

Calcul différentiel : fonction à plusieurs variables

Florent Capelli

19/10/2026

1 Fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$

On s'est concentré la semaine dernière sur les fonctions à une variables. On peut généraliser la plupart des notions vues dans ce cadre aux fonctions ayant plus d'une variables, c'est-à-dire aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. Pour cela, il faut cependant travailler dans un espace normé. On va donc se placer dans l'espace Euclidien $\mathbb{R}^n$ muni de son produit scalaire canonique et donc de la norme Euclidienne.

1.1 Limites et continuité

On peut dès lors parler de limites et de continuité d'une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$. En effet, on va dire que la limite de f en a est $\ell \in \mathbb{R}$ si si pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que pour tout $b \in \mathbb{R}^n$, si $\|b - a\|_2 < \eta$ alors $|f(b) - \ell| < \epsilon$. On comprend facilement l'analogie : dès lors que b s'approche de a , on a $f(b)$ qui s'approche de ℓ . On dit alors que f est continue si la limite de f en a vaut $f(a)$.

Nous pouvons formuler cela parce que nous avons une notion de distance. Sans cela, la notion de continuité serait plus difficile à définir (mais pas impossible, on a simplement besoin d'une notion de limite, ce que l'on peut avoir dans des espaces qui ne sont pas nécessairement métriques comme les espaces topologiques).

Exemple 1.1. Quelques exemples de fonctions continues : - La fonction $(x, y) \mapsto x + y$ est continue en tout point de $\mathbb{R}^2$. - La fonction $(x, y) \mapsto xy$ est continue en tout point de $\mathbb{R}^2$. - La fonction $(x, y) \mapsto x^2 + y^2$ est continue en tout point de $\mathbb{R}^2$. - La fonction $(x, y, z) \mapsto xy + e^{yz}$ est continue en tout point de $\mathbb{R}^3$.

Comme dans le cas à une variable, le produit, la somme et la composition de fonctions continues est aussi continu. Une fois la composition établie d'ailleurs, on peut retrouver la somme et la multiplication en remarquant simplement la continuité de $(x, y) \mapsto x + y$ et $(x, y) \mapsto xy$.

Théorème 1.1. Si $f_1, \dots, f_k$ sont des fonctions continues de $\mathbb{R}^n$ et f une fonction continue de $\mathbb{R}^k$ alors $f(f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))$ est continue.

1.2 Différentiabilité

On peut de même généraliser la notion de dérivée à des fonctions de plusieurs variables. L'intuition est la même : si on regarde une fonction de $\mathbb{R}^n \rightarrow \mathbb{R}$ autour d'un point où elle est différentiable, alors elle "ressemble" à une forme linéaire de $\mathbb{R}^n$. On peut l'imaginer dans $\mathbb{R}^3$: si on s'approche très près d'une surface lisse, elle a l'air plane (breaking news: la terre n'est pas plate !). La différentielle Df de f en un point x_0 est alors la forme linéaire correspondante.

On va formaliser cette intuition et montrer que ce n'est pas si simple et que des choses inattendues apparaissent en dimension $n > 1$. Déjà, on peut remarquer que la formule $\frac{f(a+h)-f(a)}{h}$ n'a pas de sens en dimension $n > 1$ car on ne peut pas diviser par h . On ne peut donc pas directement généraliser la notion de dérivabilité en appliquant la même formule et en changeant juste la notion de limite.

On va donc commencer par s'intéresser à la dérivée en un point a en suivant un vecteur h :

Définition 1.2. La dérivée $D_h f(a)$ de f en a en suivant un vecteur $h \in \mathbb{R}^n \setminus \{0\}$ est définie par la dérivée en 0 de la fonction $t \mapsto f(a + th)$.

Exemple 1.2. Cette notion n'implique pas de régularité particulière de f en a . Par exemple, la fonction

$$f(x, y) = \begin{cases} \frac{y^2}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases}$$

f est dérivable en $(0, 0)$ en suivant n'importe quel vecteur $h = (u, v) \in \mathbb{R}^2$. En effet, si $u \neq 0$, $f(0 + t(u, v)) = \frac{t^2 v^2}{tu} = t \frac{v^2}{u}$. Donc $D_{(u,v)} f(0, 0) = \frac{v^2}{u}$. Si $u = 0$, $f(t(u, v)) = 0$ donc $D_{(0,0)} f(0, 0) = 0$. Mais f n'est pas continue en $(0, 0)$! En effet $f(\frac{1}{n^2}, \frac{1}{n}) = 1$ pour tout $n \in \mathbb{N}$ et ne tend pas vers $f(0, 0) = 0$ quand $n \rightarrow \infty$.

Maintenant si on munit $\mathbb{R}^n$ de sa base canonique $e_1, \dots, e_n$, on notera $\frac{\partial f}{\partial x_i}(a)$ la dérivée de f en a en suivant le vecteur e_i (on peut généraliser cette notion à n'importe quelle base mais nous ne rentrerons pas dans les détails ici). Plus simplement, $\frac{\partial f}{\partial x_i}$ est la dérivée de l'application $x_i \mapsto f(a_1, \dots, x_i, \dots, a_n)$ en a_j .

Exemple 1.3. On définit $f(x, y) = x^2 + y^2 + 2xy$. On a $\frac{\partial f}{\partial x} = 2x + 2$ et $\frac{\partial f}{\partial y} = 2y + 2$.

Le problème des dérivées en suivant une direction est qu'elles ne donnent pas d'informations locales sur la fonction puisque ces dérivées peuvent être très différentes en fonction de la direction par laquelle on arrive. On va donc essayer de généraliser la définition suivante de la dérivée :

$$f(a + h) = f(a) + f'(a)h + o(h)$$

On généralise d'abord $o(h)$. Une fonction $u: \mathbb{R}^n \rightarrow \mathbb{R}$ est négligeable devant $\|h\|_2$, notée $o(\|h\|_2)$ en 0 si pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que si $\|h\|_2 < \eta$ alors $\|u(h)\|_2 \leq \epsilon \|h\|_2$. Autrement dit, la fonction $h \mapsto \|u(h)\|_2$ est un $o(\|h\|_2)$.

Définition 1.3. Une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable en a s'il existe une application **linéaire** u telle que $f(a+h) - f(a) - u(h) = o(\|h\|_2)$. Dans ce cas, u est unique et on la note $df(a)$.

Démonstration. On prouve l'unicité de u . Supposons que $f(a+h) - f(a) - u_1(h) = o(\|h\|_2)$ et $f(a+h) - f(a) - u_2(h) = o(\|h\|_2)$. Alors $v = u_1 - u_2$ est une application linéaire. On peut facilement montrer que $v = o(\|h\|_2)$. De plus, soit $h \in \mathbb{R}^n$. Pour tout $\epsilon > 0$ et t suffisamment petit, on a $\|v(th)\|_2 \leq \epsilon t \|h\|_2$. Or, le terme de gauche se réécrit en $t\|v(h)\|_2$. Donc $\|v(h)\|_2 \leq \epsilon \|h\|_2$ pour tout ϵ . Donc $v(h) = 0$. Donc $v = 0$ puisque l'égalité précédente est valable pour tout h . $\square$

- La différentielle d'une fonction linéaire est la fonction elle-même.
- La différentielle de $x, y \mapsto x^2 + y^2 + 2x + 2y$ en 0 est $(x, y) \mapsto 2x + 2y$.
- Si $f: \mathbb{R} \rightarrow \mathbb{R}$ et $a \in \mathbb{R}$, alors $df(a)(h) = f'(a)h$.

Supposons que f est différentiable en a . Alors on peut voir que pour $h \in \mathbb{R}^n$, $D_h f(a) = df(a)(h)$. En effet, $f(a+th) = f(a) + df(a)(th) + o(\|th\|_2)$. Donc $\frac{f(a+th) - f(a)}{t} = df(a)(h) + \frac{o(\|th\|_2)}{t} \xrightarrow{t \rightarrow 0} df(a)(h)$.

L'application linéaire $df(a)$ est une application de $\mathbb{R}^n$ dans $\mathbb{R}$ (autrement appelé une forme linéaire). On peut donc complètement la caractériser par ses valeurs $df(a)(e_1), \dots, df(a)(e_n)$ dans la base canonique de $\mathbb{R}^n$. Or, par ce qui précède, ces valeurs sont $\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)$. On reconnaît la transposée du gradient de f !

Définition 1.4. Le gradient de f , noté $\nabla f(a)$, est le vecteur colonne défini par $\begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}$. Si f est différentiable en a , $\nabla f(a) = df(a)^T$.

Le gradient est cependant souvent défini comme la direction dans l'espace où la fonction change le plus alors que la différentielle modélise le comportement d'une fonction autour d'un point. On a :

$$f(a+h) = f(a) + (\nabla f(a)|h) + o(\|h\|)$$

Exemple 1.4. La fonction suivante

$$f(x, y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

n'est pas différentiable en $(0, 0)$ mais elle admet un gradient.

Pour calculer une différentielle, le plus simple est souvent de calculer son gradient coordonnées par coordonnées en appliquant les règles de dérivation vues au cours précédent. Cependant, vu l'exemple précédent, on peut se demander légitimement si ce qu'on a calculé est bien la différentielle ou juste un gradient. On peut souvent utiliser la notion de continuellement différentiable.

Définition 1.5. Soit $a \in \mathbb{R}^n$, $\epsilon > 0$ et f une fonction différentiable sur $B(a, \epsilon) = \{b \mid \|b - a\|_2 < \epsilon\}$. On dit que f est continuellement différentiable sur $B(a, \epsilon)$ si $b \mapsto df(b)$ est une fonction continue sur $B(a, \epsilon)$.

On peut caractériser le fait qu'une fonction est continuellement différentiable par l'existence d'un gradient *continu* :

Théorème 1.6. Une fonction f est continuellement différentiable sur $B(a, \epsilon)$ si et seulement si ∇f est continue sur $B(a, \epsilon)$.

Ainsi, si en calculant le gradient de $\nabla f(a)$ (vu comme une fonction de a), on obtient une fonction continue autour de a , on sait que f est différentiable en a et donc que $\nabla f(a) = df(a)$.

$\frac{f(x,y)=x^2+y^2}{2x}$ est différentiable en tout point (x, y) de $\mathbb{R}_+$ et $df(x, y) = (1 - \frac{x^2+y^2}{x^2}, \frac{y}{x})$. En effet, l'existence des dérivées partielles s'obtient par un calcul simple. On vérifie facilement qu'elles sont continues sur $\mathbb{R}_+$ (comme composition de fonctions continues) et donc on en déduit, en appliquant le théorème précédent, que la transposée du gradient nous donne la différentielle.

On retrouve, comme pour la dérivée des règles formelles pour calculer les différentielles d'un produit, d'une somme ou d'une composée. On va s'y intéresser dans la section suivante, mais avant cela, il est plus simple de généraliser ce que l'on vient de voir aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^m$.

2 Fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^m$

On considère désormais des fonctions de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$. Si $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, alors on peut voir chacune de ses composantes comme une fonction f_i de $\mathbb{R}^n$ dans $\mathbb{R}^m$, c'est-à-dire $f = (f_1, \dots, f_m)$.

On généralise facilement de limite puisqu'on a une notion de normes sur $\mathbb{R}^m$: la limite de $f(x)$ quand x tend vers $a \in \mathbb{R}^n$ est $\ell \in \mathbb{R}^m$ si et seulement pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que pour tout $b \in \mathbb{R}^n$, si $\|b - a\|_2 < \eta$, alors $\|f(b) - \ell\|_2 < \epsilon$.

On définit la continuité comme d'habitude : une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continue en a si la limite de f en a est $f(a)$.

On utilise la notion de différentiabilité comme vu au-dessus : f est différentiable en a si il existe une fonction linéaire $df(a)$ telle que $f(a + h) - f(a) - df(a)(h) = o(\|h\|_2)$. On peut reprendre la preuve du cas $m = 1$ pour montrer que si une telle fonction linéaire existe alors elle est unique et c'est pour cela que nous

avons pris la liberté de la noter $df(a)$ directement. On appelle alors $df(a)$ la différentielle de f en a .

Comme $df(a)$ est une fonction linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, on peut l'écrire comme une matrice $m \times n$ dans la base canonique. De plus, en écrivant f comme $(f_1, \dots, f_m)$, on voit facilement que si $df(a)$ existe, sa i ème composante doit être $df_i(a)$. Donc $df(a)$ peut s'écrire (dans la base canonique de $\mathbb{R}^n$ dans $\mathbb{R}^m$) comme la matrice dont les lignes sont $df_1(a), \dots, df_m(a)$. On appelle cette matrice **la matrice jacobienne de f** et on la note $\nabla f(a)$.

On peut écrire $\nabla f(a)$ comme
$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Comme précédemment, l'existence d'une matrice Jacobienne ne garantit pas l'existence d'une différentielle. Cependant, le théorème précédent est toujours vrai : la continuité de la matrice jacobienne est équivalente au fait que f soit continuellement différentiable.

Dans le cas de fonctions continuellement différentiables, on peut alors calculer les jacobiens de façon systématique :

- Si f, g sont continuellement différentiables en a , alors $\nabla(f+g)(a) = \nabla f(a) + \nabla g(a)$.
- Si $f: \mathbb{R}^k \rightarrow \mathbb{R}^m$ est continuellement différentiable en $g(a)$ et $g: \mathbb{R}^n \rightarrow \mathbb{R}^k$ est continuellement différentiable en a , on a $\nabla(f \circ g)(a) = \nabla f(g(a)) \times \nabla g(a)$ où pour tout $x \in \mathbb{R}^n$, $(f \circ g)(x) = f(g_1(x), \dots, g_k(x))$.
- Si $f: \mathbb{R}^n \rightarrow \mathbb{R}^k$ est une fonction linéaire $x \mapsto Ax$, alors $\nabla f = A$.

Exemple 2.1. 1. Soit $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par
$$\begin{pmatrix} 2x + 3y \\ x + y \\ x - y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

On a $\nabla f(x, y) = \begin{pmatrix} 2 & 3 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$ pour tout $x, y \in \mathbb{R}^2$.

2. Soit W une matrice $m \times n$ et f une fonction de $\mathbb{R}^m$ dans $\mathbb{R}^k$. On définit $g(X) = f(WX)$. On a $\nabla g(X) = W \times \nabla f(WX)$.
3. Pour $h: \mathbb{R} \rightarrow \mathbb{R}$ et $q: \mathbb{R}^n \rightarrow \mathbb{R}^k$, on note $h \circ q = (h \circ q_1, \dots, h \circ q_k)$, c'est-à-dire qu'on applique h à chaque composante de $q = (q_1, \dots, q_k)$. On définit $g(X) = h(WX)$. On a $\nabla g = Wh' \circ (WX)$