb{R}^n$, d'après le théorème 2.26, f est convexe sur $\mathbb{R}^n$ si et seulement si $\nabla^2 f(x) = A$ est positive pour tout $x \in \mathbb{R}^n$.

d) De nouveau grâce au théorème 2.26, si A est définie positive alors f est strictement convexe. Par contre nous avons vu qu'en général la réciproque est fausse (exemple $x \mapsto x^4$). Cependant pour les fonctionnelles quadratiques, elle est vraie. En effet d'après le théorème 2.25, f est strictement convexe est équivalent à pour tout $x, y \in \mathbb{R}^n$, $x \neq y$, $f(y) > f(x) + \langle \nabla f(x),\, y - x \rangle$, ce qui se réécrit après simplifications

$$\forall x, y \in \mathbb{R}^n, x \neq y, \quad \langle A(y-x),\, y-x \rangle > 0,$$

c'est-à-dire A définie positive (c'est la définition).

e) Commençons par le sens indirect. Le fait que A soit positive signifie par c) que f est convexe et le fait que $Ax = b$ admet une solution $x^* \in \mathbb{R}^n$ implique $\nabla f(x^*) = 0$ par b), c'est-à-dire x^* est un point critique de f. Par le théorème 2.27, x^* est un minimum global de f d'où $\inf_{\mathbb{R}^n} f = f(x^*) > -\infty$.

Pour le sens direct, montrons tout d'abord que A est nécessairement positive. Supposons que ce n'est pas le cas par l'absurde. Alors comme A est symétrique il existe (une valeur propre) $\lambda < 0$ et (un vecteur propre) $v \in \mathbb{R}^n \setminus \{0\}$ tel que $Av = \lambda v$. Pour tout $t \in \mathbb{R}$, on a alors

$$f(tv) = \frac{t^2}{2} \lambda \|v\|_2^2 - t \langle b,\, v \rangle + c,$$

et comme $\|v\|_2^2 \neq 0$, on en déduit $\lim\limits_{t \to +\infty} f(tv) = -\infty$, ce qui contredit que $\inf_{\mathbb{R}^n} f > -\infty$. D'où A positive. Montrons maintenant que $Ax = b$ admet au moins une solution. Raisonnons de nouveau par l'absurde. Cela signifie que $b \notin \operatorname{Im} A$. Comme A est symétrique, $\mathbb{R}^n$ se décompose suivant la somme directe orthogonale suivante

$$\mathbb{R}^n = \operatorname{Ker} A \overset{\perp}{\oplus} \operatorname{Im} A.$$

On a donc $b = b_1 + b_2$ avec $b_1 \in \operatorname{Ker} A$, $b_2 \in \operatorname{Im} A$ et $\langle b_1,\, b_2 \rangle = 0$. On a forcément $b_1 \neq 0$ car sinon $b = b_2 \in \operatorname{Im} A$, d'où pour tout $t \in \mathbb{R}$

$$f(tb_1) = \frac{1}{2} t^2 \left\langle \underbrace{Ab_1}_{=0},\, b_1 \right\rangle - t \langle b,\, b_1 \rangle + c = -t \|b_1\|_2^2 + c,$$

ce qui implique $\lim\limits_{t \to +\infty} f(tb_1) = -\infty$ car $\|b_1\|_2^2 \neq 0$. Ceci contredit $\inf_{\mathbb{R}^n} f > -\infty$. D'où $Ax = b$ admet au moins une solution.

On a donc démontré que $\inf_{\mathbb{R}^n} f > -\infty$ est équivalent à A symétrique et $Ax = b$ admet au moins une solution et les solutions de ce systèmes sont les minimiseurs globaux de f.

□

Bibliographie

[ROUVIÈRE] François ROUVIÈRE, *Petit guide du calcul différentiel à l'usage de la license et de l'agrégation*, Cassini, 2009

[ALLAIRE & KABER] Grégoire ALLAIRE et Sidi Mahmoud KABER, *Algèbre linéaire numérique*, Ellipses, 2002

[CIARLET] Philippe G. CIARLET, *Introduction à l'analyse numérique matricielle et à l'optimisation*, cinquième édition, Dunod, 1998

[BOYD & VANDENBERGHE] Stephen BOYD and Lieven VANDENBERGHE *Convex Optimization*, Cambridge University Press, 2004