MAT 2410: Optimisation

Robert Guénette

Département de Mathématiques et de Statistique

Chapitre 2
Dérivabilité et convexité

Références:

- Notes de cours: chapitre 2, sections 2.1 et 2.2.
- Livre de M. Delfour: chapitre 2.

Notations

On notera par $\{e_i\}_{i=1}^n$ la base canonique de $\mathbb{R}^n$

$$\begin{array}{ccc} e_1 & = & (1,0,0,\ldots,0) \\ e_2 & = & (0,1,0,\ldots,0) \\ \vdots & = & \vdots \\ e_n & = & (0,0,0,\ldots,1) \end{array}$$

Soit $f : U \to \mathbb{R}$ une fonction numérique définie sur un ouvert U de $\mathbb{R}^n$.
Définition: dérivées partielles
On dit que f admet des dérivées partielles au point $x_0 \in U$ si les limites suivantes existent

$$\frac{\partial f}{\partial x_i}(x_0) \equiv f_{x_i}(x_0) = \lim_{t\to 0} \frac{f(x_0+te_i) - f(x_0)}{t} = \frac{d}{dt} f(x_0+te_i)_{|t=0} \; \forall i = 1,2,\ldots,n$$

Définition: fonction de classe C^1
On dit que f est de classe C^1 sur U si

- $\frac{\partial f}{\partial x_i}(x)$ existent $\forall x \in U$ et $\forall i = 1,2,\ldots,n$,
- les fonctions $\frac{\partial f}{\partial x_i}(x)$ sont continues dans U.

Rappel: fonction d'une variable

Théorème de la moyenne
Si $F : [a, b] \to \mathbb{R}$ est continue et dérivable partout dans $]a, b[$, il existe un point $\bar{x} \in]a, b[$ tel que

$$F(b) - F(a) = F'(\bar{x})(b - a)$$

Théorème de Taylor
Soit $F : [a, b] \to \mathbb{R}$ admettant des dérivées $F', F'', \ldots, F^{(n)}$ continues dans $]a, b[$ et que la dérivée $F^{(n+1)}(t)$ existe pour tous les $t \in]a, b[$. Si x_0 et x sont deux points de $]a, b[$, alors il existe un point α entre x_0 et x tel que

$$\begin{aligned} F(x) &= F(x_0) + F'(x_0)(x - x_0) + F''(x_0)\frac{(x - x_0)^2}{2} + \ldots \\ &+ F^{(n)}(x_0)\frac{(x - x_0)^n}{n!} + F^{(n+1)}(\alpha)\frac{(x - x_0)^{n+1}}{(n+1)!} \end{aligned}$$

Rappel: fonction de plusieurs variables

Soit $f : U \to \mathbb{R}$ une fonction numérique définie sur un ouvert U de $\mathbb{R}^n$ de classe C^1.

Définition: gradient

Le vecteur gradient au point x_0 correspond au vecteur

$$\nabla f(x_0) = (\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0))$$

Définition: dérivée directionnelle

La dérivée directionnelle de f dans la direction v est définie par

$$D_v f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \frac{d}{dt} f(x_0 + tv)_{|t=0}$$

Formule pour la dérivée directionnelle : $D_v f(x_0) = (\nabla f(x_0), v)$.

Théorème de la moyenne

Soit $f : U \to \mathbb{R}$ une fonction numérique de classe C^1 définie sur un ouvert U de $\mathbb{R}^n$.

Théorème de la moyenne

Soit $x, y \in U$. Si le segment de droite $[x, y] \subset U$, alors il existe un point $\bar{t} \in]0, 1[$ tel que

$$f(y) = f(x) + (\nabla f(x + \bar{t}(y - x)), \;\; y - x)$$

Preuve: on considère la fonction d'une variable

$$F(t) = f(x + t(y - x))$$

On a que

$$\begin{array}{rcl} F(0) & = & f(x) \\ F(1) & = & f(y) \\ F'(t) & = & (\nabla f(x + t(y - x)), y - x) \end{array}$$

On applique le théorème de la moyenne pour F sur l'intervalle $[0, 1]$:

$$\exists \bar{t} \in]0, 1[\qquad F(1) = F(0) + F'(\bar{t}) \Leftrightarrow f(y) = f(x) + (\nabla f(x + \bar{t}(y - x)), y - x)$$

Dérivées d'ordre supérieure

Définition: f est une fonction de classe C^2 si

- f est continue,
- $\frac{\partial f}{\partial x_i}$ est continue $\forall i = 1, 2, \dots, n$,
- $\frac{\partial^2 f}{\partial x_i \partial x_j}$ est continue $\forall i, j = 1, 2, \dots, n$.

Définition: matrice hessienne
La matrice hessienne au point $x_0 \in U$ est définie par

$$H_f(x_0) = \left(\, H_{i,j}(x_0)\, \right)_{i,j} = \left(\, \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\, \right)_{i,j}$$

la matrice est toujours symétrique!

On a la formule suivante

$$(Hu, v) = \frac{\partial^2}{\partial s \partial t} f(x_0 + su + tv)_{|t=s=0}$$

Théorème de Taylor: fonction de plusieurs variables

Théorème de Taylor: cas d'ordre 2
Soit $f : U \to \mathbb{R}$ de classe C^2 définie sur un ouvert U de $\mathbb{R}^n$. On se donne 2 points $x, y \in U$. Si le segment de droite $[x, y] \subset U$, alors il existe un point $\bar{t} \in]0, 1[$ tel que

$$f(y) = f(x) + (\nabla f(x), y - x) + \frac{1}{2} \left(H(x + \bar{t}(y - x))\,(y - x), (y - x)\right)$$

Preuve: on considère la fonction d'une variable

$$F(t) = f(x + t(y - x))$$

On a que

$$\begin{aligned} F(0) &= f(x) \\ F(1) &= f(y) \\ F'(t) &= (\nabla f(x + t(y - x)), y - x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x + t(y - x))(y_i - x_i) \end{aligned}$$

Théorème de Taylor: cas d'ordre 2 (suite)

Calcul de $F''(t)$:

$$\begin{aligned} F''(t) &= \tfrac{d}{dt} \sum_{i=n}^{n} \frac{\partial f}{\partial x_i}(x + t(y-x))(y_i - x_i) \\ &= \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x + t(y-x))(y_i - x_i)(y_j - x_j) \end{aligned}$$

On applique le théorème de Taylor pour F sur l'intervalle $[0,1]$:

$$\exists \bar{t} \in]0,1[\quad \text{tel que} \quad F(1) = F(0) + F'(0) + \frac{F''(\bar{t})}{2}$$

En remplacant les valeurs ci-dessus, on obtient le résultat.

Critère de convexité d'ordre un

Théorème: critère de convexité d'ordre 1
Soit $f : U \longrightarrow \mathbb{R}$ une fonction numérique de classe C^1 définie sur un ensemble ouvert convexe $U \subset \mathbb{R}^n$.

(a) f est convexe si et seulement si

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

(b) f est strictement convexe si et seulement si

$$f(y) > f(x) + (\nabla f(x), y - x) \qquad \forall x \neq y \in U$$

Preuve de (a): $\Longrightarrow$ $\qquad f$ convexe implique

$$\begin{array}{rcl}
f(x + t(y-x)) = f((1-t)x + ty) & \leq & (1-t)f(x) + tf(y) \\
& = & f(x) + t(f(y) - f(x)) \\
\Longrightarrow \quad \dfrac{f(x+t(y-x)) - f(x)}{t} & \leq & f(y) - f(x) \\
\Longrightarrow \quad \lim\limits_{t \to 0^+} \dfrac{f(x+t(y-x)) - f(x)}{t} & \leq & f(y) - f(x) \\
(\nabla f(x), y - x) & \leq & f(y) - f(x)
\end{array}$$

(suite) preuve du critère de convexité d'ordre un

Preuve de (a): $\Longleftarrow$
Par hypothèse on a

$$f(z) \geq f(w) + (\nabla f(w), z-w) \qquad \forall z, w \in U \tag{1}$$

On pose $z = x$ et $w = x + t(y-x) \Longrightarrow z - w = -t(y-x)$. On remplace dans (1)

$$f(x) \geq f(x + t(y-x)) - t(\nabla f(x + t(y-x)), y - x)$$

On pose $z = y$ et $w = x + t(y-x) \Longrightarrow z - w = (1-t)(y-x)$. On remplace dans (1)

$$f(y) \geq f(x + t(y-x)) + (1-t)(\nabla f(x + t(y-x)), y - x)$$

On forme la combinaison convexe des deux lignes précédentes pour obtenir

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

d'où la convexité de f.

Critère de convexité d'ordre deux

Théorème: critère de convexité d'ordre 2

Soit $f : U \longrightarrow \mathbb{R}$ une fonction numérique de classe C^2 définie sur un ensemble ouvert convexe $U \subset \mathbb{R}^n$.

(a) f est convexe si et seulement si

$$H(x) \geq 0 \qquad \forall x \in U$$

où $H(x)$ est la matrice Hessienne de f au point x. La notation $A \geq 0$ signifie que $(Au, u) \geq 0 \quad \forall u \in \mathbb{R}^n$.

(b) s'il existe un point $x \in U$ tel que $H(x) > 0$ alors on peut trouver un voisinage du point x tel que f soit strictement convexe dans ce voisinage.

Preuve de (a): $\Longrightarrow$

On utilise le théorème de Taylor

$$f(y) = f(x) + (\nabla f(x), y - x) + \frac{1}{2}\left(H(x + \bar{t}(y - x))(y - x), (y - x)\right)$$

Or f est convexe. Le critère de convexité d'ordre 1 est

$$f(y) \geq f(x) + (\nabla f(x), y - x)$$

Preuve de (a): $\Longrightarrow$ (suite)

En combinant les deux relations, on obtient

$$(H(x+\bar{t}(y-x))\,(y-x),(y-x))\geq 0$$

Notons $\bar{x}=x+\bar{t}(y-x)$. Pour tout $v\in\mathbb{R}^n$, on peut toujours choisir $y=x+r\,v\in U$ pour $r>0$ suffisamment petit car U est ouvert. Par conséquent

$$(H(\bar{x})\,(y-x),(y-x))=(H(\bar{x})\,r\,v,r\,v)\geq 0\Longleftrightarrow(H(\bar{x})\,v,v)\geq 0\Longleftrightarrow H(\bar{x})\geq 0.$$

Mais l'inégalité est encore vraie si y s'approche de x tout en restant sur le segment de droite, autrement dit $\lim\limits_{r\to 0}\bar{x}=x$,

$$\Longrightarrow\qquad (H(x)\,v,v)\geq 0\qquad \forall v\in\mathbb{R}^n$$

d'où le résultat.

Preuve de (a): $\Longleftarrow$

De nouveau, on applique le théorème de Taylor

$$f(y) = f(x) + (\nabla f(x), y - x) + \frac{1}{2}\,(H(x + \bar{t}(y - x))\,(y - x), (y - x)).$$

Mais $H(x) \geq 0 \quad \forall x \in U \Longrightarrow H(x + \bar{t}(y - x)) \geq 0.$ On obtient

$$f(y) \geq f(x) + (\nabla f(x), y - x)$$

ce qui est exactement le critère de convexité d'ordre 1, d'où la convexité de f.