CH2 : Fonctions convexes multidimensionnelles

Dans ce chapitre, on va étudier les fonctions multidimensionnelles convexes, concaves, le vecteur gradient, la matrice hessienne et quelques résultats qui sont les plus importants. Soit $f : S \subseteq R^n \to R$, $S \neq \phi$, supposons f une fonction deux fois continûment différentiable, sur un ensemble S convexe, f admet un vecteur gradient et une matrice hessienne.

Définition1 : vecteur gradient

Le vecteur colonne des dérivées partielles de f s'appelle vecteur gradient de f, noté par :

$\nabla f(x) = (\partial f(x)/\partial x_1, \ldots, \partial f(x)/\partial x_n)^t$ vecteur gradient de f

Définition2 : matrice hessienne

Les dérivées partielles de f sont elles mêmes des fonctions de x , $x \in R^n$, si ces fonctions sont continûment différentiables, on peut donc en tirer chaque fois n dérivées partielles ; on obtient alors une matrice carrée d'ordre n, cette matrice s'appelle matrice hessienne de f , notée par :

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} \partial^2 f(x)/(\partial x_1)^2 & \partial^2 f(x)/(\partial x_1 \partial x_2) & \partial^2 f(x)/(\partial x_1 \partial x_n) \\ \partial^2 f(x)/(\partial x_2 \partial x_1) & \partial^2 f(x)/(\partial x_2)^2 & \partial^2 f(x)/(\partial x_2 \partial x_n) \\ & & \\ \partial^2 f(x)/(\partial x_n \partial x_1) & \partial^2 f(x)/(\partial x_n \partial x_2) & \partial^2 f(x)/(\partial x_n)^2 \end{bmatrix}$$ matrice hessienne

Remarque:

Pour simplifier la notation, on écrit : $\partial f(x)/\partial x_i$ $i = 1,\ldots,n$ $\qquad \frac{\partial f(x)}{\partial x_1} = f_1$

$\partial^2 f(x)/\partial x_i \partial x_j = \partial/\partial x_i(\partial f(x)/\partial x_j) = f_{ij}$ $i,j = 1,\ldots,n$

On a : $\partial/\partial x_i(\partial f(x)/\partial x_j) = \partial/\partial x_j \partial f(x)/\partial x_i)$ $i, j = 1,\ldots,n$ égalité de schwartz

Une propriété très importante de la matrice hessienne est donnée par le théorème suivant connu sous le nom de théorème de schwartz

Théorème1 :

Soit $f : R^n \to R$, si f est deux fois continûment différentiable sur R^n, alors la matrice hessienne de f est symétrique.

Preuve :

$H(x) = \nabla^2 f(x)$ matrice hessienne , d'après l'égalité de schwartz :

$\partial^2 f(x)/\partial x_i \partial x_j = \partial^2 f(x)/\partial x_j \partial x_i$ $i , j = 1,\ldots,n$ ce qui implique que $H(x) = H^t(x)$ donc H(x) est symétrique.

Exemple :

Soit $f : R^2 \to R$

$x \to f(x) = 2x_1^2 + 3x_2^2$

f est 2 fois continûment différentiable sur R^2

$\nabla f(x) = (\partial f(x)/\partial x_1 , \partial f(x)/\partial x_2)^t = (4x_1 , 6x_2)$ vecteur gradient

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} \partial^2 f(x)/(\partial x_1)^2 & \partial^2 f(x)/(\partial x_1 \partial x_2) \\ \partial^2 f(x)/(\partial x_2 \partial x_1) & \partial^2 f(x)/(\partial x_2)^2 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$$

Définition 3:

Soit une matrice carrée symétrique d'ordre n

i) A est dite définie positive si $x^t A x > 0 \quad \forall x ; x \in R^n ; x \neq 0$

A est dite semi définie positive si $x^t A x \geq 0 \quad \forall x ; x \in R^n$

ii) A est dite définie négative si $x^t A x < 0 \quad \forall x ; x \in R^n ; x \neq 0$

A est dite semi définie négative si $x^t A x \leq 0 \quad \forall x ; x \in R^n$

Théorème2 :

Une matrice symétrique A est définie négative (semi définie négative) *si et seulement si* - A est définie positive (semi définie positive).

Deux résultats importants : a) en utilisant les valeurs propres

b) en utilisant la méthode des mineurs principaux

En utilisant les valeurs propres :

Théorème3 :

Soit une matrice symétrique, alors :

i) A est définie positive si et seulement si toutes ses valeurs propres sont strictement positives. A est semi définie positive si et seulement si toutes ses valeurs propres sont positives ou nulles.

ii) A est définie négative si et seulement si toutes ses valeurs propres sont strictement négatives. A est semi définie négative si et seulement si toutes ses valeurs propres sont négatives ou nulles.

iii) A est indéfinie si seulement si certaines de ses valeurs propres sont strictement positives et d'autres strictement négatives.

Preuve :

Soit une matrice carrée symétrique d'ordre n ; λ valeur propre de A ; x vecteur propre associé à λ ; $x \neq 0$

$Ax = \lambda x \Rightarrow x^t A x = x^t \lambda x = \lambda x^t x$ or $x^t x = \sum x_i^2 > 0$; $x \neq 0$ donc le signe de $x^t A x$ est le même que celui de λ ; $x^t A x = \lambda \| x \|^2$

En utilisant la méthode des mineurs principaux :

Soit $f : R^n \rightarrow R$ fonction 2 fois continûment différentiable sur R^n

$\nabla f(x) = (\partial f(x)/\partial x_1 , \ldots\ldots, \partial f(x)/\partial x_n)^t$ vecteur gradient de f

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} f_{11} & & f_{1n} \\ & \cdot & \\ f_{n1} & & f_{nn} \end{bmatrix} \qquad \text{où } f_{ij} = \frac{\partial^2 f(x)}{\partial x_i \, \partial x_j} \quad i, j = 1,\ldots,n \quad \text{matrice hessienne}$$

On note par H_j la matrice hessienne dont on a éliminé n-j colonnes et n-j lignes.

Le déterminant de H_j ($| H_j |$) est appelé mineur principal.

Exemple :

$H_1 = f_{11}$ (on a éliminé de H,n-1 lignes et colonnes, à partir de la dernière).

$$H_2 = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$ (on a éliminé de H, n-2 lignes et colonnes à partir de la dernière ligne et la dernière colonne)

$$H_3 = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

Proposition1:

*) Si $|H_j| > 0$; $\forall x \in R^n$; $\forall j$; $j = 1,\dots,n$ alors H est dite définie positive

*) Si $(-1)^j |H_j| > 0$; $\forall x \in R^n$; $\forall j$; $j = 1,\dots,n$ alors H est dite définie négative

Proposition2 :

Soit x^* un point stationnaire $\Rightarrow \nabla f(x^*) = 0$

i) Si $|H_j(x^*)| > 0$; $\forall j$; $j = 1,\dots,n$ alors $(x^*, f(x^*))$ est un minloc, pour la fonction f

ii) Si $(-1)^j |H_j(x^*)| > 0$; $\forall j$; $j = 1,\dots,n$ alors $(x^*, f(x^*))$ est un maxloc, pour la fonction f

iii) $(x^*, f(x^*))$ est un point selle si $H(x^*)$ est indéfinie ie H n'est ni définie positive, ni définie négative.

Remarque :

Si $f_{11} > 0$, $|H_j| \geq 0$; $\forall j$; $j = 2,\dots,n$; $\forall x \in R^n$, H est semi définie positive

Si $f_{11} < 0$; $(-1)^j |H_j| \geq 0$; $\forall j$; $j = 2,\dots,n$ alors H est dite semi définie négative $\forall x \in R^n$

Théorème 4 :

Soit $f : R^n \to R$ fonction 2 fois continûment différentiable sur R^n et soit H(x) la matrice hessienne de f, on a :

a) f convexe sur R^n si et seulement si H(x) est semi définie positive, en tout point x, $x \in R^n$

b) f est concave sur R^n si et seulement si H(x) est semi définie négative, en tout point x, $x \in R^n$

Théorème 5 :

Soit $f : R^n \to R$ fonction 2 fois continûment différentiable sur R^n et soit H(x) la matrice hessienne de f, on a :

a) f est strictement convexe sur R^n si et seulement si H(x) est définie positive, en tout point x, $x \in R^n$ (<, >)

b) f est strictement concave sur R^n si et seulement si H(x) est définie négative, en tout point x, $x \in R^n$