

ANALYSE DANS \mathbb{R}^n - LICENCE 2

Aldric Labarthe - Université Paris 1

TD7 - Dérivées secondes et de nouveaux problèmes d'optimisation (± 2 séances)

Cours 1: Dérivées secondes

Dérivée seconde Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable. Si f est deux fois dérivable, la dérivée seconde est définie comme la dérivée de la dérivée première. Pour une fonction $f(x, y)$, les dérivées secondes sont :

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right), \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right).$$

Matrice Hessienne Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction deux fois continûment différentiable. La matrice hessienne de f en un point $a \in \mathbb{R}^n$, notée $\mathbf{H}_f(a)$, est la matrice carrée d'ordre n dont les éléments sont les dérivées secondes partielles de f en a . Elle est donnée par :

$$\mathbf{H}_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

Théorème de Schwarz Si les dérivées secondes croisées $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existent et sont continues sur D , alors elles sont égales :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Différentielle seconde Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable de classe C^2 dans un voisinage d'un point $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$. La différentielle seconde de f au point \mathbf{a} est une forme bilinéaire associée à la matrice hessienne de f , qui mesure la variation du gradient de f en fonction de la variation des coordonnées.

La différentielle seconde de f en un point \mathbf{a} est définie par :

$$d^2 f_{\mathbf{a}}(\mathbf{h}) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) h_i h_j = \mathbf{h}^T H_f(\mathbf{a}) \mathbf{h}$$

où $\mathbf{h} = (h_1, h_2, \dots, h_n)$ est le vecteur de direction dans \mathbb{R}^n , et $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$ représente les éléments de la matrice hessienne $H_f(\mathbf{a})$.

La différentielle seconde peut être interprétée géométriquement comme une mesure de la courbure de la fonction dans une direction donnée.

Formule de Taylor (ordre 2) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable de classe C^2 dans un voisinage de $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$. La formule de Taylor à l'ordre 2 de f en \mathbf{a} s'écrit :

$$f(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a})^T (\mathbf{x} - \mathbf{a}) + \frac{1}{2} (\mathbf{x} - \mathbf{a})^T H_f(\mathbf{a}) (\mathbf{x} - \mathbf{a}) + o(\|\mathbf{x} - \mathbf{a}\|^2),$$

Cette formule est une approximation de $f(\mathbf{x})$ près du point \mathbf{a} , qui inclut la valeur de la fonction, les dérivées premières et secondes.

Cours 2: Formes quadratiques

Forme quadratique associée à une matrice carrée Soit A une matrice carrée d'ordre n , c'est-à-dire une matrice $A \in \mathbb{R}^{n \times n}$. Une *forme quadratique* associée à A est une expression du type :

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x},$$

où $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ est un vecteur colonne de \mathbb{R}^n et A est une matrice symétrique, c'est-à-dire $A = A^T$.

Définie positive, définie négative, ou indéfinie

- Si A est définie positive, la forme quadratique $Q(\mathbf{x})$ est strictement positive pour tous les vecteurs $\mathbf{x} \neq 0$.
- Si A est définie négative, $Q(\mathbf{x})$ est strictement négative pour tous les vecteurs $\mathbf{x} \neq 0$.
- Si A est indéfinie, $Q(\mathbf{x})$ peut prendre des valeurs positives et négatives pour des vecteurs \mathbf{x}

Autres caractérisations par la matrice associée

- *Par les valeurs propres* : Pour tout $n > 0$ fini, une matrice réelle $A \in \mathcal{M}(n, n)$ symétrique est (définie) positive, ssi toutes ses valeurs propres sont (resp. strictement) positives.
- *Par l'opposée* : Pour tout $n > 0$ fini, une matrice réelle $A \in \mathcal{M}(n, n)$ symétrique est (définie) négative ssi $-A$ est (définie) positive.
- *Critère de Sylvester* : Pour tout $n > 0$ fini, une matrice réelle $A \in \mathcal{M}(n, n)$ symétrique est définie positive ssi les n matrices $A_p = (a_{ij})_{1 \leq i, j \leq p}$ pour $p = 1, \dots, n$ ont leur déterminant (appelé mineur diagonal) strictement positif.

Cours 3: Condition de second ordre en optimisation

Point selle Un point col ou point-selle (en anglais : *saddle point*) d'une fonction f définie sur un produit cartésien $X \times Y$ de deux ensembles X et Y est un point $(\bar{x}, \bar{y}) \in X \times Y$ tel que :

- $y \mapsto f(\bar{x}, y)$ atteint un maximum en \bar{y} sur Y ;
- $x \mapsto f(x, \bar{y})$ atteint un minimum en \bar{x} sur X .

Caractérisation forte des extrema locaux Pour f une fonction deux fois différentiable sur A , et $\bar{x} \in A$ un point critique de cette fonction, et $q_{\bar{x}}^f : h \mapsto \mathbf{h}^T H_f(\bar{\mathbf{x}}) \mathbf{h}$ la forme quadratique associée à la matrice hessienne évaluée au point \bar{x} , on a :

- Si $H_f(\bar{\mathbf{x}})$ est définie positive alors \bar{x} est un minimum local (fort).
- Si $H_f(\bar{\mathbf{x}})$ est définie négative alors \bar{x} est un maximum local (fort).
- Si $H_f(\bar{\mathbf{x}})$ est indéfinie positive alors \bar{x} est un point selle.

Caractérisation faible des extrema locaux Pour f une fonction deux fois différentiable sur A , et $\bar{x} \in \text{int}(A)$ un point critique de cette fonction, et $q_{\bar{x}}^f : h \mapsto \mathbf{h}^T H_f(\bar{\mathbf{x}}) \mathbf{h}$ la forme quadratique associée à la matrice hessienne évaluée au point \bar{x} , on a :

- Si \bar{x} est un minimum local alors $H_f(\bar{\mathbf{x}})$ est semi-positive.
- Si \bar{x} est un maximum local alors $H_f(\bar{\mathbf{x}})$ est semi-négative.

Ces deux caractérisations donnent la méthodologie suivante :

1. Trouver les points critiques.
2. Calculer la hessienne localement.
3. Déterminer si elle est définie, et conclure.
4. Si elle ne l'est pas (c'est-à-dire si elle est semi-définie), alors étudier par les courbes le comportement de la fonction.