

Cours: Optimisation sans contrainte.

Déstiné aux étudiants de 3^{ème} année licence mathématiques

Département de mathématiques, univ de p.jel.

Enseigné par: Menniche Linda.

Chapitre 1: Quelques rappels de calcul différentiel, convexité.

1. Rappel de calcul différentiel

1.1. Notations

- Pour tout $n \in \mathbb{N}^*$, $\mathbb{R}^n$ ($n < \infty$) désigne l'espace euclidien

$\mathbb{R} \times \mathbb{R} \dots \times \mathbb{R}$ (Produit n fois).

un vecteur x de $\mathbb{R}^n$ sera noté $(x_1, \dots, x_n)$.

- On note $e_1, \dots, e_n$ les éléments de la base canonique de $\mathbb{R}^n$.

$$(e_i)_j = \delta_{ij} = \begin{cases} 1 & \text{si } i=j; \forall i, j=1, \dots, n. \\ 0 & \text{si } i \neq j, \end{cases} \quad (\text{symbole de Kronecker}).$$

- Pour tout $x \in \mathbb{R}^n$, on note $\|x\|$ la norme euclidienne de

$$x \text{ donnée par } \|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}.$$

- $\forall x, y \in \mathbb{R}^n: |\langle x, y \rangle| \leq \|x\| \cdot \|y\|$. (Inégalité de Cauchy-Schwarz)

1.2 Dérivée partielle

Définition 1.1: Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue. La fonction notée $\frac{\partial f}{\partial x_i}$ est appelée $i^{\text{ème}}$ dérivée partielle de f et définie

$$\text{par: } \lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_i + t, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{t}.$$

Remarque 1.2.1: Cette limite peut ne pas exister.

1.3. Gradient:

si les dérivées partielles $\frac{\partial f}{\partial x_i}$ existent pour tout i , le gradient de f est défini comme suit:

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

Exemple: Soit $f(x_1, x_2) = e^{x_1} + x_1^2 x_2$.

$$\nabla f(x_1, x_2) = \begin{pmatrix} e^{x_1} + 2x_1 x_2 \\ x_1^2 \end{pmatrix}.$$

Remarque 1.3.1

1) $\frac{\partial f(x)}{\partial 0} = 0$.

2) $\frac{\partial f(x)}{\partial x_i} = \frac{\partial f(x)}{\partial e_i}$.

Définition 1.3.2 (Dérivée directionnelle).

Soit $\Omega \subset \mathbb{R}^n$ et soit $x \in \Omega$ et $h \in \mathbb{R}^n$, on note

$$\frac{\partial f}{\partial h}(x) = \lim_{t \rightarrow 0} \frac{1}{t} [f(x+th) - f(x)] = \langle \nabla f(x), h \rangle.$$

La dérivée directionnelle de f en x de direction h .

Définition 1.3.3 (Fonction différentiable)

Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 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.

1.3 Matrice Hessienne

Définition 1.3.4 On appelle Hessian de f la matrice symétrique de $\mathcal{U}_n(\mathbb{R})$.

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

Remarque: si $f \in C^2(\Omega)$ alors $\nabla^2 f(x)$ est une matrice symétrique
 $\forall x \in \Omega$ (c'est le Théorème de Schwarz).

Proposition 1.31

on a $\nabla^2 f(x)h = \nabla \langle \nabla f(x), h \rangle$, $\forall x \in \Omega$, $\forall h \in \mathbb{R}^n$.

Preuve

$$\begin{aligned} \text{on a: } \frac{\partial}{\partial x_i} \langle \nabla f(x), h \rangle &= \frac{\partial}{\partial x_i} \left(\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) h_j \right) \\ &= \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_j \\ &= \langle \nabla^2 f(x), h \rangle. \end{aligned}$$

1.3 Développement de Taylor

la formule de Taylor est un outil important en convexité.

Nous rappelons dans le cas général.

Soit $\Omega \subset \mathbb{R}^n$ ouvert, $f: \Omega \rightarrow \mathbb{R}$, $a \in \Omega$ et $h \in \mathbb{R}^n$ tels que $[a, a+h]$. Alors:

1) Si $f \in C^1(\Omega)$ alors

a) Formule de Taylor à l'ordre 1 avec reste intégral

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

b) Formule de Taylor-Maclaurin à l'ordre 1

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle$$

c) Formule de Taylor-Young à l'ordre 1.

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + o(\|h\|).$$

2) Si $f \in C^2(\Omega)$ alors :

a) Formule de Taylor à l'ordre 2 avec reste intégral

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

b) Formule de Taylor-Maclaurin à l'ordre 2.

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a)h, h \rangle + o(\|h\|^2).$$

c) Formule de Taylor-Young à l'ordre 2.

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a)h, h \rangle + o(\|h\|^2).$$

Remarque la notation $o(\|h\|^k)$ pour $k \in \mathbb{N}^+$ signifie une expression qui tend vers 0 plus vite que $\|h\|^k$ c.-à-d.

$$\lim_{\|h\| \rightarrow 0} \frac{o(\|h\|^k)}{\|h\|^k} = 0.$$

1.4 Fonctions Convexes

Définition 1.4.1 :

1) un ensemble $C \subset \mathbb{R}^n$ est dit convexe si :

$$\forall (x, y) \in C^2, \forall \lambda \in [0, 1] : \lambda x + (1-\lambda)y \in C.$$

2) Soit $C \subset \mathbb{R}^n$ un ensemble non vide, une fonction $f: C \rightarrow \mathbb{R}$ est convexe si et seulement si :

$$\forall (x, y) \in C^2, \forall t \in [0, 1] : f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

3) on dit que f est strictement convexe si :

$$\forall (x, y) \in C^2, \forall t \in [0, 1] : f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

4) on dit que f est concave si $(-f)$ est convexe.

Théorème 1.4.1 (caractérisation des fcts convexe).

Soit Ω un ouvert de $\mathbb{R}^n$, et soit $C \subset \Omega$ un convexe de Ω ,

Soit $f: \Omega \rightarrow \mathbb{R}$ une fct de classe C^1 . Alors les assertions suivantes sont équivalentes :

- 1) f est convexe sur C .
- 2) $f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle, \forall x, y \in C$.
- 3) $\langle \nabla f(y) - \nabla f(x), y-x \rangle \geq 0, \forall x, y \in C$.

Remarque 1.4.1 : Lorsque f est définie sur un intervalle ouvert $I \subset \mathbb{R}$, $f \in C^2(I)$, l'équivalence de ① et ③ dans le théorème 1.4.1 s'écrit :

$$\begin{aligned} f \text{ est convexe sur } I &\Leftrightarrow \langle \nabla f(y) - \nabla f(x), y-x \rangle \geq 0 \\ &\Leftrightarrow (f'(y) - f'(x))(y-x) \geq 0, \forall x, y \in I. \\ &\Leftrightarrow f' \text{ est croissante sur } I. \\ &\Leftrightarrow f''(x) \geq 0, \forall x \in I \text{ et si } f \in C^2. \end{aligned}$$

Remarque 1.4.2 : on a l'équivalence entre la convexité stricte et les inégalité 2 et 3 rendues strictes pour $x \neq y$.

Théorème 1.4.2 (caractérisation de la convexité avec la matrice hessienne).

Soit f une fct de classe C^2 sur un ouvert convexe $C \subset \mathbb{R}^n$. Alors f est convexe si la matrice hessienne $H(x) = \nabla^2 f(x)$ est semi-définie positive pour tout $x \in C$ c.à.d $\langle \nabla^2 f(x) h, h \rangle \geq 0, \forall h \in \mathbb{R}^n$.

Théorème 1.4.3: Soit f une fct de classe C^2 sur un ouvert convexe C de $\mathbb{R}^n$, Alors f est strictement convexe si la matrice hessienne est définie positive pour tout $x \in C$. $C = \bar{a} - d$:

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

Preuve: supposons que $\nabla^2 f(x)$ est définie positive pour tout $x \in C$, la formule de Taylor-Maclaurin donne:

$$\forall (x, y) \in C^2, x \neq y: f(y) = f(x) + \langle \nabla f(x), y-x \rangle + \frac{1}{2} \langle \nabla^2 f(\underbrace{x+t(y-x)}_{\in C}) (y-x), y-x \rangle$$

$$\text{D'où } f(y) > f(x) + \langle \nabla f(x), y-x \rangle, \quad \forall x, y \in C, x \neq y.$$

donc f est strictement convexe d'après le Théorème 1.4.1.

Remarque 1.4.3: la réciproque est fautive.

Proposition 1.4.4: Soit $f: C \rightarrow \mathbb{R}$ une fonction convexe, C étant un convexe de $\mathbb{R}^n$. Alors:

- 1) Tout minimum local de f sur C est un minimum global.
- 2) Si f est strictement convexe, il y a au plus un point de minimum global.