Calcul Différentiel I

STEP, MINES ParisTech*

8 octobre 2020 (`#bd0f014`)

Table des matières

*Ce document est un des produits du projet `boisgera/CDIS`, initié par la collaboration de (S)ébastien Boisgérault (CAOR), (T)homas Romary et (E)milie Chautru (GEOSCIENCES), (P)auline Bernard (CAS), avec la contribution de Gabriel Stoltz (Ecole des Ponts ParisTech, CERMICS).

Objectifs d'apprentissage

Cette section s'efforce d'expliciter et de hiérarchiser les acquis d'apprentissages associés au chapitre. Ces objectifs sont organisés en paliers :

(∘) Prérequis (•) Fondamental (••) Standard (•••) Avancé (••••) Expert

Sauf mention particulière, les objectifs "Expert", les démonstrations du document [1] et les contenus en annexe ne sont pas exigibles ("hors-programme").

Prérequis Les éléments du calcul différentiel supposés déjà maîtrisés :

— ∘ dérivabilité et dérivée de fonctions d'une variable réelle,
— ∘ dérivée sur des sous-ensembles ouverts et des intervalles fermés de $\mathbb{R}$,
— ∘ dérivées des fonctions scalaires ($\mathbb{R}$) et vectorielles ($\mathbb{R}^m$),
— ∘ dérivée et développement limité au premier ordre,

Les fondements du calcul matriciel supposés maîtrisés :

— ∘ vecteurs de $\mathbb{R}^n$ et applications linéaires $\mathbb{R}^n \to \mathbb{R}^m$,
— ∘ matrices, vecteurs lignes et colonnes,
— ∘ produit matriciel, transposition de matrice,
— ∘ interprétation géométrique du calcul matriciel.

Matrice jacobienne et différentielle La notion de dérivée n'est applicable que pour les fonctions d'une variable, scalaires ou vectorielles. L'objet généralisant la dérivée dans le cas multivariable, défini au moyen des dérivées partielles, est la matrice jacobienne ; pour les fonctions scalaires, on préférera souvent utiliser le gradient que la matrice jacobienne.

— • savoir calculer dérivées partielles, matrices jacobiennes et gradients.

Toutefois, la seule existence de la matrice jacobienne est insuffisante pour exploiter la plupart des résultats du calcul différentiel. Pour cette raison, on exige souvent que cette matrice existe et dépende continûment de son argument ; c'est la "continue différentiabilité". Avec la continue différentiabilité, même sans savoir ce que signifie le terme, on peut alors néanmoins exploiter tous les résultats qui nécessitent la simple "différentiabilité".

— • savoir que l'existence de la matrice jacobienne est souvent insuffisante,
— • savoir qu'on peut alors avoir recours à la continue différentiabilité,
— • savoir caractériser les fonctions continûment différentiables,
— • savoir exploiter que continûment différentiable implique différentiable.

La différentiabilité n'est autre que l'existence d'un développement limité au premier ordre. C'est la généralisation de la notion de dérivabilité au cas multivariable.

— •• savoir que différentiabilité signifie existence d'un tel développement limité,

1. l'étude des démonstrations du cours peut toutefois contribuer à votre apprentissage, au même titre que la résolution d'exercices.

— •• savoir que différentiabilité équivaut à dérivabilité dans le cas monovariable,
— •• savoir caractériser et exploiter un développement limité au premier ordre,
— •• savoir ce qu'est la différentielle et son lien avec la matrice jacobienne.

Calcul différentiel La définition de la différentielle et son lien avec la matrice jacobienne et les dérivées partielles permettent le calcul des différentielles des fonctions élémentaires, fournissant autant de règles élémentaires de calcul. Il est donc nécessaire avant toute chose de

— • connaître et savoir mettre en œuvre quelques règles élémentaires,
— •• savoir en élaborer de nouvelles en exploitant les définitions.

Ensuite, pour faire face à des calculs plus complexes, deux stratégies sont à mener en parallèle. Au niveau pratique, il convient de

— • savoir exploiter les notations simplifiant le calcul différentiel,
— •• savoir expliciter ce qu'elles signifient.

Au niveau théorique, il s'agit de compléter les règles élémentaires par des règles génériques, applicables non plus à une fonction particulière mais à des classes de fonctions, donc de

— • connaître et savoir mettre en œuvre quelques règles de calcul génériques,
— •• connaître les règles d'assemblage/désassemblage,
— •• connaître la règle de dérivation en chaîne,
— ••• savoir élaborer de nouvelles règles de calcul génériques.

Variation des fonctions En intégrant les variations infinitésimales d'une fonction entre deux points, on peut évaluer sa variation entre ces points. Utiliser pleinement cette technique suppose de :

— • connaître le théorème fondamental du calcul (monovariable),
— •• connaître le théorème fondamental du calcul (multivariable),
— •• connaître la démonstration du cas multivariable,
— ••• savoir l'adapter quand c'est nécessaire,
— • connaître l'inégalité des accroissements finis (monovariable),
— •• connaître l'inégalité des accroissements finis (multivariable),
— •• savoir les déduire du théorème fondamental du calcul [2],
— •••• connaître la variante non-euclidienne de ces inégalités,
— ••• savoir exploiter ces résultats dans des contextes variés.

2. sous une hypothèse renforcée, par exemple de continue différentiabilité.

Conventions

Un vecteur $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ sera implicitement identifié, dans le contexte d'un calcul matriciel, au vecteur colonne

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}.$$

Une application linéaire $A : \mathbb{R}^m \to \mathbb{R}^n$, sera quant à elle implicitement identifiée avec la matrice à m lignes et n colonnes représentant l'application linéaire dans les bases canoniques de $\mathbb{R}^m$ et $\mathbb{R}^n$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Dans ce document, nous utiliserons le point "$\cdot$" pour désigner le produit entre matrices. Les identifications que nous venons d'exposer nous incitent donc également à noter $A \cdot x$ l'image du vecteur x par l'application linéaire A et $A \cdot B$ la composition des applications linéaires B par A.

Sauf mention contraire, $\langle x, y \rangle$ désignera le produit scalaire usuel entre les vecteurs x et y de $\mathbb{R}^n$ et $\|x\| = \|x\|_2 = \sqrt{\langle x, x \rangle}$ la norme euclidienne associée ; $\|A\|$ désignera la norme d'opérateur $\|A\|_{22}$ de A, induite par ces normes euclidiennes sur $\mathbb{R}^m$ et sur $\mathbb{R}^n$.

Matrice jacobienne et différentielle

Définition – Dérivées partielles

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et $x = (x_1, \cdots, x_n) \in U$. Lorsque la j-ème fonction partielle de f en x, $y \mapsto f(x_1, \cdots, x_{j-1}, y, x_{j+1}, \cdots, x_n)$, est dérivable en $y = x_j$, on appelle j-ème *dérivée partielle de f en x* et on note $\partial_j f(x)$ sa dérivée :

$$\partial_j f(x) := (y \mapsto f(x_1, \cdots, x_{j-1}, y, x_{j+1}, \cdots, x_n))' (x_j) \in \mathbb{R}^m$$

Exercice – Domaine de définition des fonctions partielles (•) Quel est le domaine de définition – implicite dans l'énoncé ci-dessus – de la j-ème fonction partielle de f associée au point x ? Montrer qu'il s'agit bien d'un ouvert de $\mathbb{R}$. (Solution p. 33.)

Définition – Matrice jacobienne

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et x un point de U. Si toutes les dérivées partielles de toutes les composantes f_i de f existent en x, on définit la *matrice* *jacobienne de f en x*, notée $J_f(x)$, comme la matrice de $\mathbb{R}^{m\times n}$ telle que

$$[J_f(x)]_{ij} = \partial_j f_i(x),$$

c'est-à-dire

$$J_f(x) = \begin{bmatrix} \partial_1 f_1(x) & \partial_2 f_1(x) & \cdots & \partial_n f_1(x) \\ \partial_1 f_2(x) & \partial_2 f_2(x) & \cdots & \partial_n f_2(x) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 f_m(x) & \partial_2 f_m(x) & \cdots & \partial_n f_m(x) \end{bmatrix}$$

On peut également utiliser les dérivées partielles de la fonction vectorielle f plutôt que ses composantes f_i, auquel cas on définit la matrice jacobienne de f comme une concaténation de vecteurs colonnes :

$$J_f(x) = \begin{bmatrix} | & | & \cdots & | \\ \partial_1 f(x) & \partial_2 f(x) & \cdots & \partial_n f(x) \\ | & | & \cdots & | \end{bmatrix}$$

Définition – Gradient

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et x un point de U. Si toutes les dérivées partielles de f existent en x, on appelle *gradient de f en x* et l'on note $\nabla f(x)$ le vecteur de $\mathbb{R}^n$ défini par

$$\nabla f(x) = (\partial_1 f(x), \partial_2 f(x), \dots, \partial_n f(x)) \in \mathbb{R}^n,$$

c'est-à-dire, après identification à un vecteur colonne, la transposée de la matrice jacobienne de f en x :

$$\nabla f(x) = J_f(x)^\top = \begin{bmatrix} \partial_1 f(x) \\ \partial_2 f(x) \\ \vdots \\ \partial_n f(x) \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

Exercice – Gradient et matrice jacobienne (•) Montrer qu'en tout point $x = (x_1, x_2)$ de $\mathbb{R}^2$, le gradient de la fonction scalaire

$$f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_2^2 - x_1)^2 + (x_1 - 1)^2 \in \mathbb{R}$$

est défini et vérifie

$$\nabla f(x_1, x_2) = (-2(x_2^2 - x_1) + 2(x_1 - 1), 4(x_2^2 - x_1)x_2) \in \mathbb{R}^2$$

c'est-à-dire, comme vecteur colonne,

$$\nabla f(x_1, x_2) = J_f(x_1, x_2)^\top = \begin{bmatrix} -2(x_2^2 - x_1) + 2(x_1 - 1) \\ 4(x_2^2 - x_1)x_2 \end{bmatrix} \in \mathbb{R}^{2\times 1}.$$

(Solution p. 34.)

Exercice – Matrice jacobienne (•) Montrer qu'en tout point $x = (x_1, x_2)$ de $\mathbb{R}^2$, la matrice jacobienne de la fonction

$$g : (x_1, x_2) \in \mathbb{R}^2 \mapsto (-2(x_2^2 - x_1) + 2(x_1 - 1), 4(x_2^2 - x_1)x_2) \in \mathbb{R}^2.$$

est définie et vérifie

$$J_g(x_1, x_2) = \left[\begin{array}{cc} 4 & -4x_2 \\ -4x_2 & 12x_2^2 \end{array} \right] \in \mathbb{R}^{2\times 2}.$$

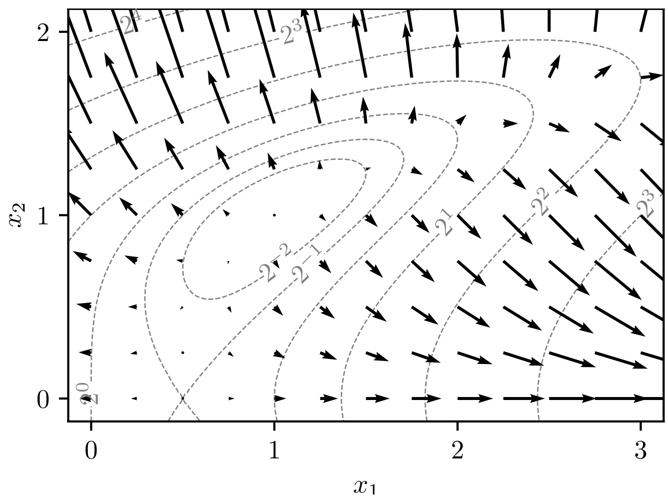

FIGURE 1 – Champ du gradient ∇f de la fonction $f : (x_1, x_2) \mapsto (x_2^2 - x_1)^2 + (x_1 - 1)^2$ (échelle des vecteurs modifiée) et en pointillés les lignes de niveau de f de la forme $\{(x_1, x_2) | \ f(x_1, x_2) = 2^n\}$; cf. exemple "Gradient et matrice jacobienne" (p. 6).

(Solution p. 34.)

Exercice – Matrice jacobienne et gradient Donner une expression de la matrice jacobienne de f en x en fonction de gradients des fonctions scalaires f_i en x. (Solution p. 34.)

Seule, l'existence de la matrice jacobienne en x offre très peu de garanties de régularité sur f en x. Il est ainsi possible que la fonction f ne soit pas même pas continue en x. (On rappelle que pour les fonctions d'une variable, l'existence de la dérivée en un point implique la continuité en ce point.)

Exercice – Fonction discontinue I (•) Construire une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ dont le gradient existe en $(0, 0)$ mais qui soit discontinue en $(0, 0)$. (Solution p. 34.)

Exercice – Fonction discontinue II (••) Construire une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ dont la dérivée dans la direction $h \in \mathbb{R}^2$

$$f'(x, h) := \lim_{t \to 0} \frac{f(x + th) - f(x)}{t}$$

existe en $x = (0,0)$ pour tout $h \in \mathbb{R}^2$, mais qui ne soit pas continue en $(0,0)$. (Solution p. 35.)

Une façon simple de renforcer la régularité de la fonction f est d'exiger qu'elle soit continûment différentiable :

Définition – Continue différentiabilité

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et x un point de U. La fonction f est *continûment différentiable* si pour tout $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, n\}$, la dérivée partielle

$$x \in U \mapsto \partial_j f_i(x) \in \mathbb{R}$$

est définie et continue en tout point de U.

Remarque – Continue différentiabilité – définitions équivalentes

Alternativement, la fonction f est continûment différentiable si (et seulement si) les dérivées partielles $x \in U \mapsto \partial_j f(x) \in \mathbb{R}^m$ sont définies et continues pour tout $i \in \{1, \dots, m\}$ ou encore si la fonction $x \in U \mapsto J_f(x) \in \mathbb{R}^{m \times n}$ est définie et continue.

On qualifiera de *différentiable* une fonction qui admet un développement limité au premier ordre. La différentiabilité est la transposition naturelle du concept de dérivabilité aux fonctions de plusieurs variables : pour jouer ce rôle, l'existence de la matrice jacobienne est une propriété trop faible et la continue différentiabilité est une propriété trop forte.

Définition – Différentiabilité

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et x un point de U. On dit que f est *différentiable en x* si la matrice jacobienne de f en x existe et que $f(x+h)$ admet en $h = 0$ le développement limité au 1er ordre $h \mapsto f(x) + J_f(x) \cdot h$: il existe sur un voisinage de $h = 0$ une fonction ε à valeurs dans $\mathbb{R}^m$ vérifiant $\lim_{h \to 0} \varepsilon(h) = 0$ telle que

$$f(x+h) = f(x) + J_f(x) \cdot h + \varepsilon(h)\|h\|.$$

On dit que f est *différentiable* (ou *différentiable sur U*) si elle est différentiable en tout point x de U.

Le plus souvent, la façon la plus simple de prouver la différentiabilité d'une fonction est d'établir sa continue différentiabilité ; en effet, on a :

Proposition – Continue différentiabilité implique différentiabilité

Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^m$. Si f est continûment différentiable, f est différentiable.

MATRICE JACOBIENNE DÉFINIE

$J_f : U \to \mathbb{R}^{m\times n}$ existe.

$$\Uparrow$$

DIFFÉRENTIABILITÉ

$J_f : U \to \mathbb{R}^{m\times n}$ existe et pour tout $x \in U$,
$f(x+h) = f(x) + J_f(x) \cdot h + \varepsilon_x(h)\|h\|$
avec $\lim_{h\to 0} \varepsilon_x(h) = 0$.

$$\Uparrow$$

CONTINUE DIFFÉRENTIABILITÉ

$J_f : U \to \mathbb{R}^{m\times n}$ existe et est continue.

FIGURE 2 – Relations entre existence de la matrice jacobienne, différentiabilité et continue différentiabilité.

Démonstration Supposons $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ continûment différentiable. Soit $x \in U$ et $r > 0$ telle que la boule fermée centrée en x et de rayon r soit incluse dans U

$$\overline{B}(x, r) = \{y \in \mathbb{R}^n \mid \|y - x\| \leq r\} \subset U$$

et soit $h \in \mathbb{R}^n$ tel que $\|h\| \leq r$. La variation de f entre x et $x + h$ satisfait

$$f(x+h) - f(x) = \sum_{i=1}^{n} f(x + (h_1, \dots, h_{i-1}, h_i, 0, \dots)) - f(x + (h_1, \dots, h_{i-1}, 0, 0, \dots)).$$

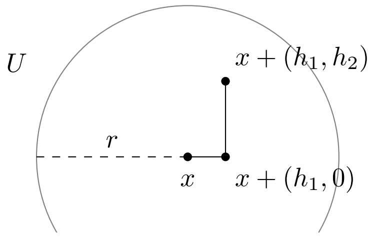

FIGURE 3 – Ligne brisée joignant x et $x + h$.

Or comme pour tout i la fonction

$$t \in [0, 1] \mapsto f(x + (h_1, \dots, th_i, 0, \dots))$$

est dérivable de dérivée $\partial_i f(x + (h_1, \dots, th_i, 0, \dots))h_i$ et que cette expression est une fonction continue – et donc intégrable – de la variable t, par le théorème fondamental du calcul (p. 19), on obtient

$$f(x + (h_1, \dots, h_{i-1}, h_i, 0, \dots)) - f(x + (h_1, \dots, h_{i-1}, 0, 0, \dots)) = h_i \int_0^1 \partial_i f(x + (h_1, \dots, h_{i-1}, th_i, 0, \dots))\, dt.$$

Par ailleurs, comme

$$f'(x) \cdot h = \sum_{i=1}^{n} \partial_i f(x) h_i = \sum_{i=1}^{n} h_i \int_0^1 \partial_i f(x)\, dt,$$

on a

$$f(x+h) - f(x) - \sum_i \partial_i f(x) h_i = \sum_{i=1}^{n} h_i \int_0^1 \left[\partial_i f(x + (h_1, \dots, h_{i-1}, th_i, 0, \dots)) - \partial_i f(x)\right] dt.$$

Par continuité des dérivées partielles en x, si r est choisi suffisamment petit pour que $\|\partial_i f(y) - \partial_i f(x)\| \leq \varepsilon/n$ quand $\|y - x\| \leq r$, alors l'inégalité triangulaire fournit

$$\left\| \int_0^1 [\partial_i f(x + (h_1, \dots, h_{i-1}, th_i, 0, \dots)) - \partial_i f(x)] \, dt \right\| \leq$$
$$\int_0^1 \|\partial_i f(x + (h_1, \dots, h_{i-1}, th_i, 0, \dots)) - \partial_i f(x)\| \, dt \leq \varepsilon/n$$

et donc, toujours par inégalité triangulaire, comme $|h_i| \leq \|h\|$,

$$\left\| f(x+h) - f(x) - \sum_{i=1}^n \partial_i f(x) h_i \right\| \leq \sum_{i=1}^n |h_i| \varepsilon/n \leq \varepsilon \|h\|.$$

La fonction f admet donc un dévelopement limité au 1er ordre en x. ∎

Proposition – Différentiabilité implique continuité

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et $x \in U$. Si f est différentiable en x, f est continue en x.

Exercice – Différentiabilité implique continuité (•) Démontrer la proposition "Différentiabilité implique continuité" (p. 11). (Solution p. 36.)

Exercice – Fonctions affines (•) Soit $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$. Calculer la matrice jacobienne de la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ définie par $f(x) = A \cdot x + b$ et montrer que cette fonction est différentiable. (Solution p. 35.)

L'existence de la matrice jacobienne de f en x ne garantit pas l'existence d'un développement au premier ordre de f en x. Par contre, si un tel développement existe, il est nécessairement obtenu à partir de la matrice jacobienne comme le montre l'exercice suivant.

Exercice – Développement limité au premier ordre (••) Montrer que pour $a \in \mathbb{R}^m$ et $B \in \mathbb{R}^{m \times n}$, si f est une fonction qui vérifie dans un voisinage de $h = 0$

$$f(x+h) = a + B \cdot h + \varepsilon(h)\|h\|$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$ alors $a = f(x)$, la matrice jacobienne $J_f(x)$ est bien définie et $B = J_f(x)$. (Solution p. 35.)

Définition – Différentielle

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et x un point de U. Si f est différentiable en x on appelle alors *différentielle de f en x* l'application linéaire $df(x)$ associée à la matrice jacobienne

$$df(x) := (h \in \mathbb{R}^n \mapsto J_f(x) \cdot h \in \mathbb{R}^m)$$

Si l'on identifie applications linéaires et matrices, la différentielle se confond avec la matrice jacobienne elle-même. Pour tout $h \in \mathbb{R}^n$,

$$df(x) \cdot h = J_f(x) \cdot h = \sum_{i=1}^{n} \partial_i f(x) h_i.$$

Remarque – Variation d'une fonction en un point

Comme le montre l'égalité caractérisant la différentiabilité de f en x, le terme $df(x) \cdot h$ représente une approximation – linéaire en h – de la variation $\Delta f(x, h) := f(x+h) - f(x)$ de f en x dans la direction h, car

$$\Delta f(x, h) = df(x) \cdot h + \varepsilon(h)\|h\|.$$

Exercice – Différentiabilité (••) Montrer que f est différentiable en x si et seulement si $J_f(x)$ est bien définie et que

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \left(\frac{f(x+h) - f(x)}{\|h\|} - J_f(x) \cdot \frac{h}{\|h\|} \right) = 0.$$

(Solution p. 36.)

Proposition – Différentielle et dérivée

Soit U un ouvert de $\mathbb{R}$, $f : U \to \mathbb{R}^m$ et $x \in U$. La fonction f est différentiable en x si et seulement si elle est dérivable en x. Dans ce cas, pour tout $h \in \mathbb{R}$,

$$df(x) \cdot h = f'(x)h.$$

Démonstration Avec la définition de la matrice jacobienne, il est clair que la dérivée de f en x existe si et seulement si $J_f(x)$ existe et qu'auquel cas on a $J_f(x) = [f'(x)]$. On sait aussi que la fonction f est dérivable en x si et seulement si elle admet un développement limité au premier ordre en x (cf. la proposition "Dérivée et développement limité" (p. 26) ainsi que sa réciproque (p. 26)), soit

$$f(x+h) = f(x) + f'(x)h + \varepsilon(h)|h|$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$, ce que l'on peut écrire sous la forme

$$f(x+h) = f(x) + [f'(x)] \cdot h + \varepsilon(h)\|h\| = f(x) + J_f(x) \cdot h + \varepsilon(h)\|h\|,$$

ce qui équivaut à la différentiabilité de f en x. La relation $df(x) \cdot h = J_f(x) \cdot h$ fournit le dernier élément de la proposition. ∎

Proposition – Différentielle et gradient

Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et $x \in U$. Si la fonction f est différentiable en x, alors pour tout $h \in \mathbb{R}^n$,

$$df(x) \cdot h = \langle \nabla f(x), h \rangle.$$

Démonstration Par définition de la différentielle (p. 11), $df(x) \cdot h = J_f(x) \cdot h$. Or, d'après la définition du gradient (p. 6), $\nabla f(x) = J_f(x)^\top$. Par conséquent, $df(x) \cdot h = \nabla f(x)^\top \cdot h = \langle \nabla f(x), h \rangle$. ■

Calcul Différentiel

Notations

Comme utilisateur du calcul différentiel vous avez probablement déjà croisé des notations assez éloignées de celles que nous avons exploitées jusqu'à présent. Vous avez peut-être mémorisé des règles de calcul élémentaires telles que

$$d(x+y) = dx + dy, \; d(xy) = y\,dx + x\,dy$$

ou appris en thermodynamique que

$$dU = TdS - PdV, \; \text{ soit } \; T = \frac{\partial U}{\partial S} \text{ et } P = -\frac{\partial U}{\partial V}.$$

Bien utilisées, ces notations ont le potentiel de simplifier significativement la pratique du calcul différentiel ; elle recèlent toutefois un potentiel d'ambiguité – et donc des risques d'erreurs – contre lequel il faut se prémunir. Nous allons donc détailler ces notations sur un exemple et les interpréter à la lumière des concepts déjà introduits.

Expressions (fonctions implicites)

La première technique consiste à favoriser l'usage d'expressions mathématiques – comme "$x^2 + y^2$" – pour désigner des grandeurs variables, sans nécessairement expliciter les fonctions correspondantes, la liste de leurs arguments ou le domaine de définition associés. Tous ces éléments doivent alors être inférés du contexte ; ainsi on peut assez naturellement associer à l'expression "$x^2 + y^2$" la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2 \in \mathbb{R}$. Mais d'autres choix sont défendables [3]. Une fois ce choix fait, on interprète le terme $d(x^2 + y^2)$ comme

$$d(x^2 + y^2) := df(x, y).$$

Variables nommées

Dans un contexte applicatif donné, il est fréquent que des noms (ou symboles) particuliers soit attachés aux grandeurs variables (plutôt que les génériques "x_1", …, "x_n") et qu'il soit plus naturel ou pratique d'utiliser ces noms pour désigner

3. Peut-être que dans le contexte où l'on exploite l'expression, les grandeurs x et y ne sont définies que si $x > 0$ et $y > 0$; peut-être a-t-on une bonne raison de plutôt lister la variable y avant x ; peut-être y a-t-il une troisième variable z et que ça n'est que "par accident" que l'expression $x^2 + y^2$ ne dépend pas de z, etc.

les variables plutôt qu'un indice [4] [5]. Pour ce qui est des dérivées partielles en particulier, dans cet esprit, on peut alors convenir de noter

$$\frac{\partial(x^2+y^2)}{\partial x} := \frac{\partial f}{\partial x}(x,y) := \partial_x f(x,y) := \partial_1 f(x,y)$$

et

$$\frac{\partial(x^2+y^2)}{\partial y} := \frac{\partial f}{\partial y}(x,y) := \partial_y f(x,y) := \partial_2 f(x,y).$$

Compte tenu de l'égalité $f(x,y) = x^2+y^2$, on a ici

$$\frac{\partial(x^2+y^2)}{\partial x} = 2x \;\text{ et }\; \frac{\partial(x^2+y^2)}{\partial y} = 2y.$$

Différentielle des variables

En poussant jusqu'au bout la logique de la différentiation des expressions, on peut encore simplifier les notations. À ce stade, nous avons établi que pour tout $(h_x, h_y) \in \mathbb{R}^2$, nous avions

$$d(x^2+y^2)\cdot(h_x,h_y) = \frac{\partial(x^2+y^2)}{\partial x}h_x + \frac{\partial(x^2+y^2)}{\partial y}h_y = 2x\,h_x + 2y\,h_y. \qquad (1)$$

Or, en utilisant les mêmes conventions, nous obtenons

$$dx\cdot(h_x,h_y) = \frac{\partial x}{\partial x}h_x + \frac{\partial x}{\partial y}h_y = h_x \;\text{ et }\; dy\cdot(h_x,h_y) = \frac{\partial y}{\partial x}h_x + \frac{\partial y}{\partial y}h_y = h_y.$$

Par conséquent, nous pouvons réécrire la relation (1) sous la forme

$$d(x^2+y^2) = \frac{\partial(x^2+y^2)}{\partial x}dx + \frac{\partial(x^2+y^2)}{\partial y}dy = 2x\,dx + 2y\,dy.$$

Exercice – Espace-temps (•) Montrer l'existence et calculer la différentielle de

$$x^2+y^2+z^2-c^2t^2$$

en utilisant les conventions de cette section. (Solution p. 37.)

Exercice – Robustesse des notations (••) On considère successivement l'expression x^2+y^2 comme une fonction non plus de (x,y), mais de (y,x) puis de (x,y,z). Comment interpréter les termes dx et dy dans chacun de ces cas ? La relation $d(x^2+y^2) = 2x\,dx + 2y\,dy$ est-elle toujours valable ? (Solution p. 37.)

4. En particulier, si la fonction considérée est implicite, car issue d'une expression, il n'y a probablement pas d'ordre "naturel" pour lister les variables.

5. C'est le même argument qui motive en Python de n'utiliser les arguments positionnels, comme dans l'appel `ask("Quit?", 3)`, qu'avec modération. L'alternative, utiliser les arguments nommés, comme dans l'appel `ask(question="Quit?", retries=3)`, peut s'avèrer plus lisible.

Règles de calcul

La définition de la différentielle, sa relation à la matrice jacobienne et avec la continue différentiabilité permettent d'élaborer un nombre quasi-illimité de "règles de calcul élémentaires" réutilisables dont nous donnons quelques exemples importants.

Proposition – Différentielle d'une application linéaire

Pour toute matrice $A \in \mathbb{R}^{m \times n}$, l'application linéaire $f : x \in \mathbb{R}^n \mapsto A \cdot x \in \mathbb{R}^m$ est différentiable et pour tout $x \in \mathbb{R}^n$, $df(x) = A$; autrement dit

$$d(A \cdot x) = A \cdot dx.$$

Démonstration Pour tout $x \in \mathbb{R}^n$, on a $f_i(x) = (A \cdot x)_i = \sum_{k=1}^n A_{ik} x_k$, donc la j-ème dérivée partielle de f_i existe et

$$\partial_j f_i(x) = \partial_j \left(x \mapsto \sum_{k=1}^n A_{ik} x_k \right)(x) = A_{ij}.$$

La matrice jacobienne $J_f(x)$ est donc définie et $J_f(x) = A$. Chaque coefficient de J_f est une constante et donc une fonction continue de x : la fonction f est continûment différentiable – et donc différentiable (p. 8) – et $df(x) = J_f(x) = A$. ■

Proposition – Règle du produit (élémentaire)

L'application produit $\pi : (x, y) \in \mathbb{R}^2 \mapsto xy \in \mathbb{R}$ est différentiable et pour tout $(h_x, h_y) \in \mathbb{R}^2$, on a $d\pi(x, y) \cdot (h_x, h_y) = x h_y + y h_x$; autrement dit

$$d(xy) = x\, dy + y\, dx.$$

Démonstration Les dérivées partielles de xy existent et sont continues : on a

$$\frac{\partial(xy)}{\partial x} = y \;\text{ et }\; \frac{\partial(xy)}{\partial y} = x.$$

Par conséquent, l'application produit est continûment différentiable, donc différentiable (p. 8), et

$$d(xy) = \frac{\partial(xy)}{\partial x} dx + \frac{\partial(xy)}{\partial y} dy = x\, dy + y\, dx.$$

■

Ce type de règles de calcul élémentaires peuvent servir de briques de base pour élaborer des règles de calcul "génériques", faisant intervenir des fonctions différentiables arbitraires et permettant de traiter des calculs plus complexes. Les deux résultats qui permettent de mener cette extension du calcul différentiel sont la règle d'assemblage (p. 17) – fondée sur la règle de différentiation composante par composante (p. 17) – et la règle de différentiation en chaîne (p. 16).

Théorème – Règle de différentiation en chaîne

Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ et $g : V \subset \mathbb{R}^n \to \mathbb{R}^m$ deux fonctions définies sur des ouverts U et V et telles que $f(U) \subset V$. Si f est différentiable en $x \in U$ et g est différentiable en $f(x) \in V$, alors la composée $g \circ f$ est différentiable en x et

$$d(g \circ f)(x) = dg(y) \cdot df(x) \ \text{ où } \ y = f(x).$$

Démonstration L'objectif de la preuve est de montrer que dans un voisinage de $h = 0$,

$$g(f(x+h)) - g(f(x)) = (dg(f(x)) \cdot df(x)) \cdot h + \varepsilon(h)\|h\|$$

où $\lim_{h \to 0} \varepsilon(h) = 0$. La fonction g étant différentiable en $f(x)$, il existe une fonction ε_1 définie dans un voisinage de 0 et telle que $\lim_{h \to 0} \varepsilon_1(h) = 0$, vérifiant

$$g(f(x) + k) - g(f(x)) = dg(f(x)) \cdot k + \varepsilon_1(k)\|k\|.$$

Choisissons $k = f(x+h) - f(x)$ dans cette équation, de telle sorte que

$$g(f(x) + k) = g\left(f(x) + (f(x+h) - f(x))\right) = g(f(x+h)).$$

Nous obtenons donc

$$g(f(x+h)) - g(f(x)) = dg(f(x)) \cdot (f(x+h) - f(x)) + \varepsilon_1(k)\|k\|.$$

Notons que la fonction $\varepsilon_2(h) := \varepsilon_1(f(x+h) - f(x))$ est définie dans un voisinage de 0 et que par continuité de f en x, $f(x+h) - f(x)$ tend vers 0 quand h tend vers 0, et par conséquent $\varepsilon_2(h) \to 0$ quand $h \to 0$. Avec cette notation, on a

$$\begin{aligned} g(f(x+h)) - g(f(x)) = dg(f(x)) \cdot (f(x+h) - f(x)) \\ + \varepsilon_2(h)\|f(x+h) - f(x)\|. \end{aligned}$$

Comme f est également différentiable en x, il existe une fonction ε_3 définie dans un voisinage de 0 et telle que $\lim_{h \to 0} \varepsilon(h) = 0$ vérifiant

$$f(x+h) - f(x) = df(x) \cdot h + \varepsilon_3(h)\|h\|.$$

En substituant cette relation dans la précédente, nous obtenons

$$g(f(x+h)) - g(f(x)) = dg(f(x)) \cdot (df(x) \cdot h) + \varepsilon(h)\|h\|$$

où $\varepsilon(0) = 0$ et dans le cas contraire,

$$\varepsilon(h) = dg(f(x)) \cdot \varepsilon_3(h) + \varepsilon_2(h) \left\| df(x) \cdot \frac{h}{\|h\|} + \varepsilon_3(h) \right\|.$$

Il suffit pour conclure de prouver que $\varepsilon(h) \to 0$ quand $h \to 0$. Or,

$$\begin{aligned} \|\varepsilon(h)\| &\leq \|dg(f(x)) \cdot \varepsilon_3(h)\| + \|\varepsilon_2(h)\| \times \|df(x) \cdot (h/\|h\|)\| + \|\varepsilon_2(h)\| \times \|\varepsilon_3(h)\| \\ &\leq \|dg(f(x))\| \times \|\varepsilon_3(h)\| + \|\varepsilon_2(h)\| \times \|df(x)\| + \|\varepsilon_2(h)\| \times \|\varepsilon_3(h)\|, \end{aligned}$$

le résultat est donc acquis. ■

Exercice – Différentiation en chaîne des fonctions continûment différentiables (•) Montrer que si dans l'énoncé de la règle de différentiation en chaîne (p. 16) les fonctions f et g sont continûment différentiables, alors $g \circ f$ l'est également. (Solution p. 37.)

Théorème – Règle de différentiation composante par composante

Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et $x \in U$. La fonction f est différentiable en x si et seulement si toutes ses composantes f_i sont différentiables en x. Dans ce cas, on a

$$(df(x))_i = df_i(x).$$

Démonstration Par la règle de dérivation composante par composante (p. 25), la matrice jacobienne de f en x existe si et seulement si toutes les matrices jacobienne $J_{f_i}(x)$ existent et on a alors $(J_f(x))_i = J_{f_i}(x)$. La différentiabilité de f en x se traduit donc par

$$f(x+h) = f(x) + J_f(x) \cdot h + \varepsilon(h)\|h\|$$

où $\lim_{h \to 0} \varepsilon(h) = 0$, ce qui entraine pour tout i,

$$f_i(x+h) = f_i(x) + (J_f(x))_i \cdot h + \varepsilon_i(h)\|h\| = f_i(x) + J_{f_i}(x) \cdot h + \varepsilon_i(h)\|h\|$$

avec $\lim_{h \to 0} \varepsilon_i(h) = 0$; chaque composante f_i est donc différentiable. La réciproque s'établit de manière similaire. On déduit alors de la relation $(J_f(x))_i = J_{f_i}(x)$ que $(df(x))_i = df_i(x)$. ■

Cette règle entraîne :

Corollaire – Règles d'assemblage et désassemblage

Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$, $g : U \to \mathbb{R}^p$ et $x \in U$. La fonction $(f, g) : U \to \mathbb{R}^{m+p}$ est différentiable en x si et seulement si f et g sont différentiables en x ; on a alors

$$d(f, g)(x) = (df(x), dg(x)).$$

Exercice – Désassemblage (••) Montrer que pour tous $m, p \in \mathbb{N}$, les fonctions $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^p \mapsto x_1 \in \mathbb{R}^m$ et $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^p \mapsto x_2 \in \mathbb{R}^p$ sont différentiables. En déduire une démonstration de "la règle de désassemblage" (p. 17). (Solution p. 37.)

Démonstration La fonction (f, g) d'une part et les fonctions f et g d'autre part ont le même jeu de composantes scalaires ; la conclusion s'ensuit par la règle de différentiation composante par composante (p. 17). ■

Exercice – Différentiation en chaîne et assemblage (•••) Combiner la règle d'assemblage (p. 17) et la règle de différentiation en chaîne (p. 16) en un résultat unique qui implique les deux résultats (Solution p. 38.)

À titre d'exemple, montrons comment ces deux résultats permettent de généraliser la règle élémentaire du produit (p. 15) :

Proposition – Règle du produit (générique)

Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$, $g : U \to \mathbb{R}$ et $x \in U$. Si les fonctions f et g sont différentiables en x, alors leur produit fg également et

$$d(fg)(x) = f(x)dg(x) + g(x)df(x).$$

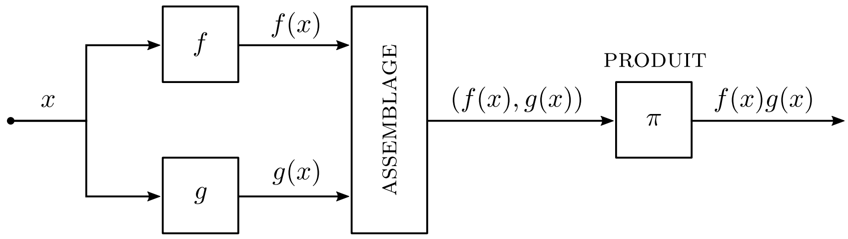

FIGURE 4 – Graphe de calcul de l'application $x \mapsto f(x)g(x)$.

Démonstration Par assemblage (p. 17), la fonction (f, g) est différentiable en x et $d(f, g)(x) = (df(x), dg(x))$. Sa composition par la fonction produit $\pi : (x, y) \mapsto xy$ est le produit fg :

$$fg = \pi \circ (f, g).$$

Par la règle de différentation en chaîne (p. 16), le produit fg est différentiable et

$$\begin{aligned} d(fg)(x) &= d\pi(f(x), g(x)) \cdot d(f, g)(x) \\ &= d\pi(f(x), g(x)) \cdot (df(x), dg(x)) \\ &= f(x)dg(x) + g(x)df(x). \end{aligned}$$

■

De façon similaire, on peut désormais tirer des conséquences élargies de la proposition "Différentielle d'une application linéaire" (p. 15) :

Proposition – Linéarité de la différentielle

Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$, $g : U \to \mathbb{R}$, $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$ et $x \in U$. Si les fonctions f et g sont différentiables en x, alors la combinaison linéaire $\lambda f + \mu g$ également et

$$d(\lambda f + \mu g)(x) = \lambda df(x) + \mu dg(x).$$

Démonstration La fonction $\lambda f + \mu g$ peut être obtenue en composant la fonction (f, g) et la fonction linéaire $A : (x, y) \in \mathbb{R}^2 \mapsto \lambda x + \mu y$. Par la règle d'assemblage (p. 17) et la règle de différentation en chaîne (p. 16), elle est donc différentiable en x et comme la différentielle de l'application linéaire A en tout point est elle-même (p. 15),

$$\begin{aligned} d(\lambda f + \mu g)(x) &= dA(f(x), g(x)) \cdot (df(x), dg(x)) \\ &= A \cdot (df(x), dg(x)) \\ &= \lambda df(x) + \mu dg(x). \end{aligned}$$

■

Exercice – Produit scalaire Montrer que la fonction

$$\langle \cdot, \cdot \rangle : (x, y) \in \mathbb{R}^{2n} \to \langle x, y \rangle \in \mathbb{R}$$

est différentiable et calculer son gradient. (Solution p. 38.)

Variation des fonctions

Lorsque la fonction f est différentiable en x, nous disposons de l'égalité

$$f(x + h) - f(x) = df(x) \cdot h + \varepsilon(h)\|h\|$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$. Cette égalité est de nature asymptotique, ce qui veut dire que pour maîtriser l'écart entre $f(x + h)$ et $f(x)$, nous devons être en mesure de faire tendre h vers 0 ; si le vecteur h est fixé et même s'il est "petit", en toute rigueur, cette relation ne nous fournit aucune information.

Mais tout n'est pas perdu : si nous savons que f est différentiable non pas uniquement en x mais sur tout le segment $[x, x + h]$, il est possible de calculer la différence entre $f(x + h)$ et $f(x)$ en intégrant les variations infinitésimales de f le long de ce segment.

Théorème – Théorème fondamental du calcul (monovariable)

Soit $x \in \mathbb{R}$ et $h \geq 0$. Si la fonction $f : [x, x + h] \subset \mathbb{R} \to \mathbb{R}^m$ est dérivable et que sa dérivée f' est intégrable alors

$$f(x + h) - f(x) = \int_x^{x+h} f'(y)\, dy.$$

Démonstration Voir l'enseignement de calcul intégral. ■

Remarque – À propos du terme "intégrable"

À ce stade, vous pouvez retenir que si f' est continue, continue par morceaux ou même intégrable au sens de Riemann, elle est "intégrable" comme le demandent les hypothèses du théorème.

Dans ce chapitre, sauf précision contraire, le terme "intégrable" doit être compris comme "intégrable au sens de Lebesgue". La définition de ce concept – ainsi que la preuve du théorème fondamental du calcul – seront fournies dans le volet calcul intégral de l'enseignement.

Remarque – Forme générale du théorème fondamental du calcul

Si l'on adopte au lieu de l'intégrale de Lebesgue l'intégrale encore plus générale de Henstock-Kurzweil (cf. calcul intégral), alors toute fonction dérivée est automatiquement intégrable. Le théorème fondamental du calcul est alors valable en toute généralité ; il prend la forme suivante : si $x \in \mathbb{R}$, $h \geq 0$ et la fonction $f : [x, x+h] \to \mathbb{R}^m$ est dérivable, alors f' est intégrable (au sens de Henstock-Kurzweil) et

$$f(x+h) - f(x) = (\mathrm{HK}) \int_x^{x+h} f'(y)\, dy.$$

Cette forme avancée du théorème est toutefois rarement nécessaire ; elle est néanmoins utile pour prouver l'inégalité des accroissements finis (p. 21) en toute généralité. Cette extension est aussi applicable à la version multivariable du théorème fondamental du calcul (p. 20) : si l'on utilise l'intégrale de Henstock-Kurzweil, il sera inutile de vérifier que l'application $t \mapsto df(x+th) \cdot h$ est intégrable pour appliquer le théorème ; comme dérivée de l'application $t \mapsto f(x+th)$, cette fonction l'est automatiquement.

Théorème – Théorème fondamental du calcul (multivariable)

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$, $x \in U$ et $h \in \mathbb{R}^n$ tels que le segment

$$[x, x+h] = \{x + th \mid t \in [0,1]\}$$

soit inclus dans U. Si f est différentiable en tout point de $[x, x+h]$ et que l'application $t \in [0,1] \mapsto df(x+th) \cdot h \in \mathbb{R}^m$ est intégrable, alors

$$f(x+h) - f(x) = \int_0^1 df(x+th) \cdot h\, dt.$$

Démonstration L'ensemble U étant ouvert, il existe un $\varepsilon > 0$ tel que l'intervalle ouvert $I := \left]-\varepsilon, 1+\varepsilon\right[$ soit inclus dans U. On note ϕ la fonction $I \to \mathbb{R}^n$ définie par

$$\phi(t) = f(x+th)$$

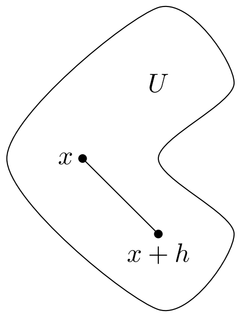

Figure 5 – Géométrie du théorème du calcul multivariable (p. 20)

La fonction ϕ est différentiable – et donc dérivable – en tout point de $[0,1]$ comme composée des fonctions différentiables f et $t \mapsto x+th$ (p. 16) ; sa dérivée est donnée par

$$\begin{aligned}\phi'(t) &= d\phi(t) \\ &= df(x+th)\cdot d(t\mapsto x+th) \\ &= df(x+th)\cdot (t\mapsto x+th)' \\ &= df(x+th)\cdot h\end{aligned}$$

Par le théorème fondamental du calcul (p. 19), comme par hypothèse ϕ' est intégrable sur $[0,1]$, on a donc

$$f(x+h)-f(x) = \phi(1)-\phi(0) = \int_0^1 \phi'(t)\,dt = \int_0^1 df(x+th)\cdot h\,dt.$$

■

Exercice – Cas des fonctions continûment différentiables (•) Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$, $x \in U$ et $h \in \mathbb{R}^n$ tels que le segment $[x, x+h] = \{x+th \mid t \in [0,1]\}$ soit inclus dans U. Montrer que si f est continûment différentiable sur U, le théorème fondamental du calcul (p. 20) est applicable. (Solution p. 38.)

Théorème – Inégalité des accroissements finis (monovariable)

Soit $x \in \mathbb{R}$, $h \geq 0$, $f : [x, x+h] \to \mathbb{R}^m$. Si f est dérivable sur $[x, x+h]$ et M est un majorant de $\|f'\|$, c'est-à-dire si

$$\text{pour tout } y \in [x, x+h],\ \|f'(y)\| \leq M.$$

Alors

$$\|f(x+h)-f(x)\| \leq Mh.$$

Démonstration Par la forme générale du théorème fondamental du calcul (p. 20), la fonction f' est intégrable au sens de Henstock-Kurzweil et

$$f(x+h) - f(x) = (\mathrm{HK}) \int_x^{x+h} f'(y)\, dy.$$

La théorie de l'intégrale de Henstock-Kurzweil nous garantit qu'il est possible d'obtenir des approximations arbitrairement précises de cette intégrale au moyen de sommes de Riemman [6]. Cela signifie que pour tout $\varepsilon > 0$, il existe des réels $x_0, \ldots, x_k, t_0, \ldots, t_{k-1}$ vérifiant

$$x = x_0 \leq t_0 \leq x_1 \leq t_1 \leq \cdots \leq x_{k-1} \leq t_{k-1} \leq x_k = x + h$$

telle que la somme

$$S = \sum_{i=0}^{k-1} f'(t_i)(x_{i+1} - x_i)$$

satisfasse

$$\left\| (\mathrm{HK}) \int_x^{x+h} f'(t)\, dt - S \right\| \leq \varepsilon.$$

En exploitant deux fois l'inégalité triangulaire, on obtient donc

$$\|f(x+h) - f(x)\| \leq \|S\| + \varepsilon \leq \sum_{i=0}^{k-1} \|f'(t_i)\| |x_{i+1} - x_i| + \varepsilon.$$

Comme $\|f'(t_i)\| \leq M$ pour tout $i \in \{0, \ldots, k-1\}$,

$$\sum_{i=0}^{k-1} \|f'(t_i)\| |x_{i+1} - x_i| \leq \sum_{i=0}^{k-1} M |x_{i+1} - x_i| \leq M \sum_{i=0}^{k-1} |x_{i+1} - x_i|$$

et comme $x = x_0 \leq x_1 \leq \cdots \leq x_k = x + h$,

$$\sum_{i=0}^{k-1} |x_{i+1} - x_i| = \sum_{i=0}^{k-1} (x_{i+1} - x_i) = x_p - x_0 = (x+h) - x = h.$$

On a donc $\|S\| \leq Mh$ et par conséquent $\|f(x+h) - f(x)\| \leq Mh + \varepsilon$; le choix de $\varepsilon > 0$ étant arbitraire, on en déduit le résultat cherché. ∎

L'énoncé de l'inégalité des accroissements finis ne fait aucune hypothèse sur la régularité de la fonction f' ; c'est cette généralité qui rend sa démonstration technique. Si l'on accepte des hypothèses un peu plus contraignantes, elle peut être simplifiée de façon significative.

Exercice – Cas des fonctions continûment différentiables (•) Trouver une démonstration simple de l'inégalité des accroissement finis (p. 21) reposant sur le théorème fondamental du calcul (p. 19) dans le cas où la fonction f' est continue (ou continue par morceaux, ou intégrable au sens de Riemann,

6. en combinant la définition de l'intégrabilité de f' au sens de Henstock-Kurzweil et le lemme de Cousin du calcul intégral.

ou intégrable au sens de Lebesgue, selon la théorie de l'intégration que vous connaissez à ce stade). (Solution p. 38.)

Il existe également une démonstration alternative de l'inégalité des accroissements finis qui ne repose pas sur le calcul intégral, mais sur les identités associées à la norme euclidienne et sur le théorème des valeurs intermédiaires :

Exercice – Inégalité des accroissements finis (version euclidienne) (••) Soit $\phi : y \in [a, a+h] \to \mathbb{R}$ la fonction définie par

$$\phi(y) = \langle f(x+h) - f(x), f(y) \rangle .$$

En appliquant le théorème des valeurs intermédiaires à ϕ, prouver l'inégalité des accroissements finis (p. 21). (Solution p. 39.)

Exercice – Inégalité de la valeur moyenne (•) Soit $f : [a, b] \subset \mathbb{R} \to \mathbb{R}^m$ une fonction intégrable et admettant une primitive ; on appelle *valeur moyenne de* f la grandeur

$$\langle f \rangle := \frac{1}{b-a} \int_a^b f(x)\, dx.$$

Quel est le lien entre $\langle f \rangle$ et la grandeur $\sup_{x \in [a,b]} \|f(x)\|$? (Solution p. 39.)

Egalité des accroissements finis ? (•) Soit $f : [0, 2\pi] \to \mathbb{R}^2$ la fonction définie par

$$f(t) = (\cos t, \sin t)$$

Peut-on trouver un $t \in [0, 2\pi]$ tel que $f(2\pi) - f(0) = f'(t) \times 2\pi$? (Solution p. 39.)

Théorème – Inégalité des accroissements finis (multivariable)

Soient U un ouvert de $\mathbb{R}^n$, et $f : U \to \mathbb{R}^m$ supposée différentiable en tout point d'un segment $[a, a+h]$ inclus dans U et dont la différentielle est majorée en norme par M sur $[a, a+h]$, c'est-à-dire telle que

$$\text{pour tout } x \in [a, a+h],\ \|df(x)\| \leq M.$$

Alors

$$\|f(a+h) - f(a)\| \leq M\|h\|.$$

Démonstration Considérons la fonction $\phi : t \mapsto f(a+th)$ déjà exploitée dans la démonstration de la proposition "Variation d'une fonction" (p. 20) ; cette fonction est dérivable sur $[0, 1]$, de dérivée $\phi'(t) = df(a+th) \cdot h$. De plus,

$$\|\phi'(t)\| = \|df(a+th) \cdot h\| \leq \|df(a+th)\|\|h\| \leq M\|h\|.$$

Par l'inégalité des accroissements finis dans le cas d'une variable réelle (p. 21),

$$\|f(a+h) - f(a)\| = \|\phi(1) - \phi(0)\| \leq M\|h\| \times 1 = M\|h\|.$$

■

Variation du logarithme (••) Il est possible de définir une version multivariable et vectorielle de la fonction logarithme, définie sur le *plan coupé*

$$U = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$$

et à valeurs dans $\mathbb{R}^2$. Cette fonction est différentiable et vérifie en tout point

$$\|d \log(x, y)\| = \frac{1}{\sqrt{x^2 + y^2}}.$$

Montrer que pour tout $(x, y) \in U$, on a

$$\| \log(x, y) - \log(x, -y)\| \leq 2\pi.$$

(Solution p. 40.)

Remarque – Normes non euclidiennes

Les versions monovariable (p. 21) et multivariable (p. 23) de l'inégalité des accroissements finis peuvent être généralisées à d'autres normes que la norme euclidienne. Dans le cas monovariable, si $\| \cdot \|_{\mathbb{R}^m}$ est une norme arbitraire sur $\mathbb{R}^m$ et que l'on dispose de la borne $\|f'(y)\|_{\mathbb{R}^m} \leq M$ sur $[x, x + h]$, alors on peut conclure que

$$\|f(x + h) - f(x)\|_{\mathbb{R}^m} \leq Mh.$$

Dans le cas multivariable, si de plus $\| \cdot \|_{\mathbb{R}^n}$ est une norme arbitraire sur $\mathbb{R}^n$ et que l'on définit pour tout $A \in \mathbb{R}^{m\times n}$ la norme d'opérateur associée

$$\|A\|_{\mathbb{R}^{m\times n}} := \sup_{x \neq 0} \frac{\|A \cdot x\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}},$$

alors on peut déduire de la borne $\|df(y)\|_{\mathbb{R}^{m\times n}} \leq M$ sur $[x, x + h]$ que

$$\|f(x + h) - f(x)\|_{\mathbb{R}^m} \leq M\|h\|_{\mathbb{R}^n}.$$

Seules des modifications mineures des démonstrations déjà présentées sont nécessaires pour établir ces résultats.

Annexes

Dérivée

Définition – Dérivée

Soient U un ouvert de $\mathbb{R}$, $f : U \to \mathbb{R}^m$ et $x \in U$. La fonction f est *dérivable en* x si la limite du *taux d'accroissement* de f en x existe ; cette limite est appelée *dérivée de* f *en* x et notée $f'(x)$:

$$f'(x) := \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x + h) - f(x)}{h}.$$

La fonction f est *dérivable (sur U)* si elle est dérivable en tout point x de U.

Cette définition accomode sans difficulté le cas des fonctions scalaires et vectorielles d'une variable réelle. Dans ce second cas, il suffit d'interpréter le terme $(f(x+h) - f(x))/h$ comme la multiplication du vecteur $f(x+h) - f(x)$ par le scalaire $1/h$. En outre, la dérivée d'une fonction vectorielle se déduit aisément de la dérivée de ses composantes :

Définition – Dérivée composante par composante

Soient U un ouvert de $\mathbb{R}$, $f : U \to \mathbb{R}^m$ et $x \in U$. La fonction f est dérivable en x si et seulement si pour tout $i \in \{1, \dots, m\}$, sa i-ème composante $f_i : U \to \mathbb{R}$ est dérivable en x. On a alors $(f'(x))_i = f_i'(x)$.

La définition de dérivée couvre le cas où la fonction f est définie sur un intervalle ouvert $]a, b[$ – ou d'ailleurs sur une réunion arbitraire d'intervalles ouverts de $\mathbb{R}$ – mais pas sur un intervalle fermé et borné $[a, b]$. Pour appréhender ce cas, on introduit classiquement les notions de dérivées à gauche et à droite.

Définition – Dérivée sur un intervalle fermé et borné

Soient $a, b \in \mathbb{R}$ avec $a < b$ et $f : [a, b] \to \mathbb{R}^m$. La fonction f est *dérivable sur* $[a, b]$ si elle est dérivable sur l'intervalle ouvert $]a, b[$ et que les dérivées de f à droite en a et à gauche en b existent. On pose alors

$$f'(a) := \lim_{\substack{h \to 0 \\ h > 0}} \frac{f(a+h) - f(a)}{h} \quad \text{et} \quad f'(b) := \lim_{\substack{h \to 0 \\ h < 0}} \frac{f(b+h) - f(b)}{h}.$$

Alternativement, on peut utiliser une pirouette pour se ramener à un domaine de définition ouvert. Cette approche nous est utile pour intégrer les dérivées au calcul différentiel multivariable qui n'est développé que pour des domaines de définition ouverts.

Proposition – Dérivée et prolongement

Soient $a, b \in \mathbb{R}$ avec $a < b$ et $f : [a, b] \to \mathbb{R}^m$. La fonction f est dérivable sur $[a, b]$ si et seulement si elle admet un prolongement g à un ensemble ouvert U de $\mathbb{R}$ contenant $[a, b]$ qui soit dérivable. Si c'est le cas, sa dérivée f' est égale à la restriction de la fonction g' à $[a, b]$.

Exercice – Dérivée et prolongement (•) Démontrer la proposition "Dérivée et prolongement" (p. 25). (Solution p. 33.)

La dérivabilité est équivalente à l'existence d'un développement limité au 1er ordre ; si elle est acquise, la valeur de la fonction et de sa dérivée au point de référence fournissent ce développement.

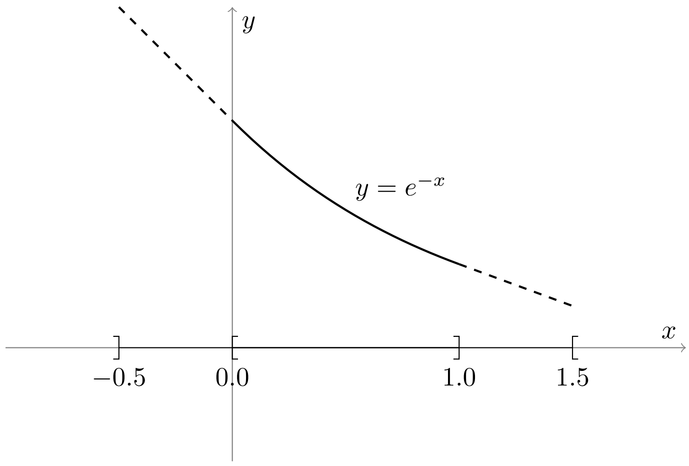

FIGURE 6 – Un prolongement dérivable de $f : x \in [0, 1] \mapsto e^{-x}$ à $]-0.5, 1.5[$. Ce prolongement g est défini par $g(x) = 1 - x$ si $x < 0$ et $g(x) = (2 - x)/e$ si $x > 1$.

Proposition – Dérivée et développement limité

Soient U un ouvert de $\mathbb{R}$, $f : U \to \mathbb{R}^m$ et $x \in U$. Si la fonction f est dérivable en x, alors dans un voisinage de x on a

$$f(x + h) = f(x) + f'(x)h + \varepsilon(h)|h|$$

où la fonction ε, définie dans un voisinage de $h = 0$ et à valeurs dans $\mathbb{R}^m$, satisfait

$$\lim_{h \to 0} \varepsilon(h) = 0.$$

Proposition – Dérivée et développement limité (réciproque)

Soient U un ouvert de $\mathbb{R}$, $f : U \to \mathbb{R}^m$ et $x \in U$. S'il existe deux vecteurs a et b de $\mathbb{R}^m$ tels que

$$f(x + h) = a + bh + \varepsilon(h)|h|$$

pour une fonction ε définie dans un voisinage de $h = 0$ telle que $\lim_{h \to 0} \varepsilon(h) = 0$, alors f est dérivable en x, $a = f(x)$ et et $b = f'(x)$.

Calcul matriciel

Les fragments de code de ce document utilisent le langage Python 3. La bibliothèque NumPy est exploitée :

```
>>> from numpy import *
```

Multiplication scalaire-vecteur

Pour tout scalaire $\lambda \in \mathbb{R}$ et vecteur $x \in \mathbb{R}^n$, on notera λx ou parfois $x\lambda$ la multiplication du vecteur x par le scalaire λ. Lorsque λ est non nul, on notera également x/λ le vecteur $(1/\lambda)x$.

Un vecteur de $\mathbb{R}^n$ est représenté dans NumPy par un tableau à une dimension :

```
>>> x = array([1, 2, 3])
>>> x.ndim
1
>>> shape(x)
(3,)
>>> size(x)
3
```

La multiplication d'un scalaire et d'un vecteur est désignée par le symbole `*` :

```
>>> 2 * x
array([2, 4, 6])
```

Matrices

Nous noterons $\mathbb{R}^{m\times n}$ l'ensemble des matrices à m lignes et n colonnes à coefficients réels. Une matrice telle que

$$\left[\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] \in \mathbb{R}^{2\times 3}$$

sera représentée avec NumPy par un tableau bi-dimensionnel:

```
>>> A = array([[1, 2, 3], [4, 5, 6]])
>>> A
array([[1, 2, 3],
       [4, 5, 6]])
>>> A.ndim
2
>>> shape(A)
(2, 3)
>>> size(A)
6
```

Applications linéaires

La raison d'être des matrices $\mathbb{R}^{m\times n}$ est de représenter les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Si A désigne une telle application linéaire, notons A_i ses m composantes scalaires :

$$A = (A_1, A_2, \ldots, A_m).$$

Si l'on désigne maintenant par e_j le j-ème vecteur de la base canonique de $\mathbb{R}^n$

$$e_1 = (1, 0, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0),\ \ldots\ ,\ e_n = (0, 0, 0, \ldots, 1),$$

il est possible d'associer à l'application linéaire $A : \mathbb{R}^n \to \mathbb{R}^m$ la matrice

$$[A] := [A_i(e_j)]_{ij} = \begin{bmatrix} A_1(e_1) & A_1(e_2) & \cdots & A_1(e_n) \\ \vdots & \vdots & \vdots & \vdots \\ A_m(e_1) & A_m(e_2) & \cdots & A_m(e_n) \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Réciproquement, étant donné une matrice

$$[a_{ij}]_{ij} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n},$$

il est possible de définir une application linéaire $A : \mathbb{R}^n \to \mathbb{R}^m$ par la relation

$$(A \cdot x)_i := \sum_j a_{ij} x_j$$

et cette opération est l'inverse de la précédente. Techniquement, cette correspondance établit un isomorphisme d'espaces vectoriels entre les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^m$ et les matrices de $\mathbb{R}^{m \times n}$.

Composition d'applications linéaires

Si A et B désignent des applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^n$ et de $\mathbb{R}^n$ dans $\mathbb{R}^m$ respectivement, la fonction composée $C = B \cdot A$ est une application linéaire qui vérifie

$$C_{ij} = \sum_k B_{ik} A_{kj}.$$

Autrement dit, la composition de fonction linéaires se traduit par la multiplication des matrices associées.

La méthode `dot` des tableaux NumPy permet de calculer ce produit matriciel :

```
>>> A = array([[1, 2, 3], [4, 5, 6]])
>>> B = array([[1, 0, 0], [0, 1, 0], [0, 0, 1]])
>>> A.dot(B)
array([[1, 2, 3],
       [4, 5, 6]])
```

Adjoint d'un opérateur

Lorsque $A : \mathbb{R}^n \to \mathbb{R}^m$ est un opérateur linéaire, on peut définir de façon unique l'opérateur adjoint $A^\top : \mathbb{R}^m \to \mathbb{R}^n$ comme l'unique opérateur tel que pour tout $x \in \mathbb{R}^n$ et tout $y \in \mathbb{R}^m$, on ait

$$\langle y, A \cdot x \rangle = \left\langle A^\top \cdot y, x \right\rangle.$$

La matrice associée à $A^\top$ est la transposée de la matrice représentant A :

```
>>> A
array([[1, 2, 3],
       [4, 5, 6]])
>>> A.T
array([[1, 4],
       [2, 5],
       [3, 6]])
```

Vecteurs colonnes et vecteur lignes

Dans le cadre du calcul matriciel, on associe souvent à un vecteur $x = (x_1, \dots, x_n)$ de $\mathbb{R}^n$ le vecteur colonne

$$\left[\begin{array}{c} x_1 \\ \vdots \\ x_n \end{array}\right] \in \mathbb{R}^{n \times 1}.$$

Dans cette terminologie, un vecteur colonne n'est pas, malgré son nom, un vecteur de $\mathbb{R}^n$, mais bien une matrice de taille $n \times 1$. Formellement, on a associé à x une matrice $X \in \mathbb{R}^{n \times 1}$, telle que $X_{i1} = x_i$. Le produit entre une matrice et un vecteur colonne de taille compatible n'est rien d'autre qu'un produit matriciel classique.

Concrètement, NumPy ne nécessite pas qu'un vecteur soit d'abord transformé en matrice pour réaliser un produit matrice-vecteur. La méthode `dot` des tableaux peut être utilisée ici aussi pour réaliser cette opération :

```
>>> A = array([[1, 2, 3], [4, 5, 6]])
>>> x = array([7, 8, 9])
>>> A.dot(x)
array([ 50, 122])
```

Le produit matriciel étant associatif, tant que l'on manipule des matrices et des vecteurs, il n'y a pas lieu de préciser si $A \cdot B \cdot C$ désigne $(A \cdot B) \cdot C$ (association à gauche) ou $A \cdot (B \cdot C)$ (association à droite). Comme le produit matrice-vecteur est un produit matriciel classique, quand x est un vecteur, $A \cdot B \cdot x$ désigne indifféremment $(A \cdot B) \cdot x$ ou $A \cdot (B \cdot x)$.

Exercices complémentaires

Spécialisation de la différentiation en chaîne

La règle générale de différentiation en chaîne (p. 16) s'applique à la composée de deux fonctions différentiables $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ et $g : V \subset \mathbb{R}^n \to \mathbb{R}^m$ quelles que soient les valeurs des entiers p, n et m.

En pratique, on préfère souvent "spécialiser" le calcul différentiel en utilisant les dérivées ou les gradients quand c'est possible et la différentielle uniquement en dernier recours.

Question 1 (○) Spécialiser la règle de différentiation en chaîne (p. 16) quand $p = n = 1$. (Solution p. 40.)

Question 2 (•) Spécialiser la règle de différentiation en chaîne (p. 16) quand $p = m = 1$, puis quand $n = m = 1$. (Solution p. 41.)

Question 3 (••) Spécialiser la règle de différentiation en chaîne (p. 16) quand $p = 1$, puis $n = 1$, puis $m = 1$. (Solution p. 41.)

Fonction quadratique

Soit $A : \mathbb{R}^n \to \mathbb{R}^n$ un opérateur linéaire, b un vecteur de $\mathbb{R}^n$ et $c \in \mathbb{R}$. On considère la fonction $f : \mathbb{R}^n \to \mathbb{R}$ définie par

$$f(x) = \frac{1}{2} \langle x, A \cdot x \rangle + \langle b, x \rangle + c.$$

Question 1 – Gradient (••) Montrer que f différentiable en tout point x de $\mathbb{R}^n$ et calculer $\nabla f(x)$. (Solution p. 42.)

Question 2 – Matrice hessienne (•) Montrer que la fonction $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ est différentiable et calculer la matrice jacobienne de ∇f en x. On note désormais $H_f(x) := J_{\nabla f}(x)$. (Solution p. 42.)

Question 3 – Point critique (•) Soit $x \in \mathbb{R}^n$; on suppose que $H_f(x)$ est inversible. Montrer qu'il existe un unique $x_0 \in \mathbb{R}^n$ où s'annule ∇f ; le calculer en fonction de x, $\nabla f(x)$ et $H_f(x)$. (Solution p. 43.)

Question 4 – Vecteur gaussien (••) La densité de probabilité associée à un vecteur gaussien $X \in \mathbb{R}^d$ est proportionnelle à la fonction

$$g : x \in \mathbb{R}^d \mapsto \exp\left(-\frac{1}{2} \left\langle x, \Sigma^{-1} \cdot x \right\rangle\right)$$

où $\Sigma : \mathbb{R}^d \to \mathbb{R}^d$ est un opérateur linéaire autoadjoint ($\Sigma^\top = \Sigma$) tel que $\langle x, \Sigma \cdot x \rangle > 0$ quand $x \neq 0$.

Montrer que la fonction g est différentiable et calculer son gradient. (Solution p. 43.)

Robot manipulateur

Les coordonnées cartésiennes x et y de l'effecteur final d'un robot dans le plan, composé de deux barres rigides de longueur ℓ_1 et ℓ_2 et d'articulations rotoïdes sont données par

$$\left| \begin{array}{lcl} x & = & \ell_1 \cos\theta_1 + \ell_2 \cos(\theta_1 + \theta_2) \\ y & = & \ell_1 \sin\theta_1 + \ell_2 \sin(\theta_1 + \theta_2) \end{array} \right.$$

où θ_1 et θ_2 sont les coordonnées articulaires du robot.

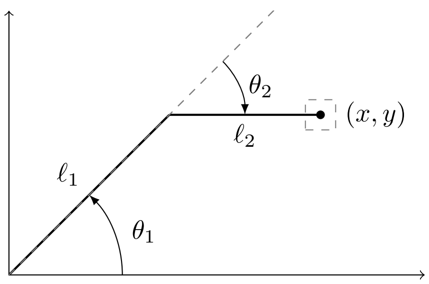

FIGURE 7 – $\ell_1 = 3$, $\ell_2 = 2$, $\theta_1 = \pi/4$, $\theta_2 = -\pi/4$.

On souhaite quantifier quel impact un jeu au niveau des articulations affecte la précision du positionnement de l'effecteur final.

Question 1 (•) Montrer que l'application $f : (\theta_1, \theta_2) \in \mathbb{R}^2 \mapsto (x, y) \in \mathbb{R}^2$ est différentiable et déterminer sa matrice jacobienne. (Solution p. 44.)

Question 2 (••) Soit $(\theta_{10}, \theta_{20}) \in \mathbb{R}^2$ et $(x_0, y_0) = f(\theta_{10}, \theta_{20})$. Montrer que si

$$|\theta_1 - \theta_{10}| \leq \varepsilon \text{ et } |\theta_2 - \theta_{20}| \leq \varepsilon$$

alors $(x, y) = f(\theta_1, \theta_2)$ appartient au carré centré en (x_0, y_0) d'arête de longueur $2(\ell_1 + 2\ell_2)\varepsilon$. (Solution p. 44.)

Dérivée directionnelle d'Hadamard

Source : (Shapiro 1990)

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ et $x \in U$. La fonction f est *directionnellement dérivable* si pour tout vecteur $h \in \mathbb{R}^n$, la dérivée directionnelle

$$f'(x, h) := (t \mapsto f(x + th))'(0) = \lim_{t \to 0} \frac{f(x + th) - f(x)}{t}$$

est bien définie.

On introduit une variante à cette définition : la fonction f est *directionnellement dérivable au sens de Hadamard* en x si pour tout chemin $\gamma : I \subset \mathbb{R} \to \mathbb{R}^n$, défini sur un intervalle ouvert I contenant 0, tel que $\gamma(I) \subset U$, $\gamma(0) = x$ et $\gamma'(0)$ existe, la dérivée $(f \circ \gamma)'(0)$ existe.

Question 1 (•) Montrer que si f est directionnellement dérivable au sens de Hadamard en x, alors f est directionnellement dérivable au sens classique. (Solution p. 45.)

Question 2 (•••) Montrer que si f est directionnellement dérivable au sens de Hadamard en x, la grandeur $(f \circ \gamma)'(0)$ ne dépend de γ qu'à travers $\gamma'(0)$ et que par conséquent

$$(f \circ \gamma)'(0) = f'(x, \gamma'(0)).$$

(Solution p. 45.)

Question 3 – Dérivation en chaîne (••) Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ et $g : V \subset \mathbb{R}^n \to \mathbb{R}^m$ deux fonctions définies sur des ouverts U et V et telles que $f(U) \subset V$. Montrer que si f est directionnellement dérivable au sens de Hadamard en $x \in U$ et que g est directionnellement dérivable au sens de Hadamard en $f(x) \in V$, alors la composée $g \circ f$ est directionnellement dérivable au sens de Hadamard en en x et

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

(Solution p. 46.)

Question 4 (••••) Montrer que f est directionnellement dérivable au sens de Hadamard en x si et seulement si la limite

$$\lim_{(t,k) \to (0,h)} \frac{f(x + tk) - f(x)}{t}$$

existe et que la limite est alors égale à $f'(x, h)$. (Solution p. 46.)

Question 5 (••••) Une fonction dérivable directionnellement au sens de Hadamard en x est *différentiable au sens de Hadamard* en x si de plus $f'(x, h)$ est une fonction linéaire de h. Montrer que f est différentiable en x au sens de Hadamard si et seulement si elle est différentiable en x. (Solution p. 47.)

Thermodynamique

Pour un gaz parfait, la pression P, le volume V, le nombre de particules N et la température T sont reliés par la relation

$$PV = Nk_B T$$

où k_B est la constante de Boltzmann. Si en outre le gaz est mono-atomique, son entropie S est donnée par l'expression [7]

$$S = Nk_B \left[\frac{5}{2} + \ln \left(\frac{V}{N} \frac{(2\pi m k_B T)^{3/2}}{h^3} \right) \right]$$

où h est la constante de Planck et m la masse d'un atome de gaz. On s'intéresse dans la suite à une quantité fixe d'un gaz donné de ce type.

7. cf. par exemple l'article consacré au “Paradoxe de Gibbs” sur Wikipédia.

Question 0 (○) Quelles sont les grandeurs variables (“variables d'état”) associées à cette expression de l'entropie S ? Quelle intervalle de valeurs peuvent prendre ces variables ? (On souhaite que l'entropie soit toujours définie.) (Solution p. 48.)

Question 1 (•) Montrer que la différentielle dS est bien définie et la calculer en utilisant les notations les plus appropriées. (Solution p. 48.)

Question 2 (•••) L'énergie interne U du gaz est une fonction des variables d'état (une “fonction d'état”) ; sa variation infinitésimale est reliée à celle de l'entropie et du volume par la relation

$$dU = TdS - PdV.$$

Quel sens donnez-vous à cette relation mathématiquement ? Pouvez-vous la réécrire en utilisant les variations associées aux variables d'état utilisées précédemment ? (Solution p. 49.)

Question 3 (••) Déduire de la question précédente une expression de l'énergie interne (définie à une constante près). (Solution p. 49.)

Solutions

Exercices essentiels

Dérivée et prolongement Si une fonction g dérivable sur un ouvert U de $\mathbb{R}$ contenant $[a, b]$ prolonge la fonction f définie sur $[a, b]$, il est clair que f est dérivable en tout point de $[a, b]$ et que $g'|_{[a,b]} = f'$.

Réciproquement, si f est dérivable sur $[a, b]$ (à droite en a et à gauche en b), alors la fonction $g :]-\infty, +\infty[\to \mathbb{R}^m$ définie par

$$g(x) = \left| \begin{array}{rl} f(a) + f'(a) \times (x - a) & \text{si } x < a \\ f(x) & \text{si } x \in [a, b] \\ f(b) + f'(b) \times (x - b) & \text{si } x > b \end{array} \right.$$

prolonge la fonction f et est dérivable par construction.

Domaine de définition des fonctions partielles La valeur de $f(x_1, \cdots, x_{j-1}, y_j, x_{j+1}, \cdots, x_n)$ est définie si et seulement si l'argument $(x_1, \cdots, x_{j-1}, y_j, x_{j+1}, \cdots, x_n)$ appartient à U. Le domaine de définition de la fonction partielle $y_j \mapsto f(x_1, \cdots, x_{j-1}, y_j, x_{j+1}, \cdots, x_n)$ est donc l'image réciproque de U par la fonction

$$y_j \in \mathbb{R} \mapsto (x_1, \cdots, x_{j-1}, y_j, x_{j+1}, \cdots, x_n) \in \mathbb{R}^n.$$

Cette fonction étant continue (il s'agit d'une fonction affine) et U ouvert par hypothèse, cet ensemble est bien ouvert.

Gradient et matrice jacobienne Les deux fonctions partielles de la fonction

$$f : (x_1, x_2) \in \mathbb{R}^2 \mapsto (x_2^2 - x_1)^2 + (x_1 - 1)^2 \in \mathbb{R}$$

sont dérivables, vérifient $\partial_1 f(x_1, x_2) = -2(x_2^2 - x_1) + 2(x_1 - 1)$ et $\partial_2 f(x_1, x_2) = 4(x_2^2 - x_1)x_2$. Par conséquent,

$$J_f(x_1, x_2) = \left[\begin{array}{cc} -2(x_2^2 - x_1) + 2(x_1 - 1) & 4(x_2^2 - x_1)x_2 \end{array} \right] \in \mathbb{R}^{1\times 2}.$$

Son gradient est donc donné par

$$\nabla f(x_1, x_2) = (-2(x_2^2 - x_1) + 2(x_1 - 1), 4(x_2^2 - x_1)x_2) \in \mathbb{R}^2$$

ou, représenté comme un vecteur colonne

$$\nabla f(x_1, x_2) = J_f(x_1, x_2)^\top = \left[\begin{array}{c} -2(x_2^2 - x_1) + 2(x_1 - 1) \\ 4(x_2^2 - x_1)x_2 \end{array} \right] \in \mathbb{R}^{2\times 1}.$$

Exercice – Matrice jacobienne (•) La fonction

$$g : (x_1, x_2) \in \mathbb{R}^2 \mapsto (-2(x_2^2 - x_1) + 2(x_1 - 1), 4(x_2^2 - x_1)x_2) \in \mathbb{R}^2.$$

a deux composantes, les fonctions (scalaires)

$$g_1 : (x_1, x_2) \in \mathbb{R}^2 \mapsto -2(x_2^2 - x_1) + 2(x_1 - 1) \in \mathbb{R}$$

et

$$g_2 : (x_1, x_2) \in \mathbb{R}^2 \mapsto 4(x_2^2 - x_1)x_2 \in \mathbb{R}.$$

En tout point $x = (x_1, x_2)$ de $\mathbb{R}^2$, les fonctions partielles de ces deux fonctions existent et vérifient $\partial_1 g_1(x_1, x_2) = 4$, $\partial_2 g_1(x_1, x_2) = -4x_2$, $\partial_1 g_2(x_1, x_2) = -4x_2$ et $\partial_2 g_2(x_1, x_2) = 12x_2^2$. Sa matrice jacobienne est donc définie en tout point et vaut

$$J_g(x_1, x_2) = \left[\begin{array}{cc} 4 & -4x_2 \\ -4x_2 & 12x_2^2 \end{array} \right] \in \mathbb{R}^{2\times 2}.$$

Matrice jacobienne et gradient D'après la définition de la matrice jacobienne (p. 5) et du gradient (p. 6), on a

$$J_f(x) = \left[\begin{array}{ccc} — & \nabla f_1(x)^\top & — \\ — & \nabla f_2(x)^\top & — \\ \vdots & \vdots & \vdots \\ — & \nabla f_m(x)^\top & — \end{array} \right].$$

Fonction discontinue I La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par :

$$f(x, y) = \left| \begin{array}{ll} 0 & \text{si } x = 0 \text{ ou } y = 0, \\ 1 & \text{sinon.} \end{array} \right.$$

est discontinue en $(0, 0)$, car $f(2^{-n}, 2^{-n}) = 1$ pour tout $n \in \mathbb{N}$, donc

$$\lim_{n\to+\infty} f(2^{-n}, 2^{-n}) = 1 \neq 0 = f(0, 0).$$

Par contre, les deux fonctions partielles de f en $(0,0)$

$$x_1 \mapsto f(x_1, 0) \text{ et } x_2 \mapsto f(0, x_2)$$

sont constantes et égales à 0. Les dérivées partielles $\partial_1 f(0,0)$ et $\partial_2 f(0,0)$ existent donc et sont nulles. Le gradient de f en $(0,0)$ est donc définie (et nul).

Fonction discontinue II Les dérivées directionnelles de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \left| \begin{array}{ll} 1 & \text{si } x > 0 \text{ et } y = x^2, \\ 0 & \text{sinon.} \end{array} \right.$$

existent en $(0,0)$ et sont nulles pour tout $h \in \mathbb{R}^2$, puisque les fonctions associées $t \in \mathbb{R} \mapsto f(th)$ sont nulles pour $|t|$ suffisamment petit. Mais f n'est pas continue en l'origine ; elle n'y est donc a fortiori pas différentiable.

Fonctions affines Comme $f(x) = A \cdot x + b$, la i-ème composante de f satisfait

$$f_i(x) = \sum_{k=1}^{n} A_{ik} x_k + b_i$$

et donc la j-ème fonction partielle de f_i en x est la fonction de la variable y_j dont la valeur est

$$f_i(x_1, \dots, x_{j-1}, y_i, x_{j+1}, \dots, x_n) = \\ A_{i1}x_1 + \dots + A_{i,j-1}x_{j-1} + A_{ij}y_j + A_{i,j+1}x_{j+1} + b_i.$$

C'est une fonction affine de y_j qui est dérivable, de dérivée A_{ij}. On a donc $\partial_j f_i(x) = A_{ij}$; la matrice jacobienne $J_f(x)$ existe en tout point $x \in \mathbb{R}^n$ et vérifie $J_f(x) = A$.

Pour tout $x \in \mathbb{R}^n$ et $h \in \mathbb{R}^n$, on a donc

$$f(x+h) - f(x) - J_f(x) \cdot h = A \cdot (x+h) - A \cdot x - A \cdot h = 0,$$

ce que l'on peut écrire sous la forme

$$f(x+h) = f(x) + J_f(x) \cdot h + \varepsilon(h)\|h\|$$

en prenant pour fonction ε la fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$ identiquement nulle. La fonction f est donc différentiable sur $\mathbb{R}^n$.

Développement limité au premier ordre De l'hypothèse

$$f(x+h) = a + B \cdot h + \varepsilon(h)\|h\|$$

on peut déduire en faisant tendre h vers 0 que $a = f(x)$, puis que pour tout $i \in \{1, \dots, n\}$, on a

$$f_i(x+h) = f_i(x) + [B \cdot h]_i + \varepsilon_i(h)\|h\|$$

et donc

$$\begin{aligned}\frac{f_i(x+te_j)-f_i(x)}{t} &= \frac{f_i(x)+[B\cdot te_j]_i+\varepsilon_i(te_j)\|te_j\|-f_i(x)}{t}\\ &= [B\cdot e_j]_i+\varepsilon_i(te_j)\\ &= B_{ij}+\varepsilon_i(te_j).\end{aligned}$$

Comme $\lim_{h\to 0}\varepsilon(h)=\varepsilon(0)=0$, cette expression a une limite quand $t\to 0$, donc $\partial_j f_i(x)$ existe et vérifie

$$\partial_j f_i(x)=\lim_{t\to 0}\frac{f_i(x+te_j)-f_i(x)}{t}=B_{ij}.$$

La matrice jacobienne $J_f(x)$ de f en x existe donc et est égale à B.

Différentiabilité implique continuité Comme f est différentiable en x, sa matrice jacobienne en x existe et

$$f(x+h)=f(x)+J_f(x)\cdot h+\varepsilon(h)\|h\|$$

avec $\lim_{h\to 0}\varepsilon(h)=0$. Or $\|J_f(x)\cdot h\|\leq\|J_f(x)\|\|h\|$ donc $\lim_{h\to 0}J_f(x)\cdot h=0$; de même $\lim_{h\to 0}\varepsilon(h)\|h\|=0$. Par conséquent, $f(x+h)\to f(x)$ quand $h\to 0$.

Différentiabilité Si f est différentiable en x, alors la matrice jacobienne de f en x existe et il existe sur un voisinage de $h=0$ une fonction ε à valeurs dans $\mathbb{R}^m$ vérifiant $\lim_{h\to 0}\varepsilon(h)=0$ telle que

$$f(x+h)=f(x)+J_f(x)\cdot h+\varepsilon(h)\|h\|.$$

On en déduit que sur ce voisinage de 0 (et hormis pour $h=0$), on a

$$\varepsilon(h)=\left(\frac{f(x+h)-f(x)}{\|h\|}-J_f(x)\cdot\frac{h}{\|h\|}\right).$$

Par hypothèse la fonction ε est continue et nulle en 0, donc le membre de droite de cette équation tend bien vers 0 quand $h\to 0$ avec $h\neq 0$. Réciproquement, si

$$\lim_{\substack{h\to 0\\ h\neq 0}}\left(\frac{f(x+h)-f(x)}{\|h\|}-J_f(x)\cdot\frac{h}{\|h\|}\right)=0,$$

on peut définir sur l'ensemble V des h tels que $x+h\in U$ (U étant le domaine de définition de f) la fonction ε par

$$\varepsilon(h)=\left|\begin{array}{ll} 0 & \text{si } h=0,\\ \dfrac{f(x+h)-f(x)}{\|h\|}-J_f(x)\cdot\dfrac{h}{\|h\|} & \text{si } h\in V \text{ et } h\neq 0.\end{array}\right.$$

Par construction, V est un voisinage de 0 et ε vérifie $\lim_{h\to 0}\varepsilon(h)=0$ ainsi que la relation $f(x+h)=f(x)+J_f(x)\cdot h+\varepsilon(h)\|h\|$; f est donc différentiable en x.

Espace-temps Etant donné le contexte ("Espace-temps"), il est probable que c désigne la vitesse de la lumière dans le vide et soit donc considérée comme une constante. L'expression $e := x^2 + y^2 + z^2 - c^2t^2$ est définie pour toutes les valeurs possibles des variables réelles x, y, z et t. Il est clair que toutes les dérivées partielles de cette expression existent et sont continues en tout point; l'application des conventions de la section "Notations" (p. 13) fournit

$$de = \frac{\partial e}{\partial x}dx + \frac{\partial e}{\partial y}dy + \frac{\partial e}{\partial z}dz + \frac{\partial e}{\partial t}dt = 2(x\,dx + y\,dy + z\,dz - c^2t\,dt).$$

Robustesse des notations Si l'on considère successivement les jeux de variables (y, x) puis de (x, y, z), on obtient

$$dy \cdot (h_y, h_x) = h_y,\ dx \cdot (h_y, h_x) = h_x$$

et

$$dx \cdot (h_x, h_y, h_z) = h_x,\ dy \cdot (h_x, h_y, h_z) = h_y,\ dz \cdot (h_x, h_y, h_z) = h_z.$$

Comme on a toujours

$$\frac{\partial(x^2 + y^2)}{\partial x} = 2x,\ \frac{\partial(x^2 + y^2)}{\partial y} = 2y,\ \frac{\partial(x^2 + y^2)}{\partial z} = 0,$$

on constate qu'on a à nouveau $d(x^2 + y^2) = 2x\,dx + 2y\,dy$ dans les deux cas.

Composition de fonctions continûment différentiables D'après la règle de différentiation en chaîne (p. 16), $d(g \circ f)(x) = dg(f(x)) \cdot df(x)$. Donc pour tout $i \in \{1, \dots, m\}$ et $j \in \{1, \dots, p\}$,

$$\begin{aligned}
[d(g \circ f)(x)]_{ij} &= [dg(f(x)) \cdot df(x)]_{ij} \\
&= \sum_{k=1}^{n} [dg(f(x))]_{ik} [df(x)]_{kj} \\
&= \sum_{k=1}^{n} \partial_k g_i(f(x)) \partial_j f_k(x).
\end{aligned}$$

Chaque coefficient $\partial_j (g \circ f)_i$ est une somme de produit de fonctions continues et est donc continu. Par conséquent, $g \circ f$ est continûment différentiable.

Désassemblage Les fonctions $p_1 : (x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^p \mapsto x_1 \in \mathbb{R}^m$ et $p_2 : (x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^p \mapsto x_2 \in \mathbb{R}^p$ sont linéaires, et donc différentiables en tout point x (p. 15) ; leurs différentielles sont données par $dp_1(x_1, x_2) = p_1$ et $dp_2(x_1, x_2) = p_2$. Par conséquent, si U est un ouvert de $\mathbb{R}^n$, $x \in U$ et que la fonction $(f, g) : U \to \mathbb{R}^m \times \mathbb{R}^p$ est différentiable en x, en appliquant à deux reprises la règle de différentiation en chaîne (p. 16), on en déduit que $f = p_1 \circ (f, g)$ et que $g = p_2 \circ (f, g)$ sont différentiables et que $df(x) = p_1 \cdot d(f, g)(x)$ et $dg(x) = p_2 \cdot d(f, g)(x)$, soit $d(f, g)(x) = (df(x), dg(x))$.

Différentiation en chaîne et assemblage La combinaison des deux résultats prend la forme suivante :

> Soient $f_1 : U \subset \mathbb{R}^p \to \mathbb{R}^{n_1}$, …, $f_m : U \subset \mathbb{R}^p \to \mathbb{R}^{n_m}$ et $g : V \subset \mathbb{R}^n \to \mathbb{R}^m$ – où $n = n_1 + \cdots + n_m$ – des fonctions définies sur des ouverts U et V et telles que $(f_1, \ldots, f_m)(U) \subset V$. Si les f_i sont différentiables en $x \in U$ et que g est différentiable en $(f_1(x), \ldots, f_m(x)) \in V$, alors la composée $g \circ (f_1, \ldots, f_m)$ est différentiable en x et
>
> $$d(g \circ (f_1, \ldots, f_m))(x) = dg(f_1(x), \cdots, f_m(x)) \cdot (df_1(x), \ldots, df_m(x)).$$

Le cas $m = 1$ de cet énoncé correspond à la règle de différentiation en chaîne (p. 16) ; le cas où g est la fonction identité correspond à la règle d'assemblage (p. 17).

Produit scalaire On a $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$. Chaque fonction $(x, y) \mapsto x_i y_i$ est (continûment) différentiable, avec $d(x_i y_i) = x_i dy_i + y_i dx_i$ (par calcul direct des dérivées partielles ; alternativement, on peut combiner le désassemblage (p. 17) de $(x, y) \mapsto (x_i, y_i)$ et la règle du produit (p. 15) par différentiation en chaîne (p. 16)). Par linéarité de la différentielle (p. 18), le produit scalaire est donc différentiable et

$$d\left\langle x, y \right\rangle = \sum_{i=1}^n x_i dy_i + y_i dx_i = \begin{bmatrix} y_1 & \cdots & y_n & x_1 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \\ dy_1 \\ \vdots \\ dy_n \end{bmatrix}.$$

Par conséquent, $\nabla \left\langle \cdot, \cdot \right\rangle (x, y) = (y, x)$.

Cas des fonctions continûment différentiables Si f est continûment différentiable sur U, elle est en particulier différentiable sur $[x, a + h]$. De plus,

$$df(x + th) \cdot h = \sum_{i=1}^n \partial_j f(x + th) h_j,$$

donc la fonction $t \in [0, 1] \mapsto df(x + th) \cdot h$ est continue et par conséquent intégrable. Le théorème fondamental du calcul (p. 20) est donc applicable.

Cas des fonctions continûment différentiables Si la fonction f' est continue (ou intégrable au sens de Riemann, ou intégrable au sens de Lebesgue), le théorème fondamental du calcul (p. 21) est applicable, donc

$$f(x + h) - f(x) = \int_x^{x+h} f'(t)\, dt.$$

L'inégalité triangulaire appliquée à l'intégrale du membre de droite fournit alors

$$\|f(x + h) - f(x)\| = \left\| \int_x^{x+h} f'(t)\, dt \right\| \leq \int_x^{x+h} \|f'(y)\|\, dy \leq \int_x^{x+h} M\, dt = Mh.$$

Inégalité des accroissements finis (version euclidienne) Comme

$$\frac{\phi(y+s)-\phi(y)}{s} = \left\langle f(x+h)-f(x), \frac{f(y+s)-f(y)}{s} \right\rangle,$$

la fonction ϕ est dérivable en tout point $y \in [a, a+h]$ et

$$\phi'(y) = \langle f(x+h)-f(x), f'(y) \rangle .$$

La fonction f étant à valeurs réelles, le théorème des valeurs intermédiaires est applicable : il existe un $y \in [x, x+h]$ tel que

$$\phi(x+h)-\phi(x) = \phi'(y)h = \langle f(x+h)-f(x), f'(y) \rangle h.$$

Comme par ailleurs

$$\begin{aligned}\phi(x+h)-\phi(x) &= \langle f(x+h)-f(x), f(x+h) \rangle - \langle f(x+h)-f(x), f(x) \rangle \\ &= \|f(x+h)-f(x)\|^2,\end{aligned}$$

on a

$$\|f(x+h)-f(x)\|^2 = \langle f(x+h)-f(x), f'(y) \rangle h \leq \|f(x+h)-f(x)\| \|f'(y)\| h.$$

Etant donné que $\|f'(y)\| \leq M$, on en déduit $\|f(x+h)-f(x)\| \leq Mh$.

Inégalité de la valeur moyenne Soit $F : [a,b] \to \mathbb{R}^m$ une primitive de f. Par le théorème fondamental du calcul (p. 19), on a

$$\langle f \rangle = \frac{1}{b-a} \int_a^b f(x)\, dx = \frac{F(b)-F(a)}{b-a}.$$

Or si $\|F'\| = \|f\|$ est borné sur $[a,b]$, par l'inégalité des accroissements finis (p. 21),

$$\|F(b)-F(a)\| \leq \sup_{x \in [a,b]} \|f(x)\| \times (b-a),$$

et donc

$$\|\langle f \rangle\| \leq \sup_{x \in [a,b]} \|f(x)\|.$$

Il va de soi que cette inégalité reste vérifiée si $\|f\|$ est non-bornée, c'est-à-dire si $\sup_{x \in [a,b]} \|f(x)\| = +\infty$.

Egalité des accroissements finis ? La dérivée de f est donnée par $f'(t) = (-\sin t, \cos t)$;
en particulier pour tout $t \in [0, 2\pi]$, $\|f'(t)\| = 1$. Or $f(2\pi) - f(0) = 0$, donc il est impossible de trouver un t tel que $f(2\pi) - f(0) = f'(t) \times 2\pi$.

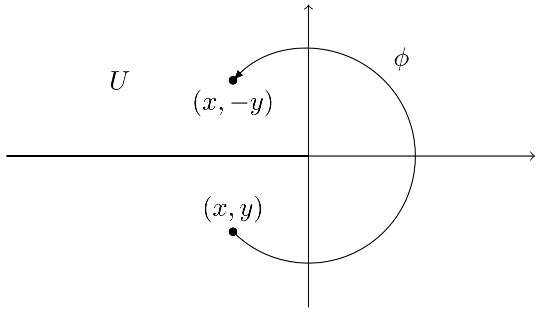

FIGURE 8 – Représentation du chemin ϕ quand $(x, y) = (-1, -1)$.

Variation du logarithme En premier lieu, on peut noter qu'en général, le segment d'extrémités (x, y) et $(x, -y)$ n'est pas inclus dans le plan coupé U. On ne peut donc pas appliquer directement la version multivariable de l'inégalité des accroissements finis (p. 20). Nous allons donc adapter la technique utilisée dans la démonstration de ce théorème en introduisant un chemin $\phi : [0, 1] \to \mathbb{R}^2$ qui joint (x, y) et $(x, -y)$ et dont l'image est incluse dans U.

On note θ l'unique détermination de l'angle polaire de (x, y) comprise dans l'intervalle $]-\pi, \pi[$ et on pose

$$\phi(t) := (r \cos((1 - 2t)\theta), r \sin((1 - 2t)\theta)) \quad \text{où} \quad r := \sqrt{x^2 + y^2}.$$

On remarque que $\phi(t) \in U$ pour tout t, que $\phi(0) = (x, y)$ et que $\phi(1) = (x, -y)$. La fonction ϕ est dérivable et

$$\phi'(t) = (2\theta r \sin((1 - 2t)\theta), -2\theta r \cos((1 - 2t)\theta))\,;$$

la fonction $\log \circ \phi$ est donc définie et dérivable et

$$d(\log \circ \phi)(t) = d \log(\phi(t)) \cdot d\phi(t) = d \log(\phi(t)) \cdot \phi'(t).$$

Par conséquent,

$$\|d(\log \circ \phi)(t)\| \leq \|d \log(\phi(t))\| \|\phi'(t)\| \leq \frac{1}{r} \times 2|\theta| r = 2|\theta| \leq 2\pi.$$

L'application de l'inégalité des accroissement finis (monovariable) (p. 21) à la fonction $\log \circ \phi$ fournit donc l'inégalité

$$\| \log(x, y) - \log(x, -y)\| \leq 2\pi.$$

Spécialisation de la différentiation en chaîne

Question 1 Comme fonctions d'une variable, f et g sont différentiables donc dérivables (p. 12), $df(x) \cdot h = f'(x)h$ et $dg(x) \cdot h = g'(x)h$. Par la règle de différentiation en chaîne (p. 16), on obtient

$$d(g \circ f)(x) = dg(f(x)) \cdot df(x).$$

On en déduit

$$\begin{aligned}
d(g \circ f)(x) \cdot h &= (dg(f(x)) \cdot df(x)) \cdot h \\
&= dg(f(x)) \cdot (df(x) \cdot h) \\
&= dg(f(x)) \cdot (f'(x)h) \\
&= g'(f(x))(f'(x)h) \\
&= (g'(f(x))f'(x))h.
\end{aligned}$$

La fonction $g \circ f$ étant également fonction d'une variable, elle est dérivable et

$$(g \circ f)'(x) = d(g \circ f)(x) \cdot 1 = g'(f(x))f'(x).$$

Question 2 La fonction f dépendant d'une variable scalaire, elle est dérivable et $df(x) \cdot h = f'(x)h$. Quant à g qui est à valeur scalaire, sa différentielle en x est reliée à son gradient par $dg(x) \cdot h = \langle \nabla g(x), h \rangle$ (p. 12). La règle de différentiation en chaîne (p. 16), $d(g \circ f)(x) = dg(f(x)) \cdot df(x)$ se décline donc ici en

$$\begin{aligned}
d(g \circ f)(x) \cdot h &= dg(f(x)) \cdot (df(x) \cdot h) \\
&= dg(f(x)) \cdot (f'(x)h) \\
&= \langle \nabla g(f(x), f'(x)h \rangle \\
&= \langle \nabla g(f(x)), f'(x) \rangle \, h,
\end{aligned}$$

soit $(g \circ f)'(x) = d(g \circ f)(x) \cdot 1 = \langle \nabla g(f(x)), f'(x) \rangle$. Quand $n = m = 1$, on a

$$\begin{aligned}
d(g \circ f)(x) \cdot h &= dg(f(x)) \cdot (df(x) \cdot h) \\
&= g'(f(x))(df(x) \cdot h) \\
&= g'(f(x)) \, \langle \nabla f(x), h \rangle \\
&= \langle g'(f(x)) \nabla f(x), h \rangle
\end{aligned}$$

donc $\nabla(g \circ f)(x) = g'(f(x))\nabla f(x)$.

Question 3 Quand $p = 1$, on a

$$\begin{aligned}
d(g \circ f)(x) \cdot h &= dg(f(x)) \cdot (df(x) \cdot h) \\
&= dg(f(x)) \cdot (f'(x)h) \\
&= (dg(f(x)) \cdot f'(x))h.
\end{aligned}$$

donc $(g \circ f)'(x) = dg(f(x)) \cdot f'(x)$. Quand $n = 1$, on a

$$\begin{aligned}
d(g \circ f)(x) \cdot h &= dg(f(x)) \cdot (df(x) \cdot h) \\
&= g'(f(x)) \, \langle \nabla f(x), h \rangle \\
&= g'(f(x)) \nabla f(x)^{\top} h.
\end{aligned}$$

donc $d(g \circ f)(x) = g'(f(x))\nabla f(x)^{\top}$. Finalement, quand $m = 1$, on a

$$\begin{aligned}
d(g \circ f)(x) \cdot h &= dg(f(x)) \cdot (df(x) \cdot h) \\
&= \langle \nabla g(f(x)), df(x) \cdot h \rangle \\
&= \left\langle (df(x))^{\top} \cdot \nabla g(f(x)), \cdot h \right\rangle
\end{aligned}$$

et donc $\nabla(g \circ f)(x) = (df(x))^{\top} \cdot \nabla g(f(x))$.

Fonction quadratique

Question 1 – Gradient La fonction présentée est la somme de la fonction $x \mapsto \langle x, A \cdot x \rangle = x^\top \cdot A \cdot x$, de l'application linéaire $x \mapsto \langle b, x \rangle = b^\top x$ et de l'application constante $x \mapsto c$. Les deux dernières fonctions sont différentiables, de différentielles les fonctions $x \mapsto b^\top$ (comme différentielle d'application linéaire (p. 15)) et $x \mapsto 0$ respectivement. Seul reste donc à étudier la fonction $g : x \mapsto x^\top \cdot A \cdot x$.

Comme

$$g(x) = \frac{1}{2} x^\top \cdot A \cdot x = \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} x_i A_{ik} x_k,$$

pour tout $j \in \{1, \ldots, n\}$ on a

$$g(x) = \frac{1}{2} \left(x_j A_{jj} x_j + \sum_{\substack{k=1 \\ k \neq j}}^{n} x_j A_{jk} x_k + \sum_{\substack{i=1 \\ i \neq j}}^{n} x_i A_{ij} x_j + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{\substack{k=1 \\ k \neq j}}^{n} x_i A_{ik} x_k \right).$$

Par conséquent, la dérivée partielle $\partial_j g(x)$ existe et vérifie

$$\begin{aligned}
\partial_j g(x) &= A_{jj} x_j + \frac{1}{2} \left(\sum_{\substack{k=1 \\ k \neq j}}^{n} A_{jk} x_k + \sum_{\substack{i=1 \\ i \neq j}}^{n} x_i A_{ij} \right) \\
&= \frac{1}{2} \sum_{k=1}^{n} A_{jk} x_k + \frac{1}{2} \sum_{i=1}^{n} x_i A_{ij} \\
&= \left[\frac{1}{2} A \cdot x \right]_j + \left[\frac{1}{2} A^\top \cdot x \right]_j \\
&= \left[\frac{1}{2} (A + A^\top) \cdot x \right]_j .
\end{aligned}$$

Toute ces dérivées partielles sont des fonctions linéaires de x, elles sont donc continues et la fonction g est continûment différentiable ; elle est donc différentiable (p. 8). De plus, l'égalité ci-dessus nous fournit

$$\nabla g(x) = \frac{1}{2} (A + A^\top) \cdot x.$$

La fonction f est donc différentiable (comme combinaison linéaire de fonctions différentiables (p. 18)) et

$$\nabla f(x) = \frac{1}{2} (A + A^\top) \cdot x + b^\top.$$

Question 2 – Matrice hessienne La fonction

$$\nabla f : x \in \mathbb{R}^n \mapsto \frac{1}{2} (A + A^\top) \cdot x + b^\top.$$

est affine, somme d'une fonction linéaire et d'une fonction constante. Elle est donc différentiable, et

$$H_f(x) := J_{\nabla f}(x) = \frac{1}{2} (A + A^\top).$$

Question 3 – Point critique Si $H_f(x)$ est inversible (cet opérateur est constant), comme

$$\nabla f(y) = \frac{1}{2}(A + A^\top) \cdot y + b = H_f(x) \cdot y + b,$$

résoudre $\nabla f(x_0) = 0$ revient à rechercher les solutions de

$$H_f(x) \cdot x_0 + b = H_f(x) \cdot x_0 + (\nabla f(x) - H_f(x) \cdot x) = 0.$$

Il existe donc un unique zéro de ∇f, donné par

$$x_0 = x - (H_f(x))^{-1} \nabla f(x).$$

Question 4 – Vecteur gaussien La fonction

$$g : x \in \mathbb{R}^d \mapsto \exp\left(-\frac{1}{2}\left\langle x, \Sigma^{-1} \cdot x \right\rangle\right)$$

apparaît comme la composée des fonctions

$$x \in \mathbb{R}^d \mapsto -\frac{1}{2}\left\langle x, \Sigma^{-1} \cdot x \right\rangle \quad \text{et} \quad \exp : \mathbb{R} \to \mathbb{R}.$$

La fonction exp est dérivable, et donc différentiable sur tout $\mathbb{R}$ avec $d(\exp(y)) = \exp'(y)dy = \exp(y)dy$, c'est-à-dire

$$d\exp(y) \cdot h = \exp(y)h.$$

Quant à la première fonction, d'après les questions qui précèdent, elle est différentiable et

$$\nabla\left(-\frac{1}{2}\left(\left\langle x, \Sigma^{-1} \cdot x \right\rangle\right)\right) = \frac{1}{2}\left(-\Sigma^{-1} - (\Sigma^{-1})^\top\right) \cdot x = -\Sigma^{-1} \cdot x.$$

Sa différentielle vérifie donc

$$d\left(-\frac{1}{2}\left(\left\langle x, \Sigma^{-1} \cdot x \right\rangle\right)\right) \cdot h = -\left\langle \Sigma^{-1} \cdot x, h \right\rangle.$$

Par conséquent, la fonction g est différentiable sur $\mathbb{R}^d$ comme composée de fonctions différentiables et

$$dg(x) \cdot h = -\exp\left(-\frac{1}{2}\left(\left\langle x, \Sigma^{-1} \cdot x \right\rangle\right)\right)\left\langle \Sigma^{-1} \cdot x, h \right\rangle = \left\langle -g(x) \times \Sigma^{-1} \cdot x, h \right\rangle,$$

le gradient de g vaut donc

$$\nabla g(x) = -g(x) \times \Sigma^{-1} \cdot x.$$

Robot manipulateur

Question 1 Des équations

$$\left| \begin{array}{lcl} x & = & \ell_1 \cos \theta_1 + \ell_2 \cos(\theta_1 + \theta_2) \\ y & = & \ell_1 \sin \theta_1 + \ell_2 \sin(\theta_1 + \theta_2) \end{array} \right.$$

on déduit que les dérivées partielles de x et de y par rapport à θ_1 et θ_2 existent et vérifient

$$\begin{array}{lcl} \partial x / \partial \theta_1 & = & -\ell_1 \sin \theta_1 - \ell_2 \sin(\theta_1 + \theta_2), \\ \partial x / \partial \theta_2 & = & -\ell_2 \sin(\theta_1 + \theta_2), \\ \partial y / \partial \theta_1 & = & \ell_1 \cos \theta_1 + \ell_2 \cos(\theta_1 + \theta_2), \\ \partial y / \partial \theta_2 & = & \ell_2 \cos(\theta_1 + \theta_2). \end{array}$$

Ces grandeurs étant continues, la fonction f est continûment différentiable et donc différentiable. Si l'on note $s_1 = \sin \theta_1$, $s_{12} = \sin(\theta_1 + \theta_2)$, $c_1 = \cos \theta_1$ et $c_{12} = \cos(\theta_1 + \theta_2)$, on obtient donc

$$J_f(\theta_1, \theta_2) = \left[\begin{array}{cc} -\ell_1 s_1 - \ell_2 s_{12} & -\ell_2 s_{12} \\ \ell_1 c_1 + \ell_2 c_{12} & \ell_2 c_{12} \end{array} \right].$$

Question 2 Soient $\delta\theta_1 := \theta_1 - \theta_{10}$ et $\delta\theta_2 := \theta_2 - \theta_{20}$. La fonction

$$\phi : t \in [0, 1] \mapsto f(\theta_{10} + t\delta\theta_1, \theta_{12} + t\delta\theta_2)$$

est continûment dérivable, de dérivée

$$\phi'(t) = df(\theta_{10} + t\delta\theta_1, \theta_{20} + t\delta\theta_2) \cdot (\delta\theta_1, \delta\theta_2)$$

Si l'on note $s_1(t) = \sin(\theta_{10} + t\delta\theta_1)$, $c_1(t) = \cos(\theta_{10} + t\delta\theta_1)$, $\ldots$, le théorème fondamental du calcul (p. 19) et l'expression de la matrice jacobienne de f nous fournissent donc

$$f(\theta_1, \theta_2) - f(\theta_{10}, \theta_{20}) = \int_0^1 \phi'(t)\, dt = \int_0^1 \left[\begin{array}{c} (-\ell_1 s_1(t) - \ell_2 s_{12}(t))\delta\theta_1 - \ell_2 s_{12}(t)\delta\theta_2 \\ (\ell_1 c_1(t) + \ell_2 c_{12}(t))\delta\theta_1 + \ell_2 c_{12}(t)\delta\theta_2 \end{array} \right] dt$$

et donc par inégalité triangulaire et majoration de l'intégrande,

$$|x - x_0| = |f_1(\theta_1, \theta_2) - f_1(\theta_{10}, \theta_{20})| \leq (\ell_1 + \ell_2)|\delta\theta_1| + \ell_2|\delta\theta_2|$$

ainsi que

$$|y - y_0| = |f_2(\theta_1, \theta_2) - f_2(\theta_{10}, \theta_{20})| \leq (\ell_1 + \ell_2)|\delta\theta_1| + \ell_2|\delta\theta_2|.$$

Si $|\delta\theta_1| \leq \varepsilon$ et $|\delta\theta_2| \leq \varepsilon$, on en déduit

$$|x - x_0| \leq (\ell_1 + 2\ell_2)\varepsilon \ \text{ et } \ |y - y_0| \leq (\ell_1 + 2\ell_2)\varepsilon.$$

Le point $(x, y) = f(\theta_1, \theta_2)$ appartient donc au carré centré en (x_0, y_0) d'arête de longueur $2(\ell_1 + 2\ell_2)\varepsilon$.

Dérivée directionnelle d'Hadamard

Question 1 Supposons que f soit directionnellement dérivable au sens de Hadamard en x. Pour tout $h \in \mathbb{R}^n$, par continuité de l'application $t \in \mathbb{R} \mapsto x+th$, pour $\varepsilon > 0$ assez petit, et parce que le domaine de définition de f est ouvert, l'image de la fonction

$$\gamma : t \in \left]-\varepsilon, \varepsilon\right[\mapsto x + th$$

est incluse dans le domaine de définition de f, est telle que $\gamma(0) = x$, $\gamma'(0) = h$. Par conséquent, la dérivée de $f \circ \gamma$ en 0 existe, et c'est par construction la dérivée directionnelle de f en x dans la direction h. La fonction f est donc directionnellement dérivable en x au sens classique.

Question 2 Supposons que f soit directionnellement dérivable au sens de Hadamard en x. Pour montrer que l'expression $(f \circ \gamma)'(0)$ ne dépend de γ qu'à travers $\gamma'(0)$, nous allons considérer un second chemin arbitraire $\beta : J \to \mathbb{R}^n$, où J est un intervalle ouvert de $\mathbb{R}$ contenant 0, tel que $\beta(0) = x$, $\beta'(0) = \gamma'(0)$ et montrer que

$$(f \circ \gamma)'(0) = (f \circ \beta)'(0).$$

L'idée de la démonstration consiste à construire un troisième chemin α qui en "mélangeant" les chemins β et γ, satisfait les hypothèses de la définition de "directionnellement dérivable au sens de Hadamard", est tel que $\alpha'(0) = \beta'(0) = \gamma'(0)$ et également tel que d'une part $(f \circ \alpha)'(0) = (f \circ \beta)'(0)$ et d'autre part $(f \circ \alpha)'(0) = (f \circ \gamma)'(0)$.

Un chemin qui permette de tenir ce raisonnement est le suivant. Tout d'abord, choisissons $\varepsilon > 0$ tel que $\left]-\varepsilon, \varepsilon\right[\subset I \cap J$, puis définissons $\alpha : \left]-\varepsilon, \varepsilon\right[\to \mathbb{R}^n$ par

$$\alpha(t) = \left| \begin{array}{cl} x & \text{si } \ t = 0, \\ \beta(t) & \text{si } \ \varepsilon/2^{2k+1} \leq |t| < \varepsilon/2^{2k}, \text{ pour un entier } \ k \in \mathbb{N}, \\ \gamma(t) & \text{si } \ \varepsilon/2^{2k+2} \leq |t| < \varepsilon/2^{2k+1}, \text{ pour un entier } \ k \in \mathbb{N}. \end{array} \right.$$

Les hypothèses de la définition sont facilement vérifiées, ainsi que la preuve que $\alpha'(0) = \beta'(0) = \gamma'(0)$. Avec l'hypothèse de différentiabilité au sens de Hadamard, nous savons donc que la dérivée $(f \circ \alpha)'(0)$ existe. On peut la calculer comme la limite de

$$(f \circ \alpha)'(0) = \lim_{k \to +\infty} \frac{f(\alpha(t_k)) - f(x)}{t_n}$$

où t_k est une suite arbitraire de valeurs non nulles tendant vers 0. Or, si l'on choisit $t_k = \varepsilon/2^{2k+1}$, on trouve

$$\lim_{k \to +\infty} \frac{f(\alpha(t_k)) - f(x)}{t_k} = \lim_{k \to +\infty} \frac{f(\beta(t_k)) - f(x)}{t_k} = (f \circ \beta)'(0)$$

et si l'on choisit $t_k = \varepsilon/2^{2k+2}$, on trouve

$$\lim_{k \to +\infty} \frac{f(\alpha(t_k)) - f(x)}{t_k} = \lim_{k \to +\infty} \frac{f(\gamma(t_k)) - f(x)}{t_k} = (f \circ \gamma)'(0),$$

ce qui prouve le résultat d'indépendance souhaité. Pour prouver que $(f \circ \gamma)'(0) = f'(x, \gamma'(0))$, il suffit d'associer à un chemin quelconque γ le chemin "canonique"

$\beta : t \mapsto x + t\gamma'(0)$ de la question 1, qui est tel que $\beta'(0) = \gamma'(0)$ d'une part et d'autre part $(f \circ \beta)'(0) = f'(x, \beta'(0))$ par construction. On en déduit que

$$(f \circ \gamma)'(0) = (f \circ \beta)'(0) = f'(x, \beta'(0)) = f'(x, \gamma'(0)).$$

Question 3 Soit $\gamma : I \subset \mathbb{R} \to \mathbb{R}^n$, un chemin défini sur un intervalle ouvert I contenant 0, tel que $\gamma(I) \subset U$, $\gamma(0) = x$ et $\gamma'(0)$ existe. Alors, sous les hypothèses du théorème de dérivée en chaîne que nous souhaitons montrer, le chemin $\beta = f \circ \gamma$ est défini sur I, vérifie $\beta(I) \subset V$, $\beta(0) = f(x)$ et par hypothèse de dérivabilité directionnelle au sens de Hadamard sur f en x, $\beta'(0) = f'(x, \gamma'(0))$. Par hypothèse de dérivabilité directionnelle au sens de Hadamard sur g en $f(x)$,

$$((g \circ f) \circ \gamma)'(0) = (g \circ \beta)'(0) = g'(f(x), \beta'(0)) = g'(f(x), f'(x, \gamma'(0))),$$

ce qui prouve la dérivabilité directionnelle au sens de Hadamard pour la composée $g \circ f$ en x. Il suffit d'associer à un vecteur h le chemin canonique $t \mapsto x + th$ pour obtenir la relation

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

Question 4 Tout d'abord, si la limite

$$\lim_{(t,k) \to (0,h)} \frac{f(x + tk) - f(x)}{t}$$

existe, elle est égale à la limite obtenue en fixant $k = h$

$$\lim_{t \to 0} \frac{f(x + th) - f(x)}{t}$$

qui est par définition $f'(x, h)$.

Supposons que cette limite existe et montrons que f a une dérivée directionnelle au sens de Hadamard. Soit γ un chemin satisfaisant les hypothèses de cette définition. La fonction $f \circ \gamma$ est dérivable en 0 si et seulement si le taux d'accroissement associé converge en 0. Or, ce taux d'accroissement peut s'écrire sous la forme

$$\frac{f(\gamma(t)) - f(\gamma(0))}{t} = \frac{f\left(x + t\frac{\gamma(t) - \gamma(0)}{t}\right) - x}{t}.$$

Le chemin γ étant dérivable en 0,

$$k(t) := \frac{\gamma(t) - \gamma(0)}{t} \to \gamma'(0) \text{ quand } t \to 0$$

donc par hypothèse, le taux d'accroissement de $f \circ \gamma$ a une limite en 0.

Réciproquement, suppose que f soit directionnellement dérivable au sens de Hadamard en 0. Pour montrer que la limite

$$\lim_{(t,k) \to (0,h)} \frac{f(x + tk) - f(x)}{t}$$

existe, il nous suffit de montrer que pour toute suite t_i de valeurs non nulles tendant vers 0 et toute suite de vecteurs k_i convergeant vers h, la limite

$$\lim_{i \to +\infty} \frac{f(x + t_i k_i) - f(x)}{t_i}$$

existe. On peut imposer la restriction que la suite $|t_i|$ soit strictement décroissante et le résultat reste valable.

Pour tout $t \in \mathbb{R}^*$ notons $j(t)$ le plus petit parmi les entiers j satisfaisant

$$|t - t_j| = \min_{i \in \mathbb{N}} |t - t_i|,$$

puis définissons $\gamma(t)$ par $\gamma(0) = x$ et si $t \neq 0$,

$$\gamma(t) = x + t k_{j(t)}.$$

S'il est défini sur un intervalle $]-\varepsilon, \varepsilon[$ assez petit, γ satisfait les hypothèses de la dérivabilité directionnelle. Le point critique à vérifier est que γ est dérivable en 0. Mais par construction

$$\frac{\gamma(t) - \gamma(0)}{t} = k_{j(t)}$$

et $j(t)$ tend vers $+\infty$ quand t tend vers 0; par conséquent la limite existe et

$$\lim_{t \to 0} \frac{\gamma(t) - \gamma(0)}{t} = h.$$

Par construction

$$\frac{f(\gamma(t_i)) - f(\gamma(0))}{t_i} = \frac{f(x + t_i k_i) - f(x)}{t_i}.$$

Comme la fonction est dérivable directionnellement au sens de Hadamard,

$$\lim_{i \to +\infty} \frac{f(x + t_i k_i) - f(x)}{t_i}$$

existe.

Question 5 Si f est différentiable, notons ε la fonction définie dans un voisinage de 0, continue et nulle en 0, telle que

$$f(x + h) = f(x) + df(x) \cdot h + \varepsilon(h)\|h\|.$$

On a alors pour tout $t \in \mathbb{R}$ non nul et tout vecteur $k \in \mathbb{R}^n$ suffisamment petits, en posant $h = tk$,

$$\begin{aligned} \frac{f(x + tk) - f(x)}{t} &= \frac{1}{t} df(x) \cdot tk + \frac{1}{t} \varepsilon(tk)\|tk\| \\ &= df(x) \cdot k + \varepsilon(tk) \frac{|t|}{t} \|k\|. \end{aligned}$$

Le terme $df(x) \cdot k$ tend vers $df(x) \cdot h$ quand $k \to h$ et le second terme du membre de droite tend vers 0 quand t et k tendent vers 0, donc

$$\lim_{(t,k) \to (0,h)} \frac{f(x + tk) - f(x)}{t} = df(x) \cdot h.$$

Par conséquent la fonction f est directionnellement dérivable au sens de Hadamard. Le membre de droite, égal à $f'(x,h)$, est linéaire en h ; la fonction f est donc différentiable au sens de Hadamard.

Réciproquement, supposons que f est différentiable au sens de Hadamard. Pour montrer que f est différentiable, de différentielle $f'(x,h)$, montrons que

$$\frac{\|f(x+h) - f(x) - f'(x,h)\|}{\|h\|} \to 0 \text{ quand } h \to 0,$$

ou de façon équivalente, que

$$\frac{f\left(x + \|h\| \frac{h}{\|h\|}\right) - f(x)}{\|h\|} - f'\left(x, \frac{h}{\|h\|}\right) \to 0 \text{ quand } h \to 0.$$

Il nous suffit de montrer que pour toute suite $t_i > 0$ telle que $t_i \to 0$ quand $i \to +\infty$ et $k_i \in \mathbb{R}^n$ telle que $\|k_i\| = 1$,

$$\frac{f(x + t_i k_i) - f(x)}{t_i} - f'(x, k_i) \to 0 \text{ quand } i \to +\infty.$$

Imaginons au contraire que cette expression ne tende pas vers 0. Alors on pourrait trouver un $\varepsilon > 0$ et une sous-suite de (t_i, k_i), notée de (t'_i, k'_i), telle que pour tout i,

$$\left\|\frac{f(x + t'_i k'_i) - f(x)}{t'_i} - f'(x, k'_i)\right\| \geq \varepsilon.$$

Mais la suite des k'_i est de norme égale à 1 ; la sphère fermée de centre 1 étant compacte, il existe des sous-suites t''_i et k''_i de t'_i et k'_i et un $h \in \mathbb{R}^n$ tels que $\|h\| = 1$ et $k''_i \to h$. Par hypothèse de dérivabilité au sens de Hadamard, on aurait

$$\frac{f(x + t''_i k''_i) - f(x)}{t''_i} \to f'(x, h) \text{ quand } i \to +\infty$$

ce qui contredit l'inégalité ci-dessus et prouve la contradiction. Par conséquent, f est bien différentiable.

Thermodynamique

Question 0 Les variables associées à l'expression fournie de l'entropie S sont le volume V et la température T. Pour que l'expression définissant l'entropie soit toujours définie, il suffit d'exiger que V et T soient strictement positives.

Question 1 L'entropie, en tant que fonction de $(V,T) \in \,]0,+\infty[^2$, est une fonction (continûment) différentiable. En effet, les dérivées partielles de $S(V,T)$ sont définies en tout point et vérifient

$$\frac{\partial S(V,T)}{\partial V} = Nk_B \frac{1}{V} \text{ et } \frac{\partial S(V,T)}{\partial T} = \frac{3}{2} Nk_B \frac{1}{T},$$

deux expressions dépendant continûment de (V,T). On a par conséquent

$$dS = \frac{\partial S(V,T)}{\partial V} dV + \frac{\partial S(V,T)}{\partial T} dT = Nk_B \left[\frac{dV}{V} + \frac{3}{2}\frac{dT}{T}\right].$$

Question 2 Si U est une fonction (différentiable) de (V, T), on peut interpréter mathématiquement la relation $dU = TdS - PdV$ comme

$$d(U(V, T)) = Td(S(V, T)) - P(V, T)dV$$

où $P(V, T) := Nk_BT/V$ résulte de la loi des gaz parfaits $PV = Nk_BT$. En exploitant la différentielle $dS(V, T)$ déjà calculée, on en déduit

$$d(U(V, T)) = TNk_B \left[\frac{dV}{V} + \frac{3}{2}\frac{dT}{T}\right] - Nk_BT\frac{dV}{V} = \frac{3}{2}Nk_BdT.$$

Question 3 Soit $(V_0, T_0) \in \left]0, +\infty\right[^2$ et $(V, T) \in \left]0, +\infty\right[^2$. Le segment reliant (V_0, T_0) et (V, T) est inclus tout entier dans $\left]0, +\infty\right[^2$ qui est convexe. Pour tout $t \in [0, 1]$, on a

$$\begin{aligned} \phi(t) &:= dU(V_0 + t(V - V_0), T_0 + t(T - T_0)) \cdot (V - V_0, T - T_0) \\ &= \frac{3}{2}Nk_BdT \cdot (V - V_0, T - T_0) \\ &= \frac{3}{2}Nk_B(T - T_0). \end{aligned}$$

La fonction ϕ est constante, donc intégrable sur $[0, 1]$. Par le théorème fondamental du calcul multivariable (p. 20), on a

$$U(V, T) = U(V_0, T_0) + \int_0^1 \phi(t)\, dt = \left[U(V_0, T_0) - \frac{3}{2}Nk_BT_0\right] + \frac{3}{2}Nk_BT,$$

ce qui démontre qu'à une constante près, on a

$$U(V, T) = \frac{3}{2}Nk_BT.$$

Références

Shapiro, A. 1990. "On Concepts of Directional Differentiability." *Journal of Optimization Theory and Applications* 66 (3): 477–87.