NOTES DE COURS : GRADIENTS ET DÉRIVÉES DIRECTIONNELLES

HERMAN GOULET-OUELLET

Table des matières

1. Rappels : dérivées partielles

Définition 1. Soit $f\colon \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables. La dérivée partielle de f par rapport à x_i est définie par la limite (si elle existe) :

$$\frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) = \lim_{h\to 0} \frac{f(x_1, \cdots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \cdots, x_n)}{h}$$

Pour une fonction de deux variables $f(x, y)$ on a deux dérivées partielles :

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h}, \qquad \frac{\partial f}{\partial y}(x, y) = \lim_{h\to 0} \frac{f(x, y+h) - f(x, y)}{h}.$$

🛈 Pour calculer la dérivée partielle par rapport à une variable donnée, on traite les autres variables comme des constantes.

Exemple 2. Calculons les dérivées partielles de $f(x, y) = 3\sin(x) - 2\cos(y) - xy$.

Réponse : $\frac{\partial f}{\partial x}(x, y) = 3\cos(x) - y$ et $\frac{\partial f}{\partial y}(x, y) = 2\sin(y) - x$.

🛈 La dérivée partielle par rapport à x au point $\boldsymbol{a}$ est égale à la pente de la droite tangente passant par $(\boldsymbol{a}, f(\boldsymbol{a}))$ dans la direction parrallèle à l'axe des x.

Date : 4 avril 2025.

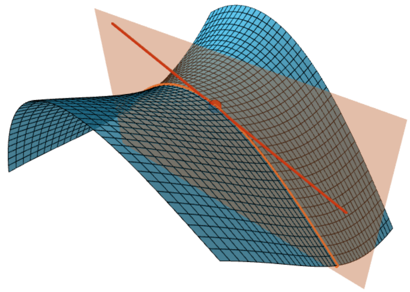

FIGURE 1. Droite tangente au graphe de $f(x, y) = 3\sin(x) - 2\cos(y) - xy$ dans la direction parallèle à l'axe des x passant par le point $(1, 1, f(1, 1))$.

2. Dérivées directionnelles

Définition 3. Soit $f\colon \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables. La dérivée directionnelle de f dans la direction $\boldsymbol{u} \in \mathbb{R}^n$, $\boldsymbol{u} \neq \boldsymbol{0}$, est définie par la limite (si elle existe)

$$\nabla_{\boldsymbol{u}} f(\boldsymbol{x}) = \lim_{h\to 0} \frac{f(\boldsymbol{x} + h\boldsymbol{u}) - f(\boldsymbol{x})}{h}, \qquad \text{où } \boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

ℹ Pour tout scalaire $k \neq 0$, on a $\nabla_{k\boldsymbol{u}} f = k\nabla_{\boldsymbol{u}} f$. On peut donc se restreindre au cas où $\boldsymbol{u}$ est un *vecteur unitaire*, c'est-à-dire dont la norme euclidienne vaut 1 :

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} = 1.$$

Dans le cas de $\mathbb{R}^2$, les vecteurs unitaires sont de la forme

$$\boldsymbol{u} = (\cos(\theta), \sin(\theta)), \quad \theta \in [0, 2\pi).$$

Pour un vecteur de cette forme, la dérivée directionnelle d'une fonction $f(x, y)$ de deux variables est, par définition,

$$\nabla_{\boldsymbol{u}} f(x, y) = \lim_{h\to 0} \frac{f(x + h\cos(\theta), y + h\sin(\theta)) - f(x, y)}{h}$$

— Quand $\theta = 0$, $\boldsymbol{u} = (1, 0)$ et on retrouve $\nabla_{\boldsymbol{u}} f = \partial f/\partial x$.

— Quand $\theta = \frac{\pi}{2}$, $\boldsymbol{u} = (0, 1)$ et on retrouve $\nabla_{\boldsymbol{u}} f = \partial f/\partial y$.

ℹ La dérivée directionnelle dans la direction $\boldsymbol{u}$ au point $\boldsymbol{a}$ est égale à la pente de la droite tangente passant par $\boldsymbol{a}$ dans la direction $\boldsymbol{u}$.

En d'autres mots, la droite tangente admet la description paramétrique :

$$D = \{(\boldsymbol{a} + t\boldsymbol{u}, f(\boldsymbol{a}) + t\nabla_{\boldsymbol{u}} f(\boldsymbol{a})) \mid t \in \mathbb{R}\}.$$

Dans le cas d'une fonction de deux variables, cette description prend la forme :

$$D = \{(a_1 + tu_1, a_2 + tu_2, f(a_1, a_2) + t\nabla_{\boldsymbol{u}} f(a_1, a_2)) \mid t \in \mathbb{R}\}.$$

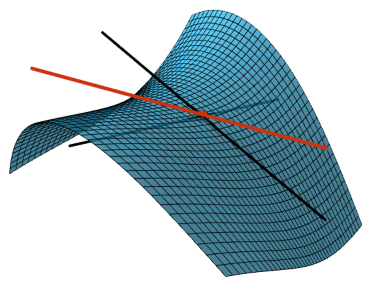

FIGURE 2. Droite tangente au graphe de $f(x) = 3\sin(x) - 2\cos(y) - xy$ dans la direction $\boldsymbol{u} = (3/5, 4/5)$ passant par le point $(1, 1, f(1, 1))$. Les droites tangentes dans la direction de l'axe des x et des y sont est indiquées en noir.

ℹ Calculer $\nabla_{\boldsymbol{u}} f(\boldsymbol{x})$ directement à partir de la définition est possible, mais difficile.

Théorème 4. Soit $f \colon \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables et $\boldsymbol{u} = (u_1, u_2) \in \mathbb{R}^2$. En tout point $(x, y) \in \mathbb{R}^2$ où f est *différentiable*,

$$\nabla_{\boldsymbol{u}} f(x, y) = u_1 \frac{\partial f}{\partial x}(x, y) + u_2 \frac{\partial f}{\partial y}(x, y).$$

(On verra la définition de *différentiable* dans un prochain cours. Pour l'instant on ne travaillera qu'avec des fonctions différentiables.)

Exemple 5. Soit la fonction $f(x, y) = 3\sin(x) - 2\cos(y) - xy$, et $\boldsymbol{u} = (\cos(\theta), \sin(\theta))$, où $\theta = \arccos(3/5)$. Calculons $\nabla_{\boldsymbol{u}} f(x, y)$ en utilsant le théorème.

Réponse : $\nabla_{\boldsymbol{u}}(x, y) = \frac{3}{5}(\cos(x) - y) + \frac{4}{5}(\sin(y) - x)$.

Corollaire 6. Soit $f \colon \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables. En tout point $(a, b) \in \mathbb{R}^2$ où f est différentiable, toutes les droites tangentes au graphe de f passant par $(a, b, f(a, b))$ sont contenues dans le plan d'équation

$$z - f(a, b) = \frac{\partial f}{\partial x}(a, b) \cdot (x - a) + \frac{\partial f}{\partial y}(a, b) \cdot (y - b)$$

appelé *plan tangent*.

Example 7. Calculons l'équation du plan tangent au graphe de $f(x, y) = -(x^2 + xy + y^2)$ passant par le point $(-1/2, 1/2, -1/4)$.

Réponse : l'équation du plan tangent est $z = x/2 - y/2 + 1/4$.

3. Gradient

Définition 8. Soit $f \colon \mathbb{R}^n \to \mathbb{R}$ une fonction réelle de n variables. Le gradient de f est la fonction

$$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f}{x_1}(\boldsymbol{x}), \ldots, \frac{\partial f}{x_n}(\boldsymbol{x}) \right), \qquad \text{où } \boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n.$$

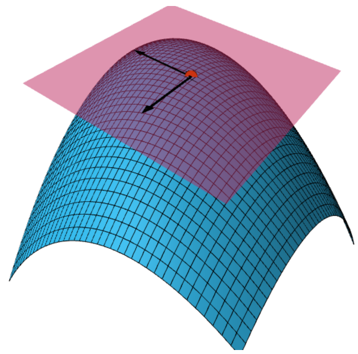

FIGURE 3. Plan tangent au graphe de $f(x, y) = -(x^2 + xy + y^2)$ passant par le point $(-1/2, 1/2, -1/4)$.

Autrement dit, $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$.

Exemple 9. Soit la fonction $f(x, y, z) = x^2 - xyz - z^2$. Calculons le gradient $\nabla f(x, y, z)$.

Réponse : $\nabla f(x, y, z) = (2x - yz, -xz, -2z - xy)$.

Définition 10. Étant donné deux vecteurs $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, on définit leur produit scalaire par

$$\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{i=1}^{n} u_i v_i.$$

Proposition 11. (Propriétés du produit scalaire.)

(1) $\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \cos(\theta)$, où θ est l'angle entre $\boldsymbol{u}$ et $\boldsymbol{v}$.

(2) $|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\|$ (inégalité de Cauchy–Schwarz).

(3) $\boldsymbol{u} \cdot \boldsymbol{v} = 0 \iff \boldsymbol{u} \perp \boldsymbol{v}$ ($\boldsymbol{u}$ et $\boldsymbol{v}$ sont perpendiculaires).

(4) $\boldsymbol{u} \cdot \boldsymbol{u} = \|\boldsymbol{u}\|^2$.

Théorème 12. (Formule du gradient.) Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction réelle de n variables et $\boldsymbol{u} \in \mathbb{R}^n$. En tout point $\boldsymbol{a} \in \mathbb{R}^n$ où f est *différentiable*,

$$\nabla_{\boldsymbol{u}} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{u}.$$

— Pour $n = 2$ on retrouve la formule énoncée précédemment.

— La preuve utilise la règle de *dérivation en chaîne* multivariée (prochain cours).

Exemple 13. Calculons le gradient $\nabla f(\boldsymbol{x})$ et la dérivée directionnelle $\nabla_{\boldsymbol{u}} f(\boldsymbol{x})$ pour les exemples suivants :

a) $f(x, y) = x^2 - xy + 3y^2, \quad \boldsymbol{u} = (\frac{3}{5}, \frac{4}{5}), \quad \boldsymbol{a} = (\frac{5}{2}, \frac{5}{7})$.
Réponse : $\nabla f(x, y) = (2x - y, -x + 6y)$ et $\nabla_{\boldsymbol{u}} f(\boldsymbol{a}) = 4$.

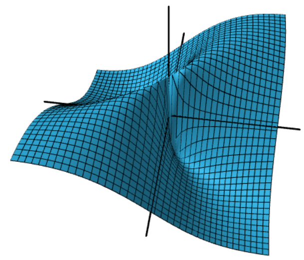

FIGURE 4. Graphe de la fonction $f(x, y) = xy/(x^2 + y^2)$.

b) $\sin(3x)\cos(3y)$, $\quad \boldsymbol{u} = (\frac{5}{13}, \frac{12}{13})$, $\quad \boldsymbol{a} = (\pi, \frac{\pi}{2})$.
Réponse : $\nabla f(x, y) = (3\cos(3x)\cos(3y), -3\sin(3x)\sin(3y))$ et $\nabla_{\boldsymbol{u}} f(\boldsymbol{a}) = 0$.

c) $f(x_1, x_2, x_3) = x_1x_3 + x_1^2 - x_2 + x_2x_3 + x_2^2 + 3x_3^2$, $\quad \boldsymbol{u} = (\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2})$, $\quad \boldsymbol{a} = (1, -2, \frac{3}{2})$.
Réponse : $\nabla f(x_1, x_2, x_3) = (2x_1 + x_3, 2x_2 + x_3 - 1, x_1 + x_2 + 6x_3)$ et $\nabla_{\boldsymbol{u}} f(\boldsymbol{a}) = 4\sqrt{3}$.

⚠ La formule du gradient ne fonctionne pas si la fonction n'est pas différentiable.

Exemple 14. Soit $f(x, y) = xy/(x^2 + y^2)$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$. Le gradient ∇f existe partout, y compris en (0,0). Cependant f n'est pas différentiable en (0,0) et les dérivées directionnelles $\nabla_{\boldsymbol{u}} f(0, 0)$ n'existent pas pour $\boldsymbol{u} \notin \{(0, 1), (1, 0)\}$.

Corollaire 15. En tout point $\boldsymbol{a} \in \mathbb{R}^n$ où f est différentiable :

(1) Le gradient $\nabla f(\boldsymbol{a})$ donne la direction qui *maximise la croissance instantanée.*

(2) Dans les directions perpendiculaires à $\nabla f(\boldsymbol{a})$, *la croissance instantané est nulle.*

Ce corollaire est très utile pour résoudre des problèmes d'optimisation. Il est à la base de l'*algorithme du gradient*, dont la variante appelée *algorithme du gradient stochastique* est utilisée en apprentissage automatique.

Soit $f(x, y) = 4 - (\cos^2(x/2) + \cos^2(y/2))^2$. La figure 5 illustre f et ∇f (représenté sur le plan $z = 0$).

4. POINTS CRITIQUES

Rappelons que les points critiques d'une fonction $f : \mathbb{R} \to \mathbb{R}$ d'une seule variable sont les points $x \in \mathbb{R}$ tels que $f'(x) = 0$.

Si la fonction est deux fois différentiable alors les points critiques peuvent prendre trois différentes formes :

(1) maximum local, quand $f''(x) < 0$.

(2) minimum local, quand $f''(x) > 0$.

(3) point d'inflexion, quand $f''(x) = 0$.

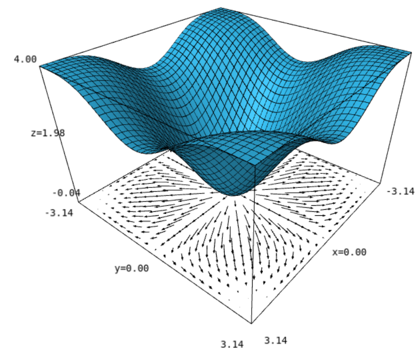

FIGURE 5. Graphe et gradient de la fonction $f(x, y) = 4 - (\cos^2(x/2) + \cos^2(y/2))^2$.

Définition 16. Soit $f\colon \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables. On appelle *point critique* un point $\boldsymbol{x} \in \mathbb{R}^n$ tel que $\nabla f(\boldsymbol{x}) = 0$.

Pour trouver les points critiques d'une fonction de n variables, il faut donc résoudre un système à n équations et n inconnues :

$$\begin{cases} \frac{\partial f}{\partial x_1}(x_1, x_2, \ldots, x_n) & = & 0 \\ \frac{\partial f}{\partial x_2}(x_1, x_2, \ldots, x_n) & = & 0 \\ \vdots & & \\ \frac{\partial f}{\partial x_n}(x_1, x_2, \ldots, x_n) & = & 0 \end{cases}$$

Exemple 17. Trouvons les points critiques de la fonction $f(x_1, x_2, x_3) = x_1x_3 + x_1^2 - x_2 + x_2x_3 + x_2^2 + 3x_3^2$.

Réponse : On a $\nabla f(x_1, x_2, x_3) = (x_3 + 2x_1,\ -1 + x_3 + 2x_2,\ x_1 + x_2 + 6x_3)$. Donc les points critiques sont les solutions du système

$$\begin{cases} 2x_1 & & & + & x_3 & = 0 \\ & & 2x_2 & + & x_3 & = 1 \\ x_1 & + & x_2 & + & 6x_3 & = 0. \end{cases}$$

En utilisant notre trousse à outils d'algèbre linéaire, on trouve un seul point critique, soit $\left(\frac{1}{20}, \frac{11}{20}, \frac{-1}{10}\right)$.