Calcul différentiel

Conventions et notations

1. Contexte.
— n est un entier naturel non nul.
— S'il n'y a pas d'ambigüité, i et j sont dans $[\![1, n]\!]$.
— On se place dans $\mathbb{R}^n$ euclidien usuel, dont la base canonique est notée $\mathscr{B} = (e_1, \ldots, e_n)$, le produit scalaire $(\cdot\,|\,\cdot)$ et la norme $\|\cdot\|$.
— Les éléments de $\mathbb{R}^n$ sont considérés indifféremment comme des points ou des vecteurs, des n-uplets ou des vecteurs colonnes : si $x \in \mathbb{R}^n$,

$$x = (x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i\, e_i.$$

— U est un *ouvert* non vide de $\mathbb{R}^n$; si une propriété est vraie en tout point de U, on dit qu'elle est vraie *sur* U ;
— a est un point quelconque de U ;
— h est un vecteur quelconque de $\mathbb{R}^n$, avec l'éventuelle contrainte que $a + h \in U$.
— v est un vecteur quelconque *non nul* de $\mathbb{R}^n$.
— On considère une fonction

$$f : U \to \mathbb{R},\ x \mapsto f(x) = f(x_1, \ldots, x_n).$$

Généralités

Dérivées partielles

2. Définition. La *dérivée de f en a selon v* est la limite, si elle existe,

$$\mathrm{D}_v f(a) = \lim_{t\to 0} \frac{f(a + t\,v) - f(a)}{t}.$$

3. Définition. La *i^e dérivée partielle de f en a* est la dérivée de f en a selon e_i :

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a) = \partial_i f(a) &= \lim_{t\to 0} \frac{f(a + t\,e_i) - f(a)}{t} \\ &= \lim_{t\to 0} \frac{1}{t}\big(f(a_1, \ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots, a_n) \\ &\qquad - f(a_1, \ldots, a_n)\big).\end{aligned}$$

Gradient

4. Contexte. Supposons que f admette des dérivées partielles en a par rapport à toutes les coordonnées.

5. Définition. On pose

$$\mathrm{d}f(a) \cdot h = \sum_{i=1}^{n} h_i\, \partial_i f(a),$$

En particulier, $\partial_i f(a) = \mathrm{d}f(a) \cdot e_i$. La *différentielle de f en a* est la forme linéaire définie par

$$\mathrm{d}f(a) : \mathbb{R}^n \to \mathbb{R},\ h \mapsto \mathrm{d}f(a) \cdot h.$$

6. Définition. D'après le théorème de représentation, cette forme linéaire est représentée par un unique vecteur, appellé *gradient de f en a* et noté $\nabla f(a)$:

$$\mathrm{d}f(a) \cdot h = (\nabla f(a)\,|\,h) = \nabla f(a)^\top h.$$

En particulier,

$$\nabla f(a) = \sum_{i=1}^{n} \partial_i f(a)\, e_i.$$

7. Définition. Si $\nabla f(a) = 0$, a est un *point critique* de f.

Classe $\mathscr{C}^1$

Définition

8. Définition. La fonction f est *de classe $\mathscr{C}^1$* sur U si toutes ses dérivées partielles sont définies et continues sur U. L'ensemble des fonctions de classe $\mathscr{C}^1$ sur U se note $\mathscr{C}^1(U, \mathbb{R})$.

9. Théorème. Si $f \in \mathscr{C}^1(U, \mathbb{R})$, elle est continue sur U. De plus, elle admet un *développement limité à l'ordre* 1 en tout a de U : pour h proche de 0,

$$f(a + h) = f(a) + \mathrm{d}f(a) \cdot h + \mathrm{o}(\|h\|).$$

Opérations

10. Contexte. Considérons f et g dans $\mathscr{C}^1(U, \mathbb{R})$.

11. Théorème. Pour tout $\lambda \in \mathbb{R}$, $f + \lambda g \in \mathscr{C}^1(U, \mathbb{R})$ et

$$\partial_i(f + \lambda g)(a) = \partial_i f(a) + \lambda\, \partial_i g(a).$$

12. Théorème. $f\,g \in \mathscr{C}^1(U, \mathbb{R})$ et

$$\partial_i(f\,g)(a) = \partial_i f(a)\, g(a) + f(a)\, \partial_i g(a).$$

13. Théorème. Si f ne s'annule pas sur U, $1/f \in \mathscr{C}^1(U, \mathbb{R})$ et

$$\partial_i\left(\frac{1}{f}\right)(a) = -\frac{1}{f^2(a)}\, \partial_i f(a).$$

14. Règle de la chaine. Soient un intervalle ouvert $I \subset \mathbb{R}$ et des fonctions $\varphi_1, \ldots, \varphi_n$ de $\mathscr{C}^1(I, \mathbb{R})$ telles que

$$\forall t \in I,\ (\varphi_1(t), \ldots, \varphi_n(t)) \in U.$$

L'application

$$\varphi : I \to \mathbb{R},\ t \mapsto f(\varphi_1(t), \ldots, \varphi_n(t))$$

est de classe $\mathscr{C}^1$ sur I et

$$\forall t \in I,\ \varphi'(t) = \sum_{i=1}^{n} \partial_i f(\varphi_1(t), \ldots, \varphi_n(t))\, \varphi_i'(t).$$

Classe $\mathscr{C}^2$

15. Définition. Si les dérivées partielles de f admettent elles-même des dérivées partielles, la fonction f admet des *dérivées partielles secondes ou d'ordre* 2 et on note

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

Si $i = j$, on abrège la notation :

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

16. Définition. La fonction f est *de classe* $\mathscr{C}^2$ sur U si toutes ses dérivées partielles d'ordre 2 sont définies et continues sur U. L'ensemble des fonctions de classe $\mathscr{C}^2$ sur U se note $\mathscr{C}^2(U,\mathbb{R})$.

17. Contexte. Considérons f et g dans $\mathscr{C}^2(U,\mathbb{R})$.

18. Théorème de Schwarz.

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

19. Théorème.
- Pour tout $\lambda \in \mathbb{R}$, $f + \lambda g \in \mathscr{C}^2(U,\mathbb{R})$;
- $f g \in \mathscr{C}^2(U,\mathbb{R})$;
- et si f ne s'annule pas sur U, $1/f \in \mathscr{C}^2(U,\mathbb{R})$.

20. Définition. La *(matrice) hessienne* de f en a est la matrice symétrique

$$H_f(a) = \left(\partial^2_{i,j} f(a)\right)_{1 \leqslant i,j \leqslant n}.$$

21. Théorème. f admet un *développement limité à l'ordre* 2 *en* a : pour h au voisinage de 0,

$$f(a+h) = f(a) + \nabla f(a)^\top h + \frac{1}{2!} h^\top H_f(a)\, h + \mathrm{o}(\|h\|^2).$$

Extrémums

22. Définitions.
- La fonction f admet un *maximum local en a* si
$$\exists r > 0, \forall x \in U \cap B(a,r),\ f(x) \leqslant f(a).$$
Dans ce cas, $f(a)$ s'appelle un *maximum local* de f.
- La fonction f admet un *minimum local en a* si
$$\exists r > 0, \forall x \in U \cap B(a,r),\ f(x) \geqslant f(a).$$
Dans ce cas, $f(a)$ s'appelle un *minimum local* de f.

Définitions. L'adjectif local est changé en *global* si l'on supprime la condition $x \in B(a,r)$.

On parle d'*extrémum* pour désigner indifféremment un maximum ou un minimum.

23. Théorème. Si $f \in \mathscr{C}^1(U,\mathbb{R})$ admet un extrémum local en a, alors a est critique.

24. Théorème. Soit $f \in \mathscr{C}^2(U,\mathbb{R})$. Supposons que a est un point critique.
- Si $H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$, alors $f(a)$ est un minimum local de f.
- Si $H_f(a) \notin \mathscr{S}_n^{+}(\mathbb{R})$, alors $f(a)$ n'est pas un minimum de f.