Sur la règle de dérivation en chaîne

Le résultat théorique
Soient $f : \mathbb{R}^n \to \mathbb{R}$ et $g : \mathbb{R}^p \to \mathbb{R}^n$ deux fonctions différentiables. Écrivons $h = f \circ g$. D'après la règle de dérivation des fonctions composées nous avons (comme pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$) :

$$h'(x) = (f \circ g)'(x) = f'(g(x)).g'(x).$$

La fonction $f \circ g$ est une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$. Sa dérivée est donc un vecteur ligne à p colonnes, la transposée de son gradient :

$$h'(x) = \begin{pmatrix} \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} & \cdots & \frac{\partial h}{\partial x_p} \end{pmatrix}.$$

La fonction g est une fonction de $\mathbb{R}^p$ dans $\mathbb{R}^n$. Sa dérivée est la matrice $n \times p$ composée des vecteurs transposés des gradients des coordonnées de g. Si $g(x) = (g_1(x), g_2(x), \ldots, g_2(x))$ (on devrait écrire ce vecteur en colonne si on voulait se conformer en toute rigueur aux choix du cours) la dérivée de g s'écrit :

$$g'(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_p} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_p} \end{pmatrix}.$$

Pour simplifier la présentation appelons $g = (g_1, g_2, \ldots, g_n)$ un point de $\mathbb{R}^n$. C'est un abus de notation, g ne désigne pas ici la fonction g mais un vecteur, un point dans $\mathbb{R}^n$. La dérivée de f en un point g est donnée par la transposée de son gradient :

$$f'(g) = \begin{pmatrix} \frac{\partial f}{\partial g_1} & \frac{\partial f}{\partial g_2} & \cdots & \frac{\partial f}{\partial g_n} \end{pmatrix}.$$

L'égalité matricielle $h'(x) = (f \circ g)'(x) = f'(g(x)).g'(x)$ signifie donc :

$$\begin{pmatrix} \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} & \cdots & \frac{\partial h}{\partial x_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial g_1} & \frac{\partial f}{\partial g_2} & \cdots & \frac{\partial f}{\partial g_n} \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_p} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_p} \end{pmatrix}.$$

Autrement dit pour tout $i = 1, \ldots, p$ on a

$$\frac{\partial h}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial f}{\partial g_k} \frac{\partial g_k}{\partial x_i}.$$

Attention ! Quand g_k apparaît au dénominateur cela signifie seulement que l'on prend la dérivée de f par rapport à sa $k$ième variable. Quand il apparaît au numérateur g_k désigne la $k$ième coordonnée de g : c'est alors une fonction.

Un exemple

Prenons $f : \mathbb{R}^3 \to \mathbb{R}$ et $g : \mathbb{R}^2 \to \mathbb{R}^3$ deux fonctions différentiables définies par

$$f(x, y, z) = 2xy - 3(x + z),$$

$$g(x, y) = (x + y^4, y - 3x^2, 2x^2 - 3y).$$

On demande de calculer les dérivées partielles de la fonction de deux variables $h = f \circ g$. Pour se ramener au théorème général et ne pas s'embrouiller, il est préférable de changer les noms des variables dans l'expression de f :

$$f(g_1, g_2, g_3) = 2g_1g_2 - 3(g_1 + g_3).$$

La formule de dérivation en chaîne donne alors

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial g_1}\frac{\partial(x + y^4)}{\partial x} + \frac{\partial f}{\partial g_2}\frac{\partial(y - 3x^2)}{\partial x} + \frac{\partial f}{\partial g_3}\frac{\partial(2x^2 - 3y)}{\partial x},$$

$$\frac{\partial h}{\partial y} = \frac{\partial f}{\partial g_1}\frac{\partial(x + y^4)}{\partial y} + \frac{\partial f}{\partial g_2}\frac{\partial(y - 3x^2)}{\partial y} + \frac{\partial f}{\partial g_3}\frac{\partial(2x^2 - 3y)}{\partial y}.$$

Pour $\frac{\partial h}{\partial x}$, on obtient :

$$\frac{\partial h}{\partial x} = (2g_2 - 3).1 + 2g_1.(-6x) + (-3).4x$$

Exprimée en fonction de x et y cette dérivée s'écrit :

$$\frac{\partial h}{\partial x} = 2y - 6x^2 - 3 - 12x(x + y^4) - 12x = -12xy^4 - 18x^2 + 2y - 12x - 3.$$

Je vous laisse le calcul de la deuxième dérivée partielle de h en exercice.
Remarque. On peut aussi écrire les choses sous la forme :

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x}\frac{\partial(x + y^4)}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial(y - 3x^2)}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial(2x^2 - 3y)}{\partial x},$$

mais c'est un peu risqué. Il ne faut surtout pas oublier de prendre les valeurs des dérivées partielles de f au point $(x + y^4, y - 3x^2, 2x^2 - 3y)$.