ISE 1, ENSAE, Dakar

Optimisation et convexité

Chapitre 1 : Rappel sur la convexité et le calcul différentiel

Dr. Oumar Diop
oumar.diop@unchk.edu.sn

16 avril 2025

Introduction

Soit U un sous ensemble de $\mathbb{R}^n$ et $J : \mathbb{R}^n \to \mathbb{R}$ une fonction. On étudie, tout au long de ce cours le problème d'optimisation suivant :

$$\textbf{Chercher } x^* \in U \text{ tel que : } J(x^*) = \min_{x \in U} J(x). \tag{1}$$

U est un ensemble de contraintes, la fonction f est par exemple, le coût d'un produit.
Les ensembles et fonctions convexes jouent un rôle fondamental en optimisation. Il est important de connaître les définitions et propriétés de base des fonctions et ensembles convexes. Les modèles convexes sont plus faciles à résoudre que les modèles non convexes. Pour la classe des fonctions convexes, tous les points stationnaires (points qui annulent les dérivées de f) sont des minima. Un ensemble convexe contient le segment de droite reliant toute paire de ses points.
Cette séquence introductive présente des rappels sur la convexité et le calcul différentiel.

Rappel sur le calcul différentiel

On considère, tout au long de séquence, l'espace euclidien $\mathbb{R}^n$. On rappelle que $x \in \mathbb{R}^n$ si $x = (x_1, x_2, ..., x_n)$, avec $x_i \in \mathbb{R}\ \forall i = 1, 2, ..., n$. Les éléments $(e_i)_{i=1,2,...,n}$ de la base canonique sont donnés par $(e_i)_j = \delta_{i,j}$(symbole de Kronecker).

1. Pour $x, y \in \mathbb{R}^n$, on définit le **produit scalaire** de x et y par :

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i.$$

2. Pour $x \in \mathbb{R}^n$, on définit la **norme** de x par :

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

3. Pour tout $x_0 \in \mathbb{R}^n$ et $r > 0$, on note par : $B(x_0, r)$ la boule ouverte de centre x et rayon r, donnée par :

$$B(x_0, r) = \{x \in \mathbb{R}^n : \|x - x_0\| < r\}.$$

Rappel sur le calcul différentiel

1. On rappelle qu'un segment $[x, y]$ de $\mathbb{R}^n$ est donné par :

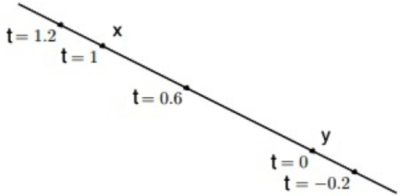
$$[x, y] = \{t\omega + (1 - t)y, \ \ t \in [0, 1]\}.$$

   Un élément ω du segment $[x, y]$ s'écrit sous la forme $\omega = y + t(x - y)$. On peut interpréter ω comme la somme d'un point initial y et d'une direction $x - y$ pondérée par le paramètre t, qui donne la fraction du chemin reliant y et x où ω se trouve, en fait, comme t varie de 0 à 1, ω varie de y à x.

   Pour obtenir les segments $[x, y[,]x, y[,]x, y]$, on remplace, respectivement le segment $[0, 1]$ par $[0, 1[,]0, 1[,]0, 1]$.

2. On appelle droite passant par deux éléments distincts x et y de $\mathbb{R}^n$ l'ensemble (infini) défini par :

$$d_{x,y} = \{tx + (1 - t)y, t \in \mathbb{R}\}.$$



Rappel sur le calcul différentiel

Soit $f : \Omega \to \mathbb{R}$ avec Ω un ouvert de $\mathbb{R}^n$. On rappelle les définitions suivantes :

1. Pour tout $x \in \Omega$, on note (s'il existe) le gradient de f par :

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \frac{\partial f}{\partial x_3}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right)^t$$

   $\frac{\partial f}{\partial x_i}(x)$ est la dérivée partielle de f en x de direction x_i.
   On note aussi le gradient de f par $J_f^t = \nabla f$, J_f est appelée la jacobienne de f.

2. On dit que f est de classe C^m si toutes ses dérivées si partielles jusqu'à l'ordre m existent et sont continues.

3. On note, quand elle existe, **la matrice hessienne** de f par :

$$\left(\nabla^2 f(x)\right)_{i,j} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right), \quad i, j = 1, 2, 3, \ldots, n.$$

D'après le théorème de Schwartz, si $f \in C^2(\Omega)$ alors la matrice hessienne est symétrique.

Exercice 1

1. *Vérifier que le gradient d'une fonction linéaire est constante,*
2. *Montrer que la hessienne d'une fonction quadratique est constante.*

Ensemble convexe

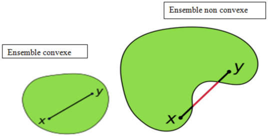
Définition 2

Un sous ensemble U de $\mathbb{R}^n$ est dit étoilé par rapport à un élément $x_0 \in U$ si pour tout $x \in U$, le segment $[x_0, x]$ reste dans U.

Définition 3

Un sous ensemble U de $\mathbb{R}^n$ est dit convexe si pour tout couple de points (x,y) de U, le segment $[x, y]$ reste dans dans U. Autrement dit : pour tout couple de points (x,y) de U, et pour tout $\lambda \in [0, 1]$, on a :

$$\lambda x + (1 - \lambda) y \in U.$$



Interprétation : Dans un salon convexe, deux personnes peuvent toujours s'apercevoir. C'est alors une pièce sans recoin.
Rque : Un ensemble convexe est un ensemble étoilé par rapport à tous ses points.

Définition 4

- *L'épigraphe d'une fonction $f : U \to \mathbb{R}$ est l'ensemble des couples de points $(x, z) \in U \times \mathbb{R}$ vérifiant :*

$$z \geq f(x),$$

ce qui revient à dire intuitivement que son graphe est le bord d'un convexe.

- *Une fonction $f : U \to \mathbb{R}$ est dite convexe si pour tout couple de points (x, y) et pour tout $\lambda \in [0, 1]$ on a :*

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

- *La fonction f est dite strictement convexe si l'inégalité est stricte pour tout couple de points (x, y), avec $x \neq y$.*
- *La fonction f est dite fortement convexe de module α si :*

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

Fonction convexe

- La fonction f est concave (strictement concave) si $-f$ est convexe (strictement convexe).
- Une fonction est (strictement) convexe si et seulement si sa restriction à tout segment inclus dans U est (strictement) convexe : pour établir la convexité on peut donc toujours se ramener à une fonction d'une variable.
- Si f est convexe l'ensemble $\{x/f(x) \leq c\}$ est convexe (la réciproque est fausse).
- La fonction f est convexe si son épigraphe est un ensemble convexe.

Fonction convexe

Exemple 5

- *Une forme linéaire est convexe.*
- *Une norme définit une fonction convexe.*
- *Une fonction quadratique positive est strictement convexe, car sa restriction à une droite quelconque est une fonction trinôme du second degré à terme directeur positif, donc strictement convexe, plus précisément nous utiliserons souvent la proposition suivante.*
- *Les fonctions affines $f : \mathbb{R}^n \to \mathbb{R}$ données par :*

$$f(x) = a^t x + b, \quad a, b \in \mathbb{R}^n,$$

sont toutes et seules les fonctions qui sont convexes et concaves (non strictement).

Convexité & Gradient

Soit Ω un ouvert de $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ une fonction et U un sous ensemble convexe de Ω. on donne les propriétés suivantes.

Proposition 6

Soit f une fonction de classe C^1. Alors on a les propriétés suivantes.

1. *f est convexe sur U si et seulement si :*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \forall\, x, y \in U$$

2. *f est strictement convexe sur U si et seulement si :*

$$f(y) > f(x) + \langle \nabla f(x), y - x \rangle, \forall\, x, y \in U \;\; \textit{avec}\; x \neq y$$

3. *f est convexe sur U si et seulement si ∇f est monotone sur U, c'est à dire*

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0 \;\forall\, x, y \in U.$$

4. *Si ∇f est strictement monotone (c'est à dire*

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle > 0 \;\forall\, x, y \in U, \;\textit{avec}\; x \neq y)$$

alors f est strictement convexe

Convexité et Matrice Hessienne

Soit Ω un ouvert de $\mathbb{R}^n$, U un ensemble convexe et $f : \Omega \to \mathbb{R}$ une fonction de classe C^2, alors nous avons les propriétés suivantes :

Proposition 7

1. *f est convexe sur U si et seulement si :*

$$\langle \nabla^2 f(x)(y-x), y-x \rangle \geq 0 \; \forall \, x, y \in U,$$

2. *Si*

$$\langle \nabla^2 f(x)(y-x), y-x \rangle > 0 \; \forall \, x, y \in U, \text{ avec } x \neq y,$$

   alors f est strictement convexe.

Remarque 8 (Cas particuliers)

1. *Dans le cas particulier où $\Omega = U = \mathbb{R}^n$ alors les deux inégalités de la proposition précédente peuvent s'écrire sous la forme suivante :*

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

   et respectivement

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

2. *Pour $n = 1$, alors Ω est un intervalle, les deux inégalités de la proposition précédente deviennent alors :*

$$f''(x) \geq 0, \quad \forall x \in \Omega$$

   respectivement

$$f''(x) > 0, \quad \forall x \in \Omega$$

Optimisation

ANSD
Agence Nationale de
la Statistique et de la Démographie

Exercice 9

Pour les fonctions suivantes calculer leur matrices Hessiennes, les valeurs propres de ces matrices et détérminer si ces fonctions sont convexes sur leur domaine de définition

$$\begin{array}{llcc} a. & f_1(x,y) & = & x^2+y^3, \\ b. & f_2(x,y) & = & x^2+y^2+2xy \\ c. & f_3(x,y) & = & e^{2xy} \end{array} \tag{2}$$

Exercice 10

Parmi les fonctions suivantes, lesquelles sont convexe sur leur ensemble de définition

$$\begin{array}{llcc} a. & f_1(x) & = & \ln(1+x) \\ b. & f_2(x) & = & x^2e^{2x} \\ c. & f_3(x) & = & 2x\cos(x) \end{array} \tag{3}$$

Fonctions elliptiques

Définition 11

Soit $f : \mathbb{R}^n \to \mathbb{R}$. On dit que f est une fonction elliptique si elle est de classe C^1 et s'il existe une constante positive γ telle que :

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \gamma \|x - y\|^2, \quad \forall x, y \in \mathbb{R}^n.$$

Proposition 12 (Ellipticité et matrice hessienne)

Soit f une fonction de classe C^2, alors f est elliptique si et seulement si il existe une constante $\beta > 0$ telle que :

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

Exemple 13

Les fonctions suivantes sont elliptiques

$$\begin{array}{llcc} a. & f_1(x) & = & 2x^2 - 5x + 17 \\ b. & f_2(x) & = & x^2 + \cos(x) \\ c. & f_3(x) & = & ax^2 + bx + c, \quad a > 0. \end{array} \tag{4}$$

Forme quadratique

Soit la forme quadratique $f : \mathbb{R}^n \to \mathbb{R}$ donnée par

$$f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c, \forall x \in \mathbb{R}^n.$$

avec A une matrice carrée réeelle de taille n, avec b un vecteur de $\mathbb{R}^n$ et c un scalaire, on appelle encore, par abus de langage, fonction (ou forme) quadratique une fonction de ce type). Ici on a :

$$\nabla f(x) = Ax - b,$$

et

$$\nabla^2 f(x) = A.$$

1. Si A est une matrice symmétrique alors on sait que toutes les valeurs propres de A sont réelles et que

   $$\langle Ah, h\rangle \geq \lambda_{min} \|h\|^2 \quad \forall h \in \mathbb{R}^n$$

   λ_{min} est la plus petite valeur propre de A.
2. Si, de plus, A est définie positive alors on en déduit que la fonction f est elliptique.

On rappelle qu'une matrice A est définie positive si :

$$\langle Ax, x\rangle > 0 \;\; \forall\, x \in \mathbb{R}^n, x \neq 0.$$

Fonction coercive

Définition 14

Soit Ω un ensemble non borné de $\mathbb{R}^n$ et une application $f : \Omega \to \mathbb{R}$. On dit que f est coercive sur Ω si on a :

$$\lim_{x\in\Omega \ \|x\|\to\infty} f(x) = +\infty$$

Rq : En pratique, on choisit la norme la plus adaptée à la fonction f étudiée. En effet, sur $\mathbb{R}^n$, toutes les normes sont équivalentes.
Exemple : La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ féfinie par $f(x, y) = x^2 + y^2 - 2xy$ est une fonction coercive sur $\mathbb{R}^2$.

Proposition 15

Soit $f : \Omega \to \mathbb{R}$ une fonction elliptique, alors elle est strictement convexe et coercive. De plus, elle vérifie l'inégalité suivante :

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|y - x\|^2.$$

Exercice 16

Considérons les application de $\mathbb{R}^3$ *dans* $\mathbb{R}$ *par :*

$$f(x,y,z) = x^2 + y^2 + z^2$$

et

$$\psi(x,y,z) = x^2 + y^2 + z^2 + 2$$

et sous ensemble de $\mathbb{R}^3$ *donné par :*

$$D = \{(x,y,z) \in \mathbb{R}^3 : \psi(x,y,z) \leq 0\}$$

1. *Montrer que la fonction* ψ *est convexe.*
2. *Soit* $X_1 = (x_1, y_1, z_1) \in \mathbb{R}^3$ *et* $X_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$ *et* $\theta \in (0,1)$*. Quelle inégalité doit-on vérifier pour prouver que* D *est convexe ?*
3. *En déduire de la question 1) que que* D *est un sous ensemble convexe.*
4. *Montrer que* f *est strictement convexe sur D.*

Exercice 17

Soit une fonction $f : \mathbb{R} \to \mathbb{R}$, répondre par Vrai ou Faux

1. *Si f n'est pas convexe, alors elle est concave.*
2. *Soit f de classe C^2 sur un intervalle I et $a \in I$. Si f'' s'annule en a alors f change de convexité en a.*
3. *Si f est convexe sur $\mathbb{R}$ alors elle est dérivable sur $\mathbb{R}$.*
4. *Soit f de classe C^2 sur un intervalle I et $a \in I$. Si f change de convexité en a, alors f'' s'annule en a.*
5. *Soit f deux fois dérivable sur un intervalle I et $a \in I$. Si f change de convexité en a, alors f'' s'annule en a.*

Exercice 18

1. *Montrer que, pour tout* $x \in [0, \pi/2]$ *, on a*

$$\frac{2}{\pi}x \leq \sin(x) \leq x.$$

2. *Montrer que* $f(x) = \ln(\ln(x))$ *est concave sur* $]1, +\infty[$.
3. *En déduire que* $\forall a, b > 1, \ln\left(\frac{a+b}{2}\right) \geq \sqrt{\ln a. \ln b}$.
4. *En utilisant la propriété des fonctions convexes :*

$$f(x) \geq f(a) + f'(a)(x-a), \forall x, a \in \mathbb{R},$$

   montrer que

$$e^x \geq 1 + x$$

   pour tout réel x *(on peut prendre* $a = 0$*).*

Exercice 19

On considère les matrices carrées suivantes.

$$A_1 = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}, A_2 = \begin{pmatrix} 2 & -2 \\ 5 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix}$$

1. *Trouver les valeurs propres de ses matrices.*
2. *Pour chaque matrice écrire une forme bilineaire correspondante.*
3. *Lesquelles parmis ses formes bilineaires correspondent aux formes quadratiques ?*
4. *Quelles matrices peuvent être des matrices Hessiennes ?*

Exercices

Exercice 20

Soit f la fonction de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ définie par :

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & si\ (x,y) \neq (0,0) \\ 0 & si\ (x,y) = (0,0) \end{cases} \tag{5}$$

1. *La fonction f est-elle $C^1(\mathbb{R}^2, \mathbb{R})$?*
2. *La fonction f est-elle coercive ?*

Exercice 21

Soit f une fonction de $\mathbb{R}^3 \backslash \{(0,0,0)^t\}$ à valeurs dans $\mathbb{R}$ définie par :

$$f(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$$

1. *Déterminer le gradient de la fonction et la Hessienne de f.*
2. *La Hessienne de f est-elle définie positive ?*

Exercices

Exercice 22

Soit la fonction f définie par :

$$f(x,y) = \frac{1}{4}x^4 - xy + y^2.$$

1. *La fonction f est-elle convexe,*
2. *Montrer que f est coercive,*
3. *Déterminer le gradient de f,*
4. *Calculer les points critiques de f,*
5. *Déterminer la Hessienne de f.*

Exercice 23

Soit $A \in \mathbb{M}_{n,p}(\mathbb{R})$ et $b \in \mathbb{R}^p$. et f la fonction définie de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$ par : par $f(x) = \|Ax - b\|^2$.

1. *Calculer le gradient et la hessienne de f*
2. *f est-elle coercive.*