Calcul Différentiel et Analyse Complexe
–
2ème séance de cours

Inégalité des accroissements finis

Nous rappelons d'abord le théorème des accroissements finis pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Théorème 1. *Si $f : [a,b] \to \mathbb{R}$ est une fonction continue, dérivable dans $]a,b[$, il existe un point $\xi \in]a,b[$ tel que $f'(\xi) = (f(b) - f(a))/(b-a)$.*

On voudrait étendre ce théorème au cas multidimensionnel. Première mauvaise surprise : il est faux même pour des fonctions d'une variable, lorsqu'elles sont à valeur dans $\mathbb{R}^m$, même pour $m = 2$. On peut considérer l'exemple suivant

$$[a,b] = [0, 2\pi], \quad f(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

En effet $f(b) - f(a) = 0$, mais $f'(t)$ ne s'annule jamais comme vecteur (ses composantes s'annulent, mais pas en un même point).

Il faut donc chercher à l'appliquer toujours à des fonctions à valeur dans $\mathbb{R}$. Aussi, on voudrait considérer le cas des fonctions définies sur des sous-ensembles de $\mathbb{R}^n$, $n > 1$, ce qu'on fera, mais le théorème prendra la forme d'une inégalité.

Pour ce faire, nous introduisons d'abord la notion de norme d'une application linéaire : si $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ (c'est-à-dire, si L est une application linéaire de $\mathbb{R}^n$ vers $\mathbb{R}^m$), on définit

$$|||L||| := \sum_{h \in \mathbb{R}^n \setminus \{0\}} \frac{|L(h)|}{|h|},$$

où $||\cdot||$ indique la norme euclidienne. La même définition pourrait être utilisée en remplaçant la norme euclidienne de $L(h)$ et/ou de h par d'autres normes, en obtenant un résultat différent. Cette quantité est une norme sur l'espace des applications linéaires (ou des matrices), différent de la norme euclidienne de la matrice vue comme un vecteur de $\mathbb{R}^{nm}$.

Théorème 2 (Inégalité des accroissements finis)**.** *Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $\bar{x}, \bar{y} \in \Omega$ deux points de Ω et supposons $[\bar{x}, \bar{y}] \subset \Omega$, où $[\bar{x}, \bar{y}]$ est le segment $\{x = (1-t)\bar{x} + t\bar{y}, t/in[0,1]\}$. Si $f : \Omega \to \mathbb{R}^m$ est une fonction différentiable dans Ω, on a*

$$|f(\bar{x}) - f(\bar{y})| \le |\bar{x} - \bar{y}| \sup\{|||Df(x)||| \,:\, x \in [\bar{x}, \bar{y}]\}.$$

Démonstration. On considère la fonction $g : [0,1] \to \mathbb{R}$ donnée par $g(t) = |f((1-t)\bar{x} + t\bar{y}) - f(\bar{x})|$, qui est continue, et différentiable (par composition) là où elle n'est pas nulle. On définit $t_0 := \sup\{t : g(t) = 0\}$. Si $t_0 = 1$ il n'y a rien à démontrer, sinon on applique le théorème des accroissements finis à g sur $[t_0, 1]$. Il reste à vérifier que l'on a $|g'(t)| \le |||Df((1-t)\bar{x} + t\bar{y})||| \, |\bar{x} - \bar{y}|$, ce qui permet d'arriver au résultat souhaité.

□

Dérivées d'ordre deux

On considère une fonction $f : \Omega \to \mathbb{R}$, avec $\Omega \subset \mathbb{R}^n$ un ouvert (sinon, on peut la regarder composante par composante, si c'est une fonction à valeurs vectoriels). On a défini son gradient $\nabla f : \Omega\mathbb{R}^n$, si cette fonction est différentiable (l'existence des dérivées partielles suffit pour définir le gradient comme vecteur). On peut regarder la différentielle du gradient en un point $\bar{x} \in \Omega$, c'est une matrice carrée $n \times n$ appelée Hessienne. Ses composantes sont données par les dérivées secondes

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\bar{x}).$$

Un point important est le fait que cette matrice est, sous une simple hypothèse de continuité, symétrique :

Théorème 3 (Théorème de Schwarz)**.** *Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $\bar{x} \in \Omega$, et $f : \Omega \to \mathbb{R}$ une fonction C^1 dans un voisinage de $\bar{x}$, telles que ses dérivées partielles admettent aussi des dérivées partielles par rapport à toutes les variables, et que ces dérivées secondes sont continues en $\bar{x}$. Alors on a*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\bar{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\bar{x}).$$

Ce théorème se démontre, entre autre, par une double application du théorème des accroissements finis pour des fonctions réelles d'une variable réelle. On indiquera désormais par $\partial^2 f/(\partial x_i \partial x_j)$ les dérivées partielles secondes, sans se soucier de l'ordre de dérivation.

Les dérivées secondes servent aussi à donner un développement limité d'ordre deux aux fonctions de plusieurs variables. Nous rappelons d'abord ce qu'un DL2 dans le cas d'une variable : si f est C^{k-1} et qu'elle admet une dérivée k-ème en $\bar{x}$ on a

$$f(\bar{x}+h) = \sum_{j=0}^{k} \frac{1}{j!} f^{(j)}(\bar{x}) h^j \; + o(h^k)$$

lorque $h \to 0$. Si on se limite aux cas $k = 1, 2$ on a

$$f(\bar{x}+h) = f(\bar{x}) + f'(\bar{x})h + o(h), \quad f(\bar{x}+h) = f(\bar{x}) + f'(\bar{x})h + \frac{f''(\bar{x})}{2}h^2 + o(h^2).$$

La première formule se généralise dans le cas de plusieurs variables en

$$f(\bar{x}+h) = f(\bar{x}) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\bar{x}) h_j + o(|h|),$$

et ce n'est rien d'autre que la définition de différentielle et son identification avec le produit scalaire avec le gradient. La deuxième par contre donne le DL que l'on trouve dans cet énoncé (qu'on présente par simplicité dans le cas C^2, alors que la différentiabilité du gradient suffirait...)

Théorème 4. *Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $\bar{x} \in \Omega$, et $f : \Omega \to \mathbb{R}$ une fonction C^2. Alors on a*

$$f(\bar{x}+h) = f(\bar{x}) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\bar{x}) h_j + \frac{1}{2} \sum_{l,k=1}^{n} \frac{\partial^2 f}{\partial x_l \partial x_k}(\bar{x}) h_l h_k + o(|h|^2).$$

Démonstration. On considère

$$R(h) := f(\bar{x}+h) - f(\bar{x}) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\bar{x}) h_j - \frac{1}{2} \sum_{l,k=1}^{n} \frac{\partial^2 f}{\partial x_l \partial x_k}(\bar{x}) h_l h_k$$

et on veut démontrer $R(h) = o(|h|^2)$. On a évidemment $R(0) = 0$, si on démontre que $\partial R/\partial h_j = o(|h|)$ pour tout j, l'inégalité des accroissements finis nous donne $|R(h)| = |R(h) - R(0)| \leq |h|o(|h|) = o(|h|^2)$. En effet, on aurait $\sup\{|||DR(x)||| : x \in [0, h]\} = o(|h|)$ aussi.

Pour estimer $\partial R/\partial h_j$ il suffit de le calculer, en regardant les termes qui contiennent h_j, et on obtient

$$\begin{aligned}\frac{\partial R}{\partial h_j} &= \frac{\partial f}{\partial x_j}(\bar{x}+h) - \frac{\partial f}{\partial x_j}(\bar{x}) - \frac{1}{2}\frac{\partial^2 f}{\partial x_j^2}(\bar{x})2h_j - \frac{1}{2}\sum_{k\neq j}\frac{\partial^2 f}{\partial x_j \partial x_k}(\bar{x})h_k - \frac{1}{2}\sum_{l\neq j}\frac{\partial^2 f}{\partial x_l \partial x_j}(\bar{x})h_l \\ &= \frac{\partial f}{\partial x_j}(\bar{x}+h) - \frac{\partial f}{\partial x_j}(\bar{x}) - \sum_{k}\frac{\partial^2 f}{\partial x_j \partial x_k}(\bar{x})h_k.\end{aligned}$$

Or, cette dernière expression est la différence entre la fonction $\partial f/\partial x_j$ et son DL1 autour de $\bar{x}$, les deux calculés en $\bar{x} + h$. Cette différence est donc $o(|h|)$.

$\square$