Chapitre 1

Quelques rappels de calcul differentiel, convexité

1.1 Différentiabilité

1.1.1 Gradient

Définition 1.1 *Soit* $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ *une fonction continue. La* $i^{ème}$ *dérivée partielle de* f *au point* h *est fonction notée* $\dfrac{\partial f}{\partial x_j}$ *et est donneé par :*

$$\frac{\partial f}{\partial x_j} = \lim_{h \to 0} \frac{f(x_1, ...x_j + h, ..., x_n) - f(x_1, ...x_j, ..., x_n)}{h}.$$

Définition 1.2 *On appelle gradient d'une fonction* $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ *au point* x*, le vecteur qu'on note* $\operatorname{grad}(f)$ *ou* ∇f *donné par :*

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n} \right)^T (x).$$

Exemple 1.1 *Soit la fonction définie de* $\mathbb{R}^2$ *dans* $\mathbb{R}$ *par :*

$$f(x_1, x_2, x_3) = 2x_1^2 - x_2 x_3 + x_1 x_3 + x_3^3.$$

On a :

$$\begin{aligned} \nabla f(x_1, x_2, x_3) &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right)^T (x_1, x_2, x_3) \\ &= \left(4x + x_3, -x_3, x_1 - x_2 + 3x_3^2 \right). \end{aligned}$$

si $x_1 = 1, x_2 = 1$ *et* $x_3 = 1$*, on a :*

$$\nabla f(1, 1, 1) = (5, -1, 3).$$

1.1.2 Matrice Hessiene

Définition 1.3 *On appelle matrice Hessiene de f en x, la matrice carrée donnée :*

$$
\begin{aligned}
H_f(x) &= \nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{1 \leq i,j \leq n} \\
&= \nabla\left(\nabla f(x)\right) = \nabla\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right)^T (x) \\
&= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.
\end{aligned}
$$

Remarque 1.1 *Si $f \in C^2(K)$ alors $H_f(x)$ est une matrice symmétrique $\forall x \in K$.*

Exemple 1.2 *Soit la fonction définie de $\mathbb{R}^2$ dans $\mathbb{R}$ par :*

$$f(x_1, x_2, x_3) = 2x_1^2 - x_2 x_3 + x_1 x_3 + x_3^3.$$

La matrice hessienne de la fonction f est donné par :

$$
\begin{aligned}
H_f(x) &= \nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{1 \leq i,j \leq n} \\
&= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{pmatrix} \\
&= \begin{pmatrix} 4 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & -1 & 6x_3 \end{pmatrix}.
\end{aligned}
$$

Remarque 1.2 *Si $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction constante, alors $\nabla f = H_f = 0$.*

1.1.3 Dérivée directionnelle

Définition 1.4 *On appelle dérivée directionnelle de f en x dans la direction d , notée $Df(x, d)$ ou $\frac{\partial f}{\partial d}$, la limite du rapport :*

$$\frac{f(x + td) - f(x)}{t} \text{ lorsque } t \text{ tendvers } 0.$$

Autrement dit :

$$Df(x, d) = \lim_{t \to 0} \frac{f(x + td) - f(x)}{t} = \nabla^T f(x) .d.$$

Exemple 1.3 *Soit* $f(x_1, x_2, x_3) = 2x_1^2 - x_2x_3 + x_1x_3 + x_3^3$ *et* $d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$.

En effet,

$$Df(x, d) = \nabla^T f(x) .d.$$

Le gradient de f *est donné par :*

$$\begin{aligned} \nabla f(x_1, x_2, x_3) &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}\right)^T (x_1, x_2, x_3) \\ &= \left(4x + x_3, -x_3, x_1 - x_2 + 3x_3^2\right). \end{aligned}$$

Alors

$$\begin{aligned} Df(x, d) &= \nabla^T f(x) .d = \left(4x + x_3, -x_3, x_1 - x_2 + 3x_3^2\right) \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \\ &= d_1(4x + x_3) - d_2x_3 + d_3\left(3x_3^2 - x_2\right). \end{aligned}$$

La dérivée directionnelle de f *dans la direction* d *est :* $d_1(4x + x_3) - d_2x_3 + d_3(3x_3^2 - x_2)$.

Définition 1.5 *On dit que* x^* *est un poit stationnaire de* f *si* $\nabla f(x^*) = 0$.

Exemple 1.4 *Soit* f *une fonction de* $\mathbb{R}^2$ *dans* $\mathbb{R}$ *définie par :*

$$f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2.$$

On calcul d.abord les dérivées partielles :

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x_1, x_2) &= 3x_1^2 - 3x_2, \\ & \quad et \\ \frac{\partial f}{\partial x_2}(x_1, x_2) &= 3x_2^2 - 3x_1. \end{aligned}$$

Alors

$$\nabla f(x_1, x_2) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2},\right)^T (x_1, x_2) = \begin{pmatrix} 3x_1^2 - 3x_2 \\ \\ 3x_2^2 - 3x_1 \end{pmatrix},$$

$$\begin{aligned} \nabla f(x_1, x_2) &= 0 \Rightarrow \begin{pmatrix} 3x_1^2 - 3x_2 \\ \\ 3x_2^2 - 3x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ \\ 0 \end{pmatrix} \\ &\Rightarrow \begin{cases} 3x_1^2 - 3x_2 = 0 \\ \\ 3x_2^2 - 3x_1 = 0 \end{cases} \Rightarrow \begin{cases} x_2 = x_1^2 \\ \\ x_1^4 - x_1 = 0 \end{cases} \Rightarrow \begin{cases} x_2 = x_1^2 \\ \\ x_1(x_1^3 - 1) = 0 \end{cases} \\ &\Rightarrow \begin{cases} x_2 = x_1^2 \\ \\ x_1 = 0 \ ou \ x_1 = 1 \end{cases}. \end{aligned}$$

On obtient alors deux points stationnaires $(0, 0)$ *et* (11).

Définition 1.6 *Soir $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue, si pour tout $d \in \mathbb{R}^n$ la dérivée directionnelle de f dans la direction d existe, alors la fonction f est dite différentiable.*

1.1.4 Direction de descente

Définition 1.7 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable et $x, d \in \mathbb{R}^n$. La direction d est une descente en x si*

$$\nabla f(x)^T d < 0.$$

Remarque 1.3 *Si $d = -\nabla f(x)$, $\nabla f(x) \neq 0$, alors d est une direction de descente.*

$$\nabla f(x)^T .d = \nabla f(x)^T .(-\nabla f(x)) = -\|\nabla f\|^2 < 0$$

Exemple 1.5 *Soit $f(x_1, x_2) = x_1^2 + x_2^2$, $x^* = (1, 1)$.*

Choisir une direction $d = (d_1, d_2)$ telle que soit d une direction de descente au point x^*.

On a :

$$\nabla f(x_1, x_2) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_1, x_2) \\ \\ \frac{\partial f}{\partial x_2}(x_1, x_2) \end{pmatrix} = \begin{pmatrix} 2x_1 \\ \\ 2x_2 \end{pmatrix}.$$

Alors

$$\nabla f(x_1, x_2)^T (-\nabla f(x_1, x_2)) = (2x_1, 2x_2) \begin{pmatrix} -2x_1 \\ \\ -2x_2 \end{pmatrix} = -4\left(x_1^2 + x_2^2\right).$$

Si $(x_1, x_2) = (1, 1)$, on a

$$\nabla f(x^*)^T d = \nabla f(1, 1)^T d = (2, 2) \begin{pmatrix} d_1 \\ \\ d_2 \end{pmatrix} = 2d_1 + 2d_2 < 0,$$

ce qui implique

$$d_1 + d_2 < 0.$$

On a une infinité de solutions.

Si $d_1 = 1$, on peut prend $d_2 = -1$.

On donne maintenant une condition suffisante pour que d soit une direction de descente.

Théorème 1.1 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable au point $x^* \in \mathbb{R}^n$. et $d \in \mathbb{R}^n$ une direction vérifiant la condition suivante :*

$$\nabla f(x^*)^T d < 0.$$

Alors d est une direction de descente au point x^*.

Preuve. Exercice. ■

1.2 Développement de Taylor

Soit $C \subset \mathbb{R}^n, f : C \to \mathbb{R}, x \in C$ et $h \in \mathbb{R}^n$ tel que $[x, x+h] \subset C$. Alors
(i) : Si $f \in C^1(C)$, on a :
(i_1) : Formule de Taylor à l'ordre 1 avec rest intégrale

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

(i_2) : Formule de Taylor-Maclaurin à l'ordre 1

$$f(x+h) = f(x) + \nabla f(x+h)^T h.$$

(i_3) : Formule de Taylor-young à l'ordre 1

$$f(x+h) = f(x) + \nabla f(x)^T h + o(\|h\|).$$

(ii) : Si $f \in C^2(C)$, on a :
(ii_1) : Formule de Taylor à l'ordre 2 avec rest intégrale

$$f(x+h) = f(x) + \nabla f(x+h)^T h + \int_0^1 (1-t) \, h^T \nabla^2 f(x+th)^T \, hdt.$$

(ii_2) : Formule de Taylor-Maclaurin à l'ordre 2

$$f(x+h) = f(x) + \nabla f(x+h)^T h + \frac{1}{2} h^T \nabla^2 f(x+\theta h) \, h, \; 0 < \theta < 1.$$

(i_3) : Formule de Taylor-young à l'ordre 2

$$f(x+h) = f(x) + \nabla f(x + \triangle h)^T h + \frac{1}{2} h^T \nabla^2 f(x) \, h + o\left(\|h\|^2\right).$$

1.2.1 Forme quadratique

Définition 1.8 *Soit $A \in M_{n,m}(\mathbb{R})$ une matrice symétrique et $b \in \mathbb{R}^n$. On appelle forme quadratique la fonction $f : \mathbb{R}^n \to \mathbb{R}$ définie par :*

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

Exemple 1.6 *Soit f une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par :*

$$f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2 - 6x_1 - 6x_2.$$

On a :

$$\begin{aligned} f(x_1, x_2) &= (x_1 x_2) \begin{pmatrix} 8 & -2 \\ -2 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - (6, 6) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \frac{1}{2} x^T A x - b^T x, \end{aligned}$$

où

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, A = \begin{pmatrix} 8 & -2 \\ -2 & 8 \end{pmatrix} \text{ et } b = (6, 6).$$

Définition 1.9 *Soit* $A \in M_{n,m}(\mathbb{R})$ *une matrice carrée. On dit que* A *est semi-définie positive* $(S.D.P)$ *et on note* $A \geq 0$ *ssi*

$$x^T Ax \geq 0, \quad \forall x \in \mathbb{R}^n.$$

Définition 1.10 *Soit* $A \in M_{n,m}(\mathbb{R})$ *une matrice carrée. On dit que* A *est définie positive* $(D.P)$ *et on note* $A > 0$ *ssi*

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

Remarque 1.4 *Soit* A *une matrice symétrique* $\left(A = A^T\right)$. *Alors :*

(1) : *La matrice* A *est semi-définie positive* $(S.D.P)$ *ssi tous les mineurs principaux sont positifs ou nuls.*

(2) : *La matrice* A *est définie positive* $(D.P)$ *ssi tous les mineurs principaux sont strictement positifs.*

Exemple 1.7 *Soit* $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

On calcule les mineurs principaux $\triangle_i, i = 1, 2, 3.$

$$\triangle_1 = 1, \triangle_2 = \triangle_3 = 0.$$

La matrice A *est semi-définie positive.*

Attention : A n'est pas nécessairement $S.D.P$ même si tous les mineurs principaux sont positifs.

Contre exemple : Soit $A = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 1 \end{pmatrix}$. On a :

$$\triangle_1 = 1, \triangle_2 = \triangle_3 = 0,$$

mais si $x = (0, -1, 2)$, alors $(0, -1, 2) \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} = -2.$

1.3 Fonctions convexes

1.3.1 Ensemble convexe

Définition 1.11 *On dit qu'un ensemble* C, $C \in \mathbb{R}^n$ *est convexe si :*

$$\forall x_1, x_2 \in C, \forall \alpha \in [0, 1], \; \alpha x_1 + (1 - \alpha) x_2 \in C.$$

Exemple 1.8 *Soit* $H = \{x \in \mathbb{R}^n, \lambda x = b\}$ *un hyperplan.*

Montrer que H *est convexe.*

Preuve. Soient $x_1, x_2 \in C$ et $\alpha \in [0,1]$, on montre que

$$\alpha x_1 + (1-\alpha) x_2 = z \in H.$$

i. e $\lambda z = b$, on a :

$$\begin{aligned} \lambda z &= \lambda (\alpha x_1 + (1-\alpha) x_2) = \alpha \lambda x_1 + (1-\alpha) \lambda x_2 \\ &= \alpha b + (1-\alpha) b = b, \end{aligned}$$

donc $\lambda z = b$.
D'où H est convexe. ■

Proposition 1.1 *Un ensemble C, $C \in \mathbb{R}^n$ est convexe ssi pour tout $m \in \mathbb{R}^*$, $x_1, x_2, ..., x_n \in C$ et $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}_+$, tq $\sum_{i=1}^n \alpha_i - 1$, on a :*

$$\sum_{i=1}^n \alpha_i x_i = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n \in C.$$

Un ensemble C est convexe ssi toute combinaison linéaire convex de points de C appartient

Preuve. Exercice. ■

Proposition 1.2 *Si C_1, C_2 sont deux ensembles convexes de $\mathbb{R}^n$ et α_1, α_2 deux réels, alors $\alpha_1 C_1 + \alpha_2 C_2$ est un convexe de $\mathbb{R}^n$.*

Proposition 1.3 *Si $(C_i), i = 1, n$ est une famille quelconque de convexes de $\mathbb{R}^n$, alors $\bigcap\limits_{i=1}^n C_i$ est un convexe de $\mathbb{R}^n$.*

Preuve. 1. Soient $x_1, x_2 \in \bigcap\limits_{i=1}^n C_i, \alpha \in [0,1]$, $\alpha x_1 + (1-\alpha) x_2 \in \bigcap\limits_{i=1}^n C_i$? ■

$x_1 \in \bigcap\limits_{i=1}^n S_i \Rightarrow x_1 \in C_i, \forall i = 1, .., n$, et $x_2 \in \bigcap\limits_{i=1}^n C_i \Rightarrow x_2 \in S_i, \forall i = 1, .., n$,
comme S_i est convexe $\forall i = 1, ..., n$, alors $\alpha x_1 + (1-\alpha) x_2 \in S_i, \forall i = 1, .., n$,
donc $\forall i = 1, .., n$, $\lambda x_1 + (1-\lambda) x_2 \in \bigcap\limits_{i=1}^n C_i$.

1.3.2 Fonction convexe

Définition 1.12 *Soit C un ensemble convexe de $\mathbb{R}^n$ et $f : C \to \mathbb{R}^n$ une fonction. On dit que la fonction f est convexe ssi :*

$$\forall\ x_1, x_2 \in C,\ \forall \alpha \in [0,1] :\ f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2).$$

On dit que la fonction f est strictement convexe ssi :

$$\forall\ x_1, x_2 \in C, x_1 \neq x_2, \forall \alpha \in [0,1] : f(\alpha x_1 + (1-\alpha) x_2) < \alpha f(x_1) + (1-\alpha) f(x_2).$$

Proposition 1.4 *Une fonction $f : C \to \mathbb{R}^n$ convexe peut être caractérisé par : $\forall x_i \in C$ $i = 1, .., n$,et $\forall \alpha_i \in \mathbb{R}_+,$ tq $\sum_{i=1}^n \alpha_i = 1,$ on a :*

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i).$$

(f est convexe$\Leftrightarrow$ $\forall x_i \in C$ $i = 1, .., n, \forall \alpha_i \in \mathbb{R}_+,$ tq $\sum_{i=1}^n \alpha_i = 1,$ on a :$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i)$.)

Preuve. Exercice. ■

Proposition 1.5 *Toute combinaison linéaire sur C convexe de fonctions convexes à coefficients positifs est une convexe.*

Preuve. On pose : $f(x) = \sum_{i=1}^n a_i f_i(x)$ avec $f_i(x)$ convexe et $a_i \geq 0$ $\forall i = 1, .., n$. Alors

$$\forall\ x_1, x_2 \in C,\ \forall \alpha \in [0,1], f(\alpha x_1 + (1-\alpha) x_2) = \sum_{i=1}^n a_i f_i(\alpha x_1 + (1-\alpha) x_2), \forall i = 1, .., n.$$

Puisque $f_i(x), i = 1, .., n$ convexe sur C, on a :

$$f_i(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f_i(x) + (1-\alpha) f_i(x_2), i = 1, .., n.$$

Comme $a_i \geq 0$, alors

$$\forall i = 1, .., n\ a_i f_i(\alpha x_1 + (1-\alpha) x_2) \leq a_i f_i(x) + a_i (1-\alpha) f_i(x_2),$$

ce qui donne

$$\sum_{i=1}^n a_i f_i(\alpha x_1 + (1-\alpha) x_2) \leq \sum_{i=1}^n a_i f_i(x) + \sum_{i=1}^n a_i (1-\alpha) f_i(x_2),$$

ce qui implique

$$f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2).$$

D'où $\sum_{i=1}^n a_i f_i(x)$ est convexe. ■

Proposition 1.6 *Une application linéaire sur C convexe est toujours convexe sur C.*

Preuve. Soit $f : C \subset \mathbb{R}^n \to \mathbb{R}$ une application linéaire, alors

$$\begin{aligned}
f(x) &= f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = f\left(x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + ... + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}\right) \\
&= x_1 f\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 f\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + ... + x_n f\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = x_1 c_1 + x_2 c_2 + ... + x_n c_n \\
&= (c_1, c_2, ..., c_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = c.x.
\end{aligned}$$

Alors

$$\begin{aligned} \forall\ x_1, x_2 &\in C,\ \forall \alpha \in [0,1]\,, f\left(\alpha x_1 + (1-\alpha)\, x_2\right) = c\left(\alpha x_1 + (1-\alpha)\, x_2\right) \\ &= \alpha c x_1 + (1-\alpha)\, c x_2 = \alpha f\left(x_1\right) + (1-\alpha)\, f\left(x_2\right). \end{aligned}$$

■

Théorème 1.2 *Soit f une fonction convexe sur C convexe et A une matrice semi-définie positive $(S.D.P)$, alors la forme qudratique :*

$$f\left(x\right) = x^T A x,$$

est convexe sur C.

Preuve. Soiet $x_1, x_2 \in C$, et $\alpha \in [0,1]\,$, alors

$$\begin{aligned} & f\left(\alpha x_1 + (1-\alpha)\, x_2\right) \\ = & \left(\alpha x_1 + (1-\alpha)\, x_2\right)^T A \left(\alpha x_1 + (1-\alpha)\, x_2\right) \\ = & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 - \alpha\,(1-\alpha)\,(x_1 - x_2)^T A\,(x_1 - x_2) \\ = & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 - \alpha\,(1-\alpha)\left(x_1^T - x_2^T\right) A\,(x_1 - x_2) \\ = & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 - \alpha\,(1-\alpha)\left(x_1^T - x_2^T\right) A x_1 + \alpha\,(1-\alpha)\left(x_1^T - x_2^T\right) A x_2 \\ = & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 - \alpha\,(1-\alpha)\, x_1^T A x_1 + \alpha\,(1-\alpha)\, x_2^T A x_1 \\ & + \alpha\,(1-\alpha)\, x_1^T A x_2 - \alpha\,(1-\alpha)\, x_2^T A x_2 \\ = & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 - \alpha x_1^T A x_1 + \alpha^2 x_1^T A x_1 - \alpha x_2^T A x_2 + \alpha^2 x_2^T A x_2 \\ & + \alpha\,(1-\alpha)\, x_2^T A x_1 + \alpha\,(1-\alpha)\, x_1^T A x_2, \end{aligned}$$

ce qui implique

$$\begin{aligned} & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 - \alpha\,(1-\alpha)\,(x_1 - x_2)^T A\,(x_1 - x_2) \\ \leq\ & \alpha x_1^T A x_1 + (1-\alpha)\, x_2^T A x_2 = \alpha f\left(x_1\right) + (1-\alpha)\, f\left(x_2\right). \end{aligned}$$

■

Théorème 1.3 *Soit f une fonction définie sur $C \subset \mathbb{R}^n$ convexe et A une matrice semi-définie positive $(S.D.P)$, alors la forme qudratique :*

$$f\left(x\right) = x^T A x + b x,$$

est convexe sur C.

Preuve. Puisque A est $(S.D.P)$, alors $x^T A x$ est convexe sur C d'aprés théorème 1.1 et bx est convexe sur C d'aprés Proposition 1.6.

Donc $f\left(x\right) = x^T A x + bx$ $f\left(x\right) = x^T A x + bx$ car c'est une somme de fonctions convexe (d'aprés Proposition 1.5). ■

Théorème 1.4 *Soit $f : C \subset \mathbb{R}^n \to \mathbb{R}$ une fonction une fois différentiable, alors f est convexe ssi*

$$f(x_1) \geq f(x_2) + \nabla f(x_2)(x_1 - x_2), \forall x_1, x_2 \in C.$$

La fonction f est convexe ssi

$$f(x_1) > f(x_2) + \nabla f(x_2)(x_1 - x_2), \forall x_1, x_2 \in C, x_1 \neq x_2.$$

Preuve. Exercice. ■

Théorème 1.5 *Soit $f : C \subset \mathbb{R}^n \to \mathbb{R}$ une fonction deux fois différentiable, alors f est convexe ssi*

$$(x_1 - x_2)^T \nabla^2 f(x_2)(x_1 - x_2) \geq 0, \forall x_1, x_2 \in C.$$

La fonction f est convexe ssi

$$(x_1 - x_2)^T \nabla^2 f(x_2)(x_1 - x_2) > 0, \forall x_1, x_2 \in C, x_1 \neq x_2.$$

Preuve. Exercice. ■

1.4 Exercices corrigés

Exercice 1.1 *On définie les fonctions $f_i : \mathbb{R}^2 \to \mathbb{R}, i = 1, 2$ par :*

$$\begin{aligned} 1. f_1(x,y) &= (2 - x - y)^2 + (1 + x + y + xy)^2, \\ 2. f_2(x,y) &= \left(x^2 + y^2\right)^2 - x - y. \end{aligned}$$

Déterminer les point critiques des fonctions $f_i, i - 1, 2$.

——————————————————————————

Solution : 1 : $f_1(x,y) = (2 - x - y)^2 + (1 + x + y + xy)^2$.

$$\nabla f_1(x,y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -2(2 - x - y) + 2(1 + x + y + xy)(1 + y) = 0, \dots (1) \\ \\ -2(2 - x - y) + 2(1 + x + y + xy)(1 + x) = 0, \dots (2) \end{cases}$$

$(1) - (2) \Leftrightarrow (1 + y + x + xy)(y - x) = 0$

$$\Leftrightarrow \begin{cases} x = y \\ \text{où} \\ x = -1 \text{ où } y = -1 \end{cases}, \begin{cases} si\ x = -1 \Rightarrow y = 3 \\ si\ y = -1 \Rightarrow x = 3 \\ si\ x = x \Rightarrow x \text{ est sol de } 2x^3 + 6x^2 + 10x - 2 = 0 \end{cases}$$

Donc les point critiques de la fonction f_1 sont : $(-1, 1)$ et $(3, 3)$.

2. : $f_2(x,y) = (x^2 + y^2)^2 - x - y$.

$$\nabla f_2(x,y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 4x(x^2 + y^2)^2 - 1 = 0, \dots\dots\dots (1) \\ \\ 4y(x^2 + y^2)^2 - 1 = 0, \dots\dots\dots (2) \end{cases} (\bullet),$$

$(1)-(2) \Leftrightarrow (x=y)$ où bien $(x^2+y=0)$

cas1 : $x=y \Rightarrow 8x^3=1 \Rightarrow x=y=\frac{1}{2} \Rightarrow \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ est un point critique.

cas1 : $x^2+y=0 \Rightarrow x=y=0 \Rightarrow (\bullet)$ n'est pas verifier $\Rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ n'est pas un point critique.

Exercice 1.2 *Montrer que les ensemle suivants :*

$$
\begin{aligned}
1.\ S_1 &= \left\{(x,y) \in \mathbb{R}^2 /\ y-x^2 \geq 0\right\}, \\
2.\ S_2 &= \left\{(x,y) \in \mathbb{R}^2 / x \geq 0, y \geq 0, x+y \leq 1\right\},
\end{aligned}
$$

sont convexes.

———————————————————————

Solution : 1 : *On montre que* $S_1 = \{(x,y) \in \mathbb{R}^2 /\ y-x^2 \geq 0\}$ *est convexe.*

Soeint $z_1, z_2 \in S_1$ *et* $\lambda \in [0,1]$, *alors*

$$z_1 \in S_1 \Rightarrow \exists (x_1,y_1) \in \mathbb{R}^2, z_1=(x_1,y_1) : y_1 - x_1^2 \geq 0,$$

et

$$z_2 \in S_1 \Rightarrow \exists (x_2,y_2) \in \mathbb{R}^2, z_2=(x_2,y_2) : y_2 - x_2^2 \geq 0,$$

On trouve

$$
\begin{aligned}
& \lambda z_1 + (1-\lambda) z_2 \\
= \ & \lambda (x_1,y_1) + (1-\lambda)(x_2,y_2) = (\lambda x_1, \lambda y_1) + ((1-\lambda) x_2, (1-\lambda) y_2) \\
= \ & (\lambda x_1 + (1-\lambda) x_2, \lambda y_1 + (1-\lambda) y_2) \in \mathbb{R}^2 \text{ tel que}
\end{aligned}
$$

$$
\begin{aligned}
& (\lambda y_1 + (1-\lambda) y_2) - (\lambda x_1 + (1-\lambda) x_2)^2 \\
& (\lambda y_1 + (1-\lambda) y_2) - \left(\lambda^2 x_1^2 + (1-\lambda)^2 x_2^2 + 2\lambda(1-\lambda) x_1 x_2\right) \\
\geq \ & \left(\lambda x_1^2 + (1-\lambda) x_2^2\right) - \left(\lambda^2 x_1^2 + (1-\lambda)^2 x_2^2 + 2\lambda(1-\lambda) x_1 x_2\right) \\
\geq \ & \lambda(1-\lambda) x_1^2 + \lambda(1-\lambda) x_2^2 + 2\lambda(1-\lambda) x_1 x_2 \\
\geq \ & \lambda(1-\lambda)\left(x_1^2 + x_2^2 + 2x_1 x_2\right) = \lambda(1-\lambda)(x_1+x_2)^2 \geq 0,
\end{aligned}
$$

d'où S_1 *est convexe.*

Exercice 1.3 *Soient* k_1, k_2 *deux ensembles convexes et* $\alpha, \beta \in \mathbb{R}$. *Montrer que les ensemle suivants :*

$$\begin{aligned} 1.\ S_1 &= \{x \in \mathbb{R}^n / x = \alpha y + \beta z, y \in k_1, z \in k_2\}, \\ 2.\ S_2 &= \left\{x / x = \frac{y+z}{2}, y \in k_1, z \in k_2\right\}, \end{aligned}$$

sont convexes.

————————————————————————

Solution : 1.*Soeint* $x_1, x_2 \in S_1$ *et* $\lambda \in [0,1]$, *alors* $\lambda x_1 + (1-\lambda) x_2 \in S_1$?, *on a :*

$$x_1 \in S_1 \Rightarrow \exists y_1 \in k_1, \exists z_1 \in k_2 : x_1 = \alpha y_1 + \beta z_1,$$

et

$$x_2 \in S_1 \Rightarrow \exists y_2 \in k_1, \exists z_2 \in k_2 : x_2 = \alpha y_2 + \beta z_2,$$

alors

$$\begin{aligned} &\lambda x_1 + (1-\lambda) x_2 \\ &= \lambda (\alpha y_1 + \beta z_1) + (1-\lambda)(\alpha y_2 + \beta z_2) \\ &= \lambda \alpha y_1 + \lambda \beta z_1 + (1-\lambda) \alpha y_2 + (1-\lambda) \beta z_2 \\ &= \lambda \alpha y_1 + (1-\lambda) \alpha y_2 + \lambda \beta z_1 + (1-\lambda) \beta z_2 \\ &= \alpha (\lambda y_1 + (1-\lambda) y_2) + \beta (\lambda z_1 + (1-\lambda) z_2), \end{aligned}$$

comme $(\lambda y_1 + (1-\lambda) y_2) \in k_1$ *et* $(\lambda z_1 + (1-\lambda) z_2) \in k_2$ *(*k_1 *et* k_2 *sont convexes), alors*

$$\lambda x_1 + (1-\lambda) x_2 \in S_1.$$

Donc S_2 *est convexe.*

Exercice 1.4 1) : *Soit* $\mathbb{k} \subseteq \mathbb{R}^n$ *un ensemble convexe. Momtrer que* $f : \mathbb{k} \to \mathbb{R}$ *est convexe si est seulement si l'ensemble suivant :* $S = \{(x,y) \in \mathbb{k} \times \mathbb{R} / \ f(x) \leq y\}$ *est convexe.*

————————————————————————

Solution : 1 : $\Rightarrow$) : *Supposons que* f *est convexe. Soient* (x_1, y_1) *et* (x_2, y_2) *appartenant à* S *:*

On a : $y_1 \geq f(x_1)$ et $y_2 \geq f(x_2)$,
soit $\lambda \in [0,1]$, alors

$$\lambda y_1 + (1-\lambda) y_2 \geq \lambda f(x_1) + (1-\lambda) f(x_2) \geq f[\lambda x_1 + (1-\lambda) x_2] \quad \text{(f est convexe)},$$

donc
$$(\lambda x_1 + (1-\lambda) x_2, \lambda y_1 + (1-\lambda) y_2) \in S.$$
Puisque
$$(\lambda x_1 + (1-\lambda) x_2, \lambda y_1 + (1-\lambda) y_2) = \lambda(x_1, y_1) + (1-\lambda)(x_2, y_2),$$
on en déduit que S est convexe.

$\Leftarrow$) : Suposons que S est convexe. Soient $x_1, x_2 \in \mathbb{k}, \lambda \in [0,1]$.

Remarquons que $(x_1, f(x_1)), (x_2, f(x_2)) \in S$
$$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in S$$

$$\begin{aligned} &\Rightarrow \quad (\lambda x_1 + (1-\lambda) x_2, \lambda f(x_1) + (1-\lambda) f(x_2)) \in S \\ &\Rightarrow \quad \lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2), \end{aligned}$$

d'où f est convexe.

Exercice 1.5 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^2 et $S \subseteq \mathbb{R}^n$ un ensemble convexe. Momtrer que ssi f est convexe sur S, alors $H_f(x)$ est S.D.P, $\forall x \in S$.*

———————————————————————

Solution : $\Leftarrow$)*On suppose que f est convexe sur S et on montre que $H_f(x)$ est S.D.P, $\forall x, y \in S$. Alors*
$$\forall x, y \in \mathbb{k}, \alpha \in [0,1], \alpha x + (1-\alpha) y = y + \alpha(x-y) \in S.$$
Le développement de Taylor au $v(y)$, donne
$$f(y + \alpha(x-y)) = f(y) + \alpha \nabla f(y)(x-y) + \frac{1}{2}\alpha^2 (x-y)^T H_f(y + \theta\alpha(x-y))(x-y).$$
Pour $\alpha = 1$, on a
$$f(x) = f(y) + \nabla f(y)(x-y) + \frac{1}{2}(x-y)^T H_f(y + \theta(x-y))(x-y), \theta \in [0,1],$$
puisque f set covexe $(f(x) \geq f(y) + \nabla f(y)(x-y))$, alors
$$\frac{1}{2}(x-y)^T H_f(y + \theta(x-y))(x-y) \geq 0, \theta \in [0,1], \theta \text{ est trés petit.}$$
Alors
$$(x-y)^T H_f(x)(x-y) \geq 0,$$
donc $H_f(x)$ est *S.D.P*, $\forall x, y \in S$.

Exercice 1.6 1) : *Montrer que si la fonction indicatrice d'un ensemble K définie par :*
$$I_K = \begin{cases} 0, & si\ x \in S, \\ +\infty, & si\ x \notin S, \end{cases}$$

est convexe, alors S est convexe.
2) : *Soit $S \subset \mathbb{R}^n$ un ensemble convexe et $f, g : S \subset \mathbb{R}^n \to \mathbb{R}$ deux fonctions convexes. Montrer que* $\sup(f, g)$ *est convexe.*
3) : *Montrer que l'ensemble $C = \{(x, y) \in \mathbb{R}^2 : xy \geq 0\}$ n'est pas convexe.*

————————————————————————

Solution :

1 : Soient $x, y \in S, \lambda \in [0, 1]$, $\lambda x + (1 - \lambda) y \in I_S$?
On suppose que I_S est convexe, alors

$$I_S(\lambda x + (1 - \lambda) y) \leq \lambda I_S(x) + (1 - \lambda) I_S(y),$$

comme $I_S(x) = I_S(y) = 0$, alors $I_S(\lambda x + (1 - \lambda) y) \leq 0$ et comme I_S prend que les valeurs 0 ou $+\infty$, alors

$$I_S(\lambda x + (1 - \lambda) y) = 0,$$

par conséquent $\lambda x + (1 - \lambda) y \in I_S$, donc S est convexe.
2 :Soient $x, y \in S, \lambda \in [0, 1]$, alors

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

et

$$g(\lambda x + (1 - \lambda) y) \leq \lambda g(x) + (1 - \lambda) g(y),$$

si on pose $h = \sup\{f, g\}$, alors

$$f(x) \leq h(x), f(y) \leq h(y), g(x) \leq h(x),$$

et

$$g(y) \leq h(y),$$

ce qui donne

$$\lambda f(x) + (1 - \lambda) f(y) \leq \lambda h(x) + (1 - \lambda) h(y), \quad (\lambda \geq 0),$$

et

$$\lambda g(x) + (1 - \lambda) g(y) \leq \lambda h(x) + (1 - \lambda) h(y),$$

donc

$$\begin{aligned} & \sup\{f(\lambda x + (1 - \lambda) y), g(\lambda x + (1 - \lambda) y)\} \\ = \; & h(\lambda x + (1 - \lambda) y) \leq \lambda h(x) + (1 - \lambda) h(y). \end{aligned}$$

D'où $\sup\{f, g\}$ est convexe.
3 : $C = \{(x, y) \in \mathbb{R}^2 : xy \geq 0\}$ n'est pas convexe ?
Soient $A = (x_A, y_A)$ et $B = (x_B, y_B)$ deux points de C, alors $x_A.y_A \geq 0$ et $x_B.y_B \geq 0$, et soit $M = (x_M, y_M)$, tq $x_M = (1 - \alpha) x_A + \alpha x_B$ et $y_M = (1 - \alpha) y_A + \alpha y_B$.
Calculons $x_M.y_M$:

$$\begin{aligned} x_M.y_M &= [(1 - \alpha) x_A + \alpha x_B] \times [(1 - \alpha) y_A + \alpha y_B] \\ &= (1 - \alpha)^2 x_A y_A + \alpha^2 x_B y_B + \alpha (1 - \alpha)(x_A y_B + x_B y_A). \end{aligned}$$

Si $\alpha = \frac{1}{2}$, alors

$$x_M.y_M = \frac{1}{4}x_Ay_A + \frac{1}{4}x_By_B + \frac{1}{4}\left(x_Ay_B + x_By_A\right).$$

Si on prend $x_A + x_B > 0$, $y_B + y_A < 0$, alors le point $x_M.y_M < 0$, ce qui implique que le point $M \notin C$,

par conséquent C n'est pas convexe.

par exemple $A = (1, 1)$ et $B = \left(-\frac{1}{2}, -2\right)$, le point $M = (1-\alpha)A + \alpha B \notin C, \alpha = \frac{1}{2}$.

Exercice 1.7 *Soit la forme quadratique $f : \mathbb{R}^n \to \mathbb{R}$ définie par :*

$$f(x) = \frac{1}{2}x^TAx - b^Tx, \ \forall x \in \mathbb{R}^n, b \in \mathbb{R}^n,$$

A : *Matrice symétrique et définie positive.*

1 : *Montrer que la forme quadratique f est strictement convexe.*

2 : *Calculer le gradient et la matrice Hessienne de f.*

————————————————————————

Solution : *Soit $f : \mathbb{R}^n \to \mathbb{R}$ définie par : $f(x) = \frac{1}{2}x^TAx - b^Tx$, $\forall x \in \mathbb{R}^n$,*

1 : Soit $x, y \in \mathbb{R}^n, x \neq y$ et $\lambda \in]0, 1[$

$$\begin{aligned}
& f(\lambda x + (1-\lambda)y) \\
= & \frac{1}{2}(\lambda x + (1-\lambda)y)^T A(\lambda x + (1-\lambda)y) - b^T(\lambda x + (1-\lambda)y) \\
= & \frac{1}{2}\left[\lambda x^TA(\lambda x + (1-\lambda)y) + (1-\lambda)y^TA(\lambda x + (1-\lambda)y)\right] - \lambda b^Tx - (1-\lambda)b^Ty \\
= & \frac{1}{2}\left[\lambda^2x^TAx + (1-\lambda)^2y^TAy + \lambda(1-\lambda)\left(x^TAy + y^TAx\right)\right] - \lambda b^Tx - (1-\lambda)b^Ty \\
= & \lambda f(x) + (1-\lambda)f(y) + \frac{1}{2}\lambda(1-\lambda)\left[-x^TAx + x^TAy + y^TAx - y^TAy\right] \\
& \lambda f(x) + (1-\lambda)f(y) - \frac{1}{2}(y-x)^T A(y-x)^T
\end{aligned}$$

$$\begin{aligned}
\text{puisque } A\ D.P, \text{ ona : } \quad & (y-x)^T A(y-x)^T > 0 \\
\text{donc } f(\lambda x + (1-\lambda)y) \quad < \quad & \lambda f(x) + (1-\lambda)f(y).
\end{aligned}$$

2.1 :

$$f(x+th) = \frac{1}{2}(x+th)^T A(x+th) - b^T(x+th)$$

$$= \frac{1}{2}x^TAx + \frac{t}{2}x^TAh + \frac{t}{2}h^TAx + \frac{t^2}{2}h^TAh - b^Tx - tb^Th$$

$$= \frac{1}{2}x^TAx - b^Tx + tx^TAh - tb^Th + \frac{t^2}{2}h^TAh.$$

Donc

$$\frac{f(x+th)-f(x)}{t} = x^T Ah - b^T h + \frac{t}{2} h^T Ah = (xA-b)^T h + \frac{t}{2} h^T Ah,$$

alors

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(x+th)-f(x)}{t} = (xA-b)^T h,$$

d'où $\nabla f(x) = Ax - b$.

2.2 : D'après 2.1, on a :

$$\begin{aligned} \nabla f(x+th) &= A(x+th) - b = Ax - b + Ath = \nabla f(x) + Ath \\ \nabla f(x+th) - \nabla f(x) &= Ath, \end{aligned}$$

donc

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{\nabla f(x+th) - \nabla f(x)}{t} = Ah,$$

d'où $\nabla^2 f(x) = A$.