Université Paris Dauphine
S4 MIDO
Année 2017-2018

Calcul différentiel et optimisation :

Cours

Responsables :
Amic Frouvelle

frouvelle@ceremade.dauphine.fr

bureau C610
Nejla Nouaili

nouaili@ceremade.dauphine.fr

bureau B618 bis

Plan du cours

On étudie le calcul différentiel dans l'espace de dimension finie $\mathbb{R}^n$. Le cours se compose de trois grandes parties :
- — Topologie et continuité (chapitres 1 à 3),
- — Différentiabilité (chapitres 4 à 6),
- — Applications (chapitres 7 à 9).

Table des matières

Quelques autres références

Il y a de nombreux bons ouvrages sur le calcul différentiel dans $\mathbb{R}^n$. Entre autres :
- — `http://www.proba.jussieu.fr/pageperso/bolley/poly-cdiff.pdf` : le cours de F. Bolley, donné à Dauphine en 2010-2011, et dont ce cours-ci s'inspire fortement.
- — C. Leruste, *Calcul différentiel*, Elsevier-Masson, 1997 : un livre d'un contenu assez proche de celui de ce cours.
- — J. Dieudonné, *Elements d'analyse, Tome 1*, Jacques Gabay, 3e éd., 1979 : très bon ouvrage classique, mais d'un niveau plus avancé que celui de ce cours.

Ce cours a bénéficié des remarques des étudiants des années 2015-2016 et 2016-2017, qui m'ont permis de corriger plusieurs erreurs.

1. Normes sur $\mathbb{R}^n$

Dans ce chapitre, on définit ce qu'est une norme sur $\mathbb{R}^n$. L'existence d'une norme permet de définir la notion de suites convergentes et de suites bornées dans $\mathbb{R}^n$. Le résultat principal du chapitre est que toutes les normes sur $\mathbb{R}^n$ sont équivalentes (car $\mathbb{R}^n$ est de dimension finie). Ce résultat sera très utile : il permet de montrer que plusieurs concepts introduits dans ce cours ne dépendent pas du choix de la norme.

Espace vectoriel normé $\mathbb{R}^n$
Pour $n \in \mathbb{N}^*$, on définit

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_k \in \mathbb{R} \text{ pour tout } 1 \leq k \leq n\}.$$

Cet ensemble forme un espace vectoriel pour l'addition et la multiplication par un scalaire définies respectivement par

(1) addition : $x + y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$
(2) multiplication scalaire : si $\lambda \in \mathbb{R}$, $\lambda x = \lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$.

Une norme sur $\mathbb{R}^n$ est une application

$$\|\cdot\| : \mathbb{R}^n \to \mathbb{R} : x \mapsto \|x\|$$

telle que

(1) $\|x\| > 0$ pour tout $0 \neq x \in \mathbb{R}^n$,
(2) $\|\lambda x\| = |\lambda| \, \|x\|$ pour tout $x \in \mathbb{R}^n$ et tout $\lambda \in \mathbb{R}$,
(3) $\|x + y\| \leq \|x\| + \|y\|$ pour tout $x, y \in \mathbb{R}^n$ (inégalité triangulaire).

En particulier $\|0\| = 0$.

Remarque : on déduit de l'inégalité triangulaire que

$$\big| \|x\| - \|y\| \big| \leq \|x - y\|$$

pout tout $x, y \in \mathbb{R}^n$. En effet

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|,$$
$$\|y\| = \|y - x + x\| \leq \|x - y\| + \|x\|,$$

d'où

$$\|x\| - \|y\| \leq \|x - y\| \qquad \text{et} \qquad \|y\| - \|x\| \leq \|x - y\|,$$

d'où le résultat.

Exemples de normes

(1) $\|x\|_1 = \sum_{k=1}^n |x_k|$. L'inégalité triangulaire suit de $|a + b| \leq |a| + |b|$ où $|\cdot|$ est la valeur absolue sur $\mathbb{R}$, valable pour tout $a, b \in \mathbb{R}$.

(2) $\|x\|_\infty = \max_{1\le k\le n} |x_k|$. Pour montrer l'inégalité triangulaire, on note qu'il existe $j \in \{1,\dots,n\}$ tel que $\|x+y\|_\infty = |x_j + y_j|$, et donc

$$\|x+y\|_\infty = |x_j+y_j| \le |x_j| + |y_j| \le \|x\|_\infty + \|y\|_\infty$$

pour tout $x, y \in \mathbb{R}^n$.

(3) $\|x\|_2 = (\sum_{k=1}^n x_k^2)^{1/2}$. Cette norme représente la "longueur" usuelle (théorème de Pythagore). Pour démontrer l'inégalité triangulaire, on remarque que

$$\|x\|_2 = \langle x, x\rangle^{1/2},$$

où $\langle x, y\rangle = \sum_{k=1}^n x_k y_k$ est le produit scalaire usuel dans $\mathbb{R}^n$, puis on utilise l'inégalité de Cauchy-Schwarz (voir la preuve ci-dessous) :

$$|\langle x, y\rangle| \le \|x\|_2\,\|y\|_2.$$

Grâce à ça

$$\|x+y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 + 2\langle x, y\rangle \le \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2\|y\|_2 = \big(\|x\|_2 + \|y\|_2\big)^2.$$

Montrons l'inégalité de Cauchy-Schwarz. Pour tout $\lambda \in \mathbb{R}$,

$$\|x+\lambda y\|_2^2 = \|x\|_2^2 + \lambda^2\|y\|_2^2 + 2\lambda\langle x, y\rangle \ge 0.$$

Cette expression est un polynôme du second degré en λ qui admet donc au plus une racine. Dès lors

$$\Delta := 4\langle x, y\rangle^2 - 4\|x\|_2^2\|y\|_2^2 \le 0,$$

ce qui implique Cauchy-Schwarz.

Remarque : les notations $\|\cdot\|_p$ avec $p = 1, 2, \infty$ viennent de ce qu'en général

$$\|x\|_p = \Big(\sum_{k=1}^n |x_k|^p\Big)^{1/p}$$

est une norme pour tout $1 \le p < \infty$ (ce n'est pas toujours évident, voir le chapitre sur la convexité).

Convergence des suites dans $\mathbb{R}^n$

Une suite dans $\mathbb{R}^n$ est une application

$$\mathbb{N}^* \to \mathbb{R}^n : k \mapsto x_k.$$

On la note $(x_k)_{k\ge1}$. Si $E \subset \mathbb{R}^n$, on écrira (abusivement) $(x_k)_{k\ge1} \subset E$ pour dire que l'image de la suite $(x_k)_{k\ge1}$ est incluse dans E.

Exemples dans $\mathbb{R}^2$: $((1, 1/k))_{k\ge1}$, $((1/k, 1/k))_{k\ge1}$, $((k, k^2))_{k\ge1}$.

On munit $\mathbb{R}^n$ d'une norme $\|\cdot\|$. A priori, la notion de convergence d'une suite dépend du choix de la norme (mais nous verrons que ce n'est en fait pas le cas). Soit une suite $(x_k)_{k\ge1} \subset \mathbb{R}^n$, et soit $x \in \mathbb{R}^n$. On dit que $x_k \to x$ pour la norme $\|\cdot\|$ si

$$\forall \epsilon > 0, \exists M \in \mathbb{R} : k > M \Rightarrow \|x_k - x\| < \epsilon.$$

L'unicité de la limite se déduit de l'inégalité triangulaire. En effet supposons que x et x' soient tous deux limites de la suite $(x_k)_{k\geq 1}$. Pour tout $k \in \mathbb{N}^*$,

$$\|x - x'\| \leq \|x - x_k\| + \|x' - x_k\|.$$

Le membre de droite peut être rendu arbitrairement proche de 0, d'où $\|x - x'\| = 0$, d'où $x = x'$.

Cas particulier : $\|\cdot\| = \|\cdot\|_\infty$. Pour éviter la confusion avec les indices de la suite, notons les composantes d'un point $y \in \mathbb{R}^n$ par des indices supérieurs : $y = (y^1, \ldots, y^n)$. D'après la définition, $x_k \to x$ si

$$\forall \epsilon > 0, \exists M \in \mathbb{R} : k > M \Rightarrow \max_{1\leq j\leq n} |x_k^j - x^j| < \epsilon.$$

Donc, pour cette norme, la convergence de la suite $(x_k)_{k\geq 1}$ vers x équivaut à la convergence de la suite de la composante $(x_k^j)_{k\geq 1}$ vers la composante x^j pour chaque $1 \leq j \leq n$. En effet d'une part, si $x_k \to x$, alors on a l'implication

$$\max_{1\leq j\leq n} |x_k^j - x^j| \leq \epsilon \quad \Rightarrow \quad |x_k^j - x^j| \leq \epsilon \text{ pour tout } 1 \leq j \leq n.$$

D'autre part, si pour chaque $1 \leq j \leq n$ on a $x_k^j \to x^j$ alors, on sait que, pour tout $\epsilon > 0$ et pour tout $1 \leq j \leq n$,

$$\exists M_j \in \mathbb{R} : k \geq M_j \Rightarrow |x_k^j - x^j| < \epsilon.$$

Prenant $M = \max_{1\leq j\leq n} M_j$, on a $x_k \to x$.

Suites bornées dans $\mathbb{R}^n$

On munit $\mathbb{R}^n$ d'une norme $\|\cdot\|$ et de nouveau, la notion de suite bornée dépend a priori du choix de la norme (et de nouveau ce ne sera en fait pas le cas). On dit qu'une suite $(x_k)_{k\geq 1} \subset \mathbb{R}^n$ est bornée pour la norme $\|\cdot\|$ s'il existe $M \in \mathbb{R}$ tel que

$$\|x_k\| \leq M \text{ pour tout } k \geq 1.$$

Dans le cas particulier où $\|\cdot\| = \|\cdot\|_\infty$, on montre à nouveau que $(x_k)_{k\geq 1}$ est bornée si et seulement si chacune des suites $(x_k^j)_{k\geq 1}$ $(1 \leq j \leq n)$ est bornée.

Normes équivalentes

Deux normes $\|\cdot\|_a$ et $\|\cdot\|_b$ sont équivalentes s'il existe deux constantes $M_1, M_2 > 0$ telles que

$$\|x\|_a \leq M_1 \|x\|_b \quad \text{et} \quad \|x\|_b \leq M_2 \|x\|_a \quad \text{pour tout } x \in \mathbb{R}^n.$$

Exemple : les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sont équivalentes. En effet :

$$\|x\|_1 = \sum_{k=1}^{n} |x_k| \leq \sum_{k=1}^{n} \max_{1\leq j\leq n} |x_j| = n\|x\|_\infty,$$

$$\|x\|_\infty = \max_{1\leq j\leq n} |x_j| \leq \sum_{k=1}^{n} |x_k| = \|x\|_1.$$

Donc $M_1 = n$ et $M_2 = 1$. On note que ces constantes ne peuvent pas être améliorées (c'est-à-dire prises plus petites). En effet, si $x = (1, \ldots, 1)$, on a $\|x\|_1 = n$ et $\|x\|_\infty = 1$, et $\|e_1\|_\infty = \|e_1\|_1 = 1$ où $e_1 = (1, 0, \ldots, 0)$.

La notion de normes équivalentes est la plus importante du chapitre. Si $\|\cdot\|_a$ et $\|\cdot\|_b$ sont deux normes équivalentes, on voit qu'une suite est bornée ou convergente pour la norme $\|\cdot\|_a$ si et seulement si elle l'est pour la norme $\|\cdot\|_b$:

Preuve pour "bornée" : Si $(x_k)_{k\geq 1}$ est bornée pour la norme $\|\cdot\|_a$, il existe $R \geq 0$ tel que $\|x_k\|_a \leq R$ pour tout $k \in \mathbb{N}^*$. Dès lors, $\|x_k\|_b \leq M_2\|x_k\|_a \leq M_2 R$ pour tout $k \in \mathbb{N}^*$, et la suite $(x_k)_{k\geq 1}$ est bornée pour la norme $\|\cdot\|_b$.

Preuve pour "convergente" : Soit $(x_k)_{k\geq 1}$ une suite convergente pour la norme $\|\cdot\|_a$, de limite $\overline{x}$. Ceci veut dire que, pour tout $\epsilon > 0$, il existe $R > 0$ tel que si $k \geq R$, on ait $\|x_k - \overline{x}\|_a < \epsilon$. Fixons $\epsilon > 0$. Il existe donc $R' > 0$ tel que, si $k \geq R'$, on ait $\|x_k - \overline{x}\|_a < \epsilon/M_2$. Mais alors on a aussi $\|x_k - \overline{x}\|_b \leq M_2\|x_k - \overline{x}\|_a < \epsilon$.

Nous allons voir à présent que toutes les normes sur $\mathbb{R}^n$ sont équivalentes. Ceci permet d'établir que beaucoup de concepts qui apparaissent dans ce cours sont indépendants du choix de la norme. Par exemple, on peut parler de suite bornée ou convegente sans référence à une norme.

Remarque : cette propriété dépend de façon cruciale du fait que $\mathbb{R}^n$ est de dimension finie. En effet, on a vu que, pour certains $x \in \mathbb{R}^n$, $\|x\|_1 = n\|x\|_\infty$, et on s'attend donc à ce que certaines normes ne soient plus équivalentes en dimension infinie.

Nous aurons d'abord besoin du résultat suivant :

Théorème de Bolzano-Weierstass

Etant donné une suite $(x_k)_{k\geq 1} \subset \mathbb{R}^n$, une sous-suite est une suite de la forme $(x_{f(j)})_{j\geq 1}$ où $f : \mathbb{N}^* \to \mathbb{N}^*$ est strictement croissante. Exemple : $((1/j^2, 1/j^2))_{j\geq 1}$ est une sous-suite de $((1/k, 1/k))_{k\geq 1}$ (la fonction f est $f(j) = j^2$).

Théorème : Toute suite bornée dans $\mathbb{R}^n$ pour la norme $\|\cdot\|_\infty$ admet une sous-suite convergente pour la norme $\|\cdot\|_\infty$.

Remarque : quand nous aurons énoncé le résultat sur l'équivalence des normes dans $\mathbb{R}^n$, nous pourrons bien sûr reformuler ce théorème sans référence à la norme $\|\cdot\|_\infty$: *Toute suite bornée dans $\mathbb{R}^n$ admet une sous-suite convergente.*

Preuve : pour $n = 1$, nous renvoyons au cours de L1(p. 35 du cours de F. Simenhaus). Supposons $n \geq 2$. On distingue les composantes d'un point $y \in \mathbb{R}^n$ par un indice supérieur : $y = (y^1, \ldots, y^n)$. On va extraire une sous-suite convergente de la suite $(x_k)_{k\geq 1}$ en n étapes. Puisque la suite $(x_k)_{k\geq 1}$ est bornée pour la norme $\|\cdot\|_\infty$, la suite réelle $(x_k^1)_{k\geq 1}$ est bornée. Donc il existe une sous-suite convergente $(x^1_{f_1(j)})_{j\geq 1}$ de la suite $(x^1_k)_{k\geq 1}$. Donc $(x_{f_1(j)})_{j\geq 1}$ est une sous-suite de $(x_k)_{k\geq 1}$ dont la première composante converge. On peut alors répéter le procédé et extraire une sous-suite $(x_{f_2\circ f_1(j)})_{j\geq 1}$ de $(x_{f_1(j)})_{j\geq 1}$ de sorte à ce que la deuxième composante converge aussi (bien sûr $(x_{f_2\circ f_1(j)})_{j\geq 1}$ est aussi une sous-suite de $(x_k)_{k\geq 1}$), et ainsi de suite n fois pour que toutes les composantes convergent. On obtient ainsi une sous-suite convergente pour la norme $\|\cdot\|_\infty$.

Toutes les normes sur $\mathbb{R}^n$ sont équivalentes

Nous montrons maintenant le résultat principal du chapitre :

Théorème : Toutes les normes sur $\mathbb{R}^n$ sont équivalentes

Preuve. Soit $\|\cdot\|$ une norme sur $\mathbb{R}^n$. Il suffit de vérifier que $\|\cdot\|$ et $\|\cdot\|_\infty$ sont équivalentes.

Montrons d'abord $\|\cdot\| \leq M_1\|\cdot\|_\infty$ pour un certain $M_1 > 0$:

$$\|x\| = \Big\| \sum_{k=1}^n x_k e_k \Big\| \leq \sum_{k=1}^n |x_k|\,\|e_k\| \leq C \sum_{k=1}^n |x_k| \leq Cn\|x\|_\infty,$$

où $C = \max_{1\leq k\leq n} \|e_k\|$. Le résultat s'obtient donc avec $M_1 = Cn$.

Montrons ensuite $\|\cdot\|_\infty \leq M_2\|\cdot\|$ pour un certain $M_2 > 0$. Cela revient à montrer que

$$m := \inf_{x\in\mathbb{R}^n/\{0\}} \frac{\|x\|}{\|x\|_\infty} = \inf_{x\in\mathbb{R}^n/\{0\}} \Big\| \frac{x}{\|x\|_\infty} \Big\| = \inf_{y\in\mathbb{R}^n:\|y\|_\infty=1} \|y\| > 0.$$

Par définition de inf, il existe une suite $(y_k)_{k\geq 1} \subset \mathbb{R}^n$ telle que

$$\|y_k\|_\infty = 1, \qquad \|y_k\| \to m \quad \text{si} \quad k \to \infty$$

(cette suite s'appelle une suite minimisante ; on peut la construire explicitement en notant par exemple que, pour tout $k \in \mathbb{N}^*$, il existe $y_k \in \mathbb{R}^n$ tel que $\|y_k\|_\infty = 1$ et $\|y_k\| \leq m+1/k$). La suite $(y_k)_{k\geq 1}$ est donc bornée pour la norme $\|\cdot\|_\infty$ et, par le théorème de Bolzano-Weierstrass, on peut en extraire une sous-suite $(y_{f(j)})_{j\geq 1}$ convergente pour la norme $\|\cdot\|_\infty$. Notons y la limite de cette sous-suite. Il suffit maintenant de voir que

$$\|y\|_\infty = 1 \quad \text{et} \quad \|y\| = m,$$

car alors, comme $\|y\|_\infty \neq 0$, on a $y \neq 0$ et donc $0 < \|y\| = m$.

On a d'une part

$$|\|y\|_\infty - 1| = |\|y\|_\infty - \|y_{f(j)}\|_\infty| \leq \|y - y_{f(j)}\|_\infty$$

et cette quantité tend vers 0 quand j tend vers l'infini. D'autre part,

$$|\|y\| - m| = |\|y\| - \|y_{f(j)}\| + \|y_{f(j)}\| - m| \leq |\|y\| - \|y_{f(j)}\|| + |\|y_{f(j)}\| - m|.$$

Le premier terme du membre de droite tend vers 0 quand $j \to \infty$ car

$$\left| \|y\| - \|y_{f(j)}\| \right| \leq \|y - y_{f(j)}\| \leq M_1 \|y - y_{f(j)}\|_\infty,$$

et cette quantité tend vers 0. Le second tend aussi vers 0 quand $j \to \infty$ par définition de la suite $(y_k)_{k \geq 1}$.

Remarque : Grâce à ce résultat, nous pouvons à présent dire qu'une suite dans $\mathbb{R}^n$ converge si et seulement si la suite de chacune de ses composantes converge, et qu'une suite dans $\mathbb{R}^n$ est bornée si et seulement si la suite de chacune de ses composantes est bornée, quelle que soit la norme utilisée.

2. Topologie de $\mathbb{R}^n$

Dans ce chapitre, nous définissons les principaux concepts relatifs à la topologie de $\mathbb{R}^n$: ouverts, fermés, bornés, compacts, adhérence et intérieur ; nous caractérisons certaines de ces notions en terme de convergence de suites. Ces concepts apparaitront à plusieurs reprises dans la suite, pour au moins deux raisons : d'une part les ouverts permettent de définir des propriétés locales (et non strictement ponctuelles) de fonctions, et d'autre part la nature topologique du domaine des fonctions jouera un rôle important en optimisation.

Boules ouvertes et voisinages

On munit $\mathbb{R}^n$ d'une norme $\|\cdot\|$. Soit $r > 0$, et soit $a \in \mathbb{R}^n$. L'ensemble

$$B(a, r) = \{x \in \mathbb{R}^n : \|x - a\| < r\}$$

s'appelle la boule ouverte de rayon r autour du point a.

Il peut être commode de formuler un concept similaire qui soit indépendant du choix de la norme sur $\mathbb{R}^n$. On dit que V est un voisinage du point a s'il existe une norme sur $\mathbb{R}^n$ telle que V contienne une boule ouverte autour de a.

Grâce au résultat établi au cours précédent selon lequel toutes les normes sur $\mathbb{R}^n$ sont équivalentes, on voit que si V est un voisinage du point a, alors V contient une boule ouverte autour du point a pour toute norme.

Preuve : Soit V un voisinage de a. Par définition il existe une norme $\|\cdot\|$ et un rayon $r > 0$ tels que tous les points x tels que $\|x - a\| < r$ sont dans V. Prenons maintenant une autre norme $\|\cdot\|'$. Il suffit de montrer qu'il existe $r' > 0$ tel que, si $\|x - a\|' < r'$, alors $\|x - a\| < r$. Or les normes $\|\cdot\|$ et $\|\cdot\|'$ sont équivalentes, et donc $\|\cdot\| \leq M\|\cdot\|'$ pour un certain $M > 0$. Dès lors, si $\|x - a\|' < r'$ pour $r' = r/M$, on a $\|x - a\| \leq M\|x - a\|' < r$.

Ouverts et fermés

On munit $\mathbb{R}^n$ d'une norme.

Un sous-ensemble $E \subset \mathbb{R}^n$ est ouvert si, pour tout $x \in E$, il existe une boule ouverte de centre x comprise dans E. Un sous-ensemble est fermé si son complémentaire est ouvert.

On vérifie que la notion d'ouvert est indépendante du choix de la norme : si E est ouvert pour une norme, il l'est pour toute norme. On peut d'ailleurs reformuler la définition en terme de voisinage : E est ouvert si E est un voisinage de x pour tout $x \in E$.

Remarque : un ensemble $E \subset \mathbb{R}^n$ peut n'être ni ouvert ni fermé ; c'est le cas de $[0, 1[\subset \mathbb{R}$ $(n = 1)$.

Exemples (les ensembles ci-dessous sont vus comme des sous-ensembles de $\mathbb{R}^n$) :

(1) L'ensemble vide et $\mathbb{R}^n$ sont à la fois ouverts et fermés.

(2) Un point est un ensemble fermé.

(3) Une boule ouverte est ouverte. En effet, soit une boule ouverte $B(a,r)$ (pour une certaine norme $\|\cdot\|$), et soit $x \in B(a,r)$. Donc $\|x-a\| = r' < r$. Soit alors $r'' = (r-r')$. Il suffit de voir que $B(x,r'') \subset B(a,r)$. Or, si $y \in B(x,r'')$, on a

$$\|y-a\| = \|y-x+x-a\| \leq \|y-x\| + \|x-a\| < r'' + r' = r.$$

(4) Une union d'ouverts est ouverte, et une intersection finie d'ouverts est ouverte. Par contre une intersection infinie d'ouverts n'est pas toujours un ouvert ; par exemple

$$\bigcap_{k\geq 1}]-1/k, 1/k[= \{0\} \subset \mathbb{R} \quad (n=1).$$

(5) Un sous-espace vectoriel de $\mathbb{R}^n$ est fermé. En effet, considérons un sous-espace V de dimension m où $1 \leq m < n$ (les cas $m=0$ et $m=n$ sont immédiats). Donc il existe m vecteurs $v_1, \ldots, v_m$ linéairement indépendants tels que

$$V = \{x \in \mathbb{R}^n : x = a_1v_1 + \cdots + a_mv_m \text{ pour certains } a_1, \ldots, a_m \in \mathbb{R}\}.$$

Quitte à utiliser le procédé de Gram-Schmidt, on peut supposer que $v_1, \ldots, v_m$ forment une base orthonormée. De plus on peut trouver des vecteurs $v_{m+1}, \ldots, v_n$ tels que $\{v_1, \ldots, v_n\}$ soit une base orthonormée de $\mathbb{R}^n$.

Montrons que V^c est ouvert. Il est commode de munir $\mathbb{R}^n$ de la norme $\|\cdot\|_2$. Prenons $x \in V^c$ et décomposons-le dans la base $\{v_1, \ldots, v_n\}$:

$$x = x_1v_1 + \cdots + x_nv_n.$$

Comme $x \in V^c$, on a forcément $x_\ell \neq 0$ pour un certain $\ell \in \{m+1, \ldots, n\}$. Donc, si $y \in V$, on a

$$\|y-x\|_2^2 = (x_1-y_1)^2 + \cdots + (x_m-y_m)^2 + x_{m+1}^2 + \cdots + x_n^2 \geq x_\ell^2 > 0.$$

Dès lors une boule de rayon inférieur à $|x_\ell|$ centrée en x ne peut contenir de points de V. Donc V^c est ouvert.

Remarque : plus tard, nous verrons une façon plus simple de montrer ce résultat.

Ensembles bornés

Un sous-ensemble $E \subset \mathbb{R}^n$ est borné s'il existe une norme $\|\cdot\|$ sur $\mathbb{R}^n$ et un nombre $M > 0$ tels que $\|x\| \leq M$ pour tout $x \in E$. Bien sûr, si E est borné, alors, pour toute norme $\|\cdot\|'$, il existe $M' > 0$ tel que $\|x\|' < M'$ pour tout $x \in E$.

Exemples : l'ensemble vide, un point, une boule ouverte, une union finie de bornés, et une intersection de bornés sont des ensembles bornés.

Contre-exemples : $\mathbb{R}^n$, un sous-espace vectoriel de $\mathbb{R}^n$ de dimension non nulle, le graphe d'une application de $\mathbb{R}$ dans $\mathbb{R}$ (sous-ensemble de $\mathbb{R}^2$) ne sont pas bornés.

Adhérence et intérieur

Soit un sous-ensemble $E \subset \mathbb{R}^n$.

L'intérieur de E, noté $\mathrm{int}(E)$ est l'union de tous les ouverts contenus dans E.
L'adhérence de E, notée $\mathrm{adh}(E)$, est l'intersection de tous les fermés contenant E.

Remarque : E est ouvert si et seulement si $E = \mathrm{int}(E)$; E est fermé si et seulement si $E = \mathrm{adh}(E)$.

Exemple : On munit $\mathbb{R}^n$ de la norme $\|\cdot\|_2$. Soit

$$E = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1 \text{ et } (\|x\|_2 < 1 \text{ si } x_1 \leq 0)\}.$$

(1) $\mathrm{int}(E) = A := \{x \in \mathbb{R}^n : \|x\|_2 < 1\}$.
En effet, d'une part A est un ouvert contenu dans E. D'autre part, imaginons par l'absurde un ouvert $U \subset E$ tel que $U \not\subset A$. Donc il existe $x \in U$ tel que $\|x\|_2 \geq 1$. Comme U est ouvert, il existe alors une boule $B(x,r) \subset U$ pour un certain $r > 0$, et donc $x + \frac{r}{2}\frac{x}{\|x\|_2} \in U$. Mais $\|x + \frac{r}{2}\frac{x}{\|x\|_2}\|_2 > 1$ et donc on aboutit à la contradiction $U \not\subset E$.

(2) $\mathrm{adh}(E) = B := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$.
En effet, d'une part B est un fermé qui contient E. D'autre part, imaginons par l'absurde un fermé V tel que $E \subset V$ mais $B \not\subset V$. Donc il existe $x \in B$ tel que $x \notin V$. Comme le complémentaire de V est ouvert, il existe donc une boule $B(x,r)$, pour un certain $r > 0$, telle qu'aucun point de cette boule ne soit dans V. En particulier, $x - \frac{r'}{2}\frac{x}{\|x\|_2} \notin V$ pour tout $0 < r' < r$. Mais $\|x - \frac{r'}{2}\frac{x}{\|x\|_2}\|_2 < 1$ pour $r' > 0$ assez petit, et on aboutit à la contradiction $E \not\subset V$.

Caractérisation par la convergence de suites
Un sous-ensemble $E \subset \mathbb{R}^n$ est compact si toute suite d'éléments de E admet une sous-suite qui converge dans E.

Proposition : $E \subset \mathbb{R}^n$ est compact si et seulement si E est fermé borné.

Preuve :
Supposons d'abord E fermé borné. Comme E est borné, toute suite $(x_k)_{k\geq 1}$ admet une sous-suite convergente $(x_{f(j)})_{j\geq 1}$, d'après le théorème de Bolzano-Weierstarss. Soit x la limite de $(x_{f(j)})_{j\geq 1}$. Il reste à voir que $x \in E$. Imaginons par l'absurde $x \notin E$. Comme E^c est ouvert, il existe $r > 0$ tel que $B(x,r) \subset E^c$. Mais alors pour tout k, on aurait $\|x - x_k\| \geq r$, ce qui contredit $x = \lim_{j\to\infty} x_{f(j)}$.
Supposons ensuite que E est compact, c'est-à-dire que toute suite de E admet une sous-suite convergente dans E. Imaginons d'abord par l'absurde que E ne soit pas fermé. Alors E^c ne serait pas ouvert, et il existerait $x \in E^c$ tel que, pour tout $r > 0$, on aurait $B(x,r) \cap E \neq \varnothing$. Mais alors, on aurait une suite d'éléments de E qui converge vers x, ce qui est une contradiction puisque $x \in E^c$. Imaginons ensuite que E ne soit pas borné. Alors il existerait une suite $(x_k)_{k\geq 1} \subset E$ telle que $\|x_k\| \geq k$. Donc toute sous-suite de (x_k) serait non-bornée, et donc divergente, ce qui est de nouveau une contradiction.

Remarque : Il existe une autre façon de définir un ensemble compact. Un ensemble $E \subset \mathbb{R}^n$ satisfait à la propriété de Borel-Lebesgue si, de tout recouvrement ouvert de E, c'est-à-dire de toute famille $\{O_i\}_{i\in I}$ d'ouverts de $\mathbb{R}^n$ telle que $E \subset \cup_{i\in I} O_i$, on peut extraire un sous-recouvrement fini, c'est-à-dire, si le recouvrement est $\{O_i\}_{i\in I}$, des indices $i_1, \ldots, i_m \in I$ tels que $E \subset O_{i_1} \cup \cdots \cup O_{i_m}$. On peut montrer, mais nous ne le ferons pas dans ce cours, que $E \subset \mathbb{R}^n$ est compact si et seulement s'il satisfait à la propriété de Borel-Lebesgue.

Proposition : Soit $E \subset \mathbb{R}^n$ *et soit* $x \in \mathbb{R}^n$. *On a* $x \in \mathrm{adh}(E)$ *si et seulement s'il existe une suite* $(x_k)_{k\geq 1} \subset E$ *telle que* $\lim_{k\to\infty} x_k = x$.

Preuve :
Supposons d'abord $(x_k)_{k\geq 1} \subset E$ telle que $\lim_{k\to\infty} x_k = x$. Montrons que x est dans l'intersection de tous les fermés contenant E. Bien sûr, $(x_k)_{k\geq 1} \subset F$ pour tout fermé $F \supset E$. De plus, comme $(x_k)_{k\geq 1}$ converge, $(x_k)_{k\geq 1}$ est bornée et il existe donc un fermé borné B tel que $(x_k)_{k\geq 1} \subset B$. Donc, $(x_k)_{k\geq 1} \subset B \cap F$ pour tout fermé $F \supset E$, et donc, par le résultat précédent, $x \in B \cap F$, donc $x \in F$.
Supposons ensuite $x \in \mathrm{adh}(E)$. Imaginons par l'absurde qu'aucune suite d'éléments de E ne converge vers x. Alors il existerait $r > 0$ tel que $B(x, r) \cap E = \varnothing$. Mais alors $\mathrm{adh}(E)/B(x, r)$ serait un fermé qui contient E mais ne contient pas x, ce qui contredit $x \in \mathrm{adh}(E)$, c'est-à-dire par définition l'intersection de tous les fermés contenant E.

Exemple
A l'aide de la caractérisation qu'on vient de dériver, montrons

$$\mathrm{adh}(A \cup B) = \mathrm{adh}(A) \cup \mathrm{adh}(B)$$

pour tout $A, B \subset \mathbb{R}^n$.
D'une part prenons $x \in \mathrm{adh}(A \cup B)$. Donc il existe une suite $(x_k)_{k\geq 1} \subset A \cup B$ qui converge vers x. Donc il existe au moins une sous-suite $(x_{f(j)})_{j\geq 1}$ contenue dans A ou contenue dans B, et bien sûr de limite x. Donc $x \in \mathrm{adh}(\mathrm{A})$ ou $x \in \mathrm{adh}(B)$.
D'autre part, prenons $x \in \mathrm{adh}(A) \cup \mathrm{adh}(B)$. Supposons $x \in \mathrm{adh}(A)$ (l'autre cas est analogue). Donc il existe $(x_k)_{k\geq 1} \subset A \subset A \cup B$ qui converge vers x, et donc $x \in \mathrm{adh}(A \cup B)$.

Attention : En général $\mathrm{adh}(A \cap B) \neq \mathrm{adh}(A) \cap \mathrm{adh}(B)$. En effet, dans $\mathbb{R}$, si $A =]-1, 0[$ et $B =]0, 1[$, alors $A \cap B = \varnothing$ d'où $\mathrm{adh}(A \cap B) = \varnothing$, alors que $\mathrm{adh}(A) \cap \mathrm{adh}(B) = [-1, 0] \cap [0, 1] = \{0\}$.

Attention : Si $(A_i)_{i\in I}$ est une famille infinie de sous-ensembles de $\mathbb{R}^n$, en général

$$\mathrm{adh}\Big(\bigcup_{i\in I} A_i\Big) \neq \bigcup_{i\in I} \mathrm{adh}(A_i).$$

En effet, dans $\mathbb{R}$, prenons $I = \mathbb{N}^*$ et $A_i = \{1/i\}$.

D'une part,

$$\mathrm{adh}\Big(\bigcup_{i\geq 1}\{1/i\}\Big) = \Big(\bigcup_{i\geq 1}\{1/i\}\Big) \cup \{0\}.$$

En effet, dénotons par E l'ensemble qui apparaît dans le membre de droite. On voit tout de suite que tout point de E est limite d'une suite de points de $\{1/i : i \in \mathbb{N}^*\}$. Montrons que, réciproquement, toute suite convergente $(x_k)_{k\geq 1} \subset \{1/i : i \in \mathbb{N}^*\}$ converge vers un point $x \in E$. Distinguons deux cas. Soit $(x_k)_{k\geq 1}$ ne prend qu'un nombre fini de valeurs. Dans ce cas, $(x_k)_{k\geq 1}$ est constante à partir d'un certain rang et donc $x \in \{1/i : i \in \mathbb{N}^*\} \subset E$. Supposons ensuite que $(x_k)_{k\geq 1}$ prenne un nombre infini de valeurs. Dans ce cas, on peut extraire de $(x_k)_{k\geq 1}$ une sous-suite strictement décroissante qui tend vers $0 \in E$.

D'autre part

$$\bigcup_{i\geq 1}\mathrm{adh}(\{1/i\}) = \bigcup_{i\geq 1}\{1/i\}.$$

3. Fonctions continues

La continuité d'une fonction permet le passage à la limite : une fonction f est continue en a si $x \to a$ implique $f(x) \to f(a)$. Par exemple, au chapitre 1, on a implicitement utilisé la continuité de la norme pour justifier un passage à la limite dans la preuve du résultat qui affirme que toutes les normes sur $\mathbb{R}^n$ sont équivalentes. Ce type de raisonnement apparaitra encore à plusieurs reprises dans la suite.

Dans tout ce chapitre, et les suivants, on munit $\mathbb{R}^n$ d'une norme ; si on considère une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$, les normes sur $\mathbb{R}^n$ et $\mathbb{R}^m$ sont en général différentes mais toutes deux notées $\|\cdot\|$.

Rappel sur les applications linéaires

Les applications linéaires apparaîtrons à plusieurs reprises dans ce cours. Une application $f : \mathbb{R}^n \to \mathbb{R}^m$ est linéaire si $f(x) = Ax$ pour une certaine matrice $A \in \mathbb{R}^{m\times n}$, si x est représenté par le vecteur colonne $(x_1, \ldots, x_n)^t$ et $f(x)$ par le vecteur colonne $(f_1(x), \ldots, f_m(x))^t$.

Toute application linéaire est bornée au sens où il existe $M > 0$ tel que

$$\|f(x)\| \leq M\|x\| \quad \text{pour tout } x \in \mathbb{R}^n.$$

En effet

$$\|f(x)\| = \|Ax\| = \Big\|A\sum_{k=1}^{n} x_k e_k\Big\| = \Big\|\sum_{k=1}^{n} x_k A e_k\Big\|$$

$$| \leq \sum_{k=1}^{n} |x_k|\|Ae_k\| \leq M_1\|x\|_1 \leq M_1 M_2\|x\| = M\|x\|,$$

où $M_1 = \max_{1\leq k\leq n} \|Ae_k\|$, M_2 est telle que $\|\cdot\|_1 \leq M_2\|\cdot\|$, et $M = M_1 M_2$.

Attention : l'usage du mot 'borné' est trompeur, car c'est en fait la fonction

$$\mathbb{R}^n\backslash\{0\} \to \mathbb{R}^m, x \mapsto \frac{f(x)}{\|x\|}$$

qui est bornée (au sens usuel) !

Soit A la matrice qui représente l'application linéaire f. Définissons

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

La quantité $\|A\|$ correspond donc à la plus petite constante M telle que $\|f(x)\| \leq M\|x\|$ pour tout $x \in \mathbb{R}^n$. En fait l'application $A \mapsto \|A\|$ définit une norme sur les matrices $m \times n$ (c'est-à-dire sur $\mathbb{R}^{m\times n}$), qu'on appelle la 'norme opérateur'. On note que le symbole $\|\cdot\|$ est utilisé dans trois sens différents : une norme sur $\mathbb{R}^n$ ($\|x\|$), une norme sur $\mathbb{R}^m$ ($\|Ax\|$) et une norme sur $\mathbb{R}^{n\times m}$ ($\|A\|$).

Fonction continue
Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, et soit $a \in E$. On dit que f est continue en a si, pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que, pour tout $x \in E$,

$$\|x - a\| < \delta \quad \Rightarrow \quad \|f(x) - f(a)\| < \epsilon.$$

Comme toutes les normes sont équivalentes sur $\mathbb{R}^n$ et $\mathbb{R}^m$, on montre que la continuité est un concept indépendant du choix de la norme.

On dit que la fonction f est continue sur un sous-ensemble $U \subset E$ si f est continue en tout point $a \in U$.

Exemples :

(1) Une norme $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ est continue en tout point $a \in \mathbb{R}^n$. Soit $\epsilon > 0$. Prenons $\delta = \epsilon$. Alors, si $\|x - a\| < \delta$, on a $|\|x\| - \|a\|| \leq \|x - a\| < \epsilon$.

(2) Une application linéaire $f : \mathbb{R}^n \to \mathbb{R}^m$ est continue en tout point $a \in \mathbb{R}^n$. Soit en effet M tel que $\|f(x)\| \leq M\|x\|$ pour tout $x \in \mathbb{R}^n$. Prenons $\epsilon > 0$, et supposons $\|x - a\| < \delta$ pour $\delta = \epsilon/M$. On a

$$\|f(x) - f(a)\| = \|f(x - a)\| \leq M\|x - a\| < \epsilon.$$

(3) On rappelle qu'une application $f : \mathbb{R}^n \to \mathbb{R}$ est une forme quadratique sur $\mathbb{R}^n$ s'il existe une matrice symétrique $A \in \mathbb{R}^{n \times n}$ telle que $f(x) = \langle x, Ax \rangle$, où x est représenté par le vecteur colonne $(x_1, \ldots, x_n)^t$. Une forme quadratique $f : \mathbb{R}^n \to \mathbb{R}$ est continue en tout point $a \in \mathbb{R}^n$. Il est commode de munir $\mathbb{R}^n$ de la norme $\|\cdot\|_2$. On a

$$\begin{aligned}
|f(x) - f(a)| = |\langle x, Ax \rangle - \langle a, Aa \rangle| &= |\langle x - a, Ax \rangle + \langle a, A(x - a) \rangle| \\
&\leq |\langle Ax, x - a \rangle| + |\langle Aa, (x - a) \rangle| \\
&\leq (\|Ax\|_2 + \|Aa\|_2)\|x - a\|_2
\end{aligned}$$

où on a utilisé l'inégalité de Cauchy-Schwarz pour obtenir la dernière inégalité. De plus, comme on a vu dans l'exemple précédent, il existe une constante $M > 0$ telle que

$$\|Ax\|_2 \leq M\|x\|_2 \leq M(\|x - a\|_2 + \|a\|_2).$$

Dès lors, si $\|x - a\|_2 < \delta$ pour un certain $\delta > 0$, on a

$$|f(x) - f(a)| < (M\delta + M\|a\|_2 + \|Aa\|_2)\delta,$$

ce qui peut être rendu plus petit que n'importe quel $\epsilon > 0$ en prenant $\delta > 0$ assez petit.

(4) La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \frac{x^3 + y^3}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \quad \text{et} \quad f(0, 0) = 0$$

est continue en $(0,0)$. Il est commode de munir $\mathbb{R}^2$ de la norme $\|\cdot\|_1$. Fixons $\epsilon > 0$ et prenons $\delta = \epsilon$. Alors, pour les points $(x,y) \in \mathbb{R}^2$ différents de $(0,0)$ tels que $\|(x,y)\|_1 < \delta$, on a

$$\begin{aligned}|f(x,y)| = \Big|\frac{x^3+y^3}{x^2+y^2}\Big| &= \Big|x\frac{x^2}{x^2+y^2} + y\frac{y^2}{x^2+y^2}\Big| \\ &\le |x|\frac{x^2}{x^2+y^2} + |y|\frac{y^2}{x^2+y^2} \le |x| + |y| = \|(x,y)\|_1 < \delta = \epsilon.\end{aligned}$$

Opérations sur les fonctions continues

Nous montrons à présent certains résultats qui permettent de ne pas revenir toujours à la définition pour s'assurer qu'une fonction est continue. En particulier, nous pourrons réutiliser nos connaissances sur les fonctions continues sur $\mathbb{R}$.

Proposition : Supposons $n \geq 2$. Soit $f : I \subset \mathbb{R} \to \mathbb{R}$ une fonction continue en $a \in I$. Alors la fonction $F : I \times \mathbb{R}^{n-1} \to \mathbb{R}$ définie par $F(x) = f(x_1)$ est continue en tout point de la forme $(a, x_2, \ldots, x_n)$.

Preuve : Il est commode de munir $\mathbb{R}^n$ de la norme $\|\cdot\|_\infty$. Soit $\overline{a} \in \mathbb{R}^n$ un point de la forme $(a, x_2, \ldots, x_n)$. Soit $\epsilon > 0$. On sait qu'il existe $\delta > 0$ tel que, si $|y - a| < \delta$ et $y \in I$, alors $|f(y) - f(a)| < \epsilon$. Alors, pour tout $x \in I \times \mathbb{R}^{n-1}$ tel que $\|x - \overline{a}\|_\infty < \delta$, on a $|x_1 - a| < \delta$ et $x_1 \in I$ et donc

$$|F(x) - F(\overline{a})| = |f(x_1) - f(a)| < \epsilon.$$

Proposition : Une fonction $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ est continue en $a \in E$ si et seulement si toutes les applications $f_j : E \subset \mathbb{R}^n \to \mathbb{R}$ sont continues en a pour tout $1 \leq j \leq m$, où on a noté $f(x) = (f_1(x), \ldots, f_m(x))$.

Preuve : Exercice ! (il est commode d'utiliser la norme $\|\cdot\|_\infty$)

Proposition : la continuité est stable pour un certain nombre d'opérations :

1. *Si $f, g : E \subset \mathbb{R}^n \to \mathbb{R}$ sont continues en $a \in E$, alors $f + g$ est continue en a.*
2. *Si $f, g : E \subset \mathbb{R}^n \to \mathbb{R}$ sont continues en $a \in E$, alors fg est continue en a.*
3. *Si $f : E \subset \mathbb{R}^n \to \mathbb{R}$ est continue en $a \in E$ et $f(a) \neq 0$, alors $1/f$ est bien définie dans un voisinage de a et est continue en a.*
4. *Si $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ est continue en $a \in E$, et si $g : \mathrm{Im}(f) \subset \mathbb{R}^m \to \mathbb{R}^p$ est continue en $f(a)$, alors $g \circ f$ est continue en a.*

Preuve : La démonstration n'est pas très différente de celle du résultat analogue pour les fonctions d'une variable. Montrons (3) à titre d'exemple. Comme f est continue en a et que $f(a) \neq 0$, il existe une boule $B(a,r)$ telle que $f(x) \neq 0$ pour tout $x \in B(a,r)$, et $1/f$ est donc bien définie sur cette boule. De plus, quitte à prendre r plus petit, on peut supposer que $|1/f(x)| \leq 2|1/f(a)|$ pour tout $x \in B(a,r)$ (vérifier ces affirmations !). Soit $\epsilon > 0$. Prenons $0 < \delta \leq r$, et supposons $\|x-a\| < \delta$. Alors

$$\Big|\frac{1}{f(x)} - \frac{1}{f(a)}\Big| = \Big|\frac{f(x)-f(a)}{f(x)f(a)}\Big| \leq \frac{2}{(f(a))^2}|f(x)-f(a)|.$$

Par continuité de f, on a

$$|f(x)-f(a)| < \epsilon\frac{(f(a))^2}{2}$$

si $\delta > 0$ est pris assez petit, et donc $|1/f(x) - 1/f(a)| < \epsilon$.

Exemple : On identifie l'espace $\mathbb{R}^{n^2}$ à l'espace des matrices réelles $n \times n$. L'application

$$\mathbb{R}^{n^2} \mapsto \mathbb{R}, A \mapsto \det(A)$$

est continue sur $\mathbb{R}^{n^2}$ (c'est un polynôme).

Remarque : Avec les propositions de ce paragraphe, on peut aussi redémontrer autrement la continuité dans les exemples (2) et (3) du paragaphe précédent, et voir que la fonction définie dans l'exemple (4) du paragraphe précédent est continue sur $\mathbb{R}^2$.

Continuité et convergence de suites

Le théorème suivant montre que les fonctions continues préservent la convergence des suites :

Théorème : Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ et soit $a \in E$. La fonction f est continue en a si et seulement si, pour toute suite $(x_k)_{k\geq 1} \subset E$ qui converge vers a, la suite $(f(x_k))_{k\geq 1}$ converge vers $f(a)$.

Preuve :

Supposons d'abord que f est continue en a et considérons une suite $(x_k)_{k\geq 1} \subset E$ qui converge vers a. Soit $\epsilon > 0$. Par continuité de f en a, il existe $\delta > 0$ tel que si $\|x-a\| < \delta$ et $x \in E$, alors $\|f(x) - f(a)\| < \epsilon$. Or $x_k \to x$ et donc, il existe $M > 0$ tel que, pour tout $k \geq M$, $\|x_k - a\| < \delta$, d'où $\|f(x_k) - f(a)\| < \epsilon$, ce qui montre que $f(x_k) \to f(a)$.

Supposons ensuite que, pour toute suite $(x_k)_{k\geq 1} \subset E$ qui converge vers a, $(f(x_k))_{k\geq 1}$ converge vers $f(a)$, et montrons que f est continue en a. Raisonnons par l'absurde, et supposons que f ne soit pas continue en a. Donc il existerait $\epsilon > 0$ et une suite $(x_k)_{k\geq 1}$ telle que $\|x_k - a\| \leq 1/k$ et $\|f(x_k) - f(a)\| \geq \epsilon$. Ceci contredit l'hypothèse.

Exemple : Soit $a \in \mathbb{R}$ et soit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$, définie par

$$f(x,y) = \frac{xy^2}{x^2+y^4} \text{ si } (x,y) \neq (0,0), \qquad f(0,0) = a.$$

Cette fonction n'est pas continue en $(0,0)$ quelle que soit la valeur de a. En effet la suite $((1/k, 1/k))_{k\geq 1}$ converge vers $(0,0)$ et $f(1/k,1/k) \to 0$ pour $k \to \infty$, alors que la suite $((1/k^2, 1/k))_{k\geq 1}$ converge aussi vers $(0,0)$ mais $f(1/k^2, 1/k) = 1/2$ pour tout $k \in \mathbb{N}_*$.

Continuité et toplogie

Nous montrons d'abord un résultat assez utile pour montrer qu'un ensemble est fermé :

Proposition : Soit $E \subset \mathbb{R}^n$ un fermé, et soit une fonction continue $f : E \to \mathbb{R}$.

(1) *L'ensemble $F = \{x \in E : f(x) = 0\}$ est fermé.*

(2) *L'ensemble $G = \{(x, y) \in \mathbb{R}^{n+1} : x \in E, y = f(x)\}$ est fermé.*

Preuve :

(1) Montrons $\mathrm{adh}(F) \subset F$. Pour cela, considérons une suite convergente $(x_k)_{k\geq 1} \subset F$ de limite x, et montrons que $x \in F$. D'abord, comme E est fermé, $x \in E$, et $f(x)$ est donc bien défini. Comme f est continue $f(x_k)$ converge vers $f(x)$. Comme $f(x_k) = 0$ pour tout $k \in \mathbb{N}_*$, $f(x) = 0$ et donc $x \in F$.

(2) On se ramène au cas précédent en considérant la fonction

$$\tilde{f} : E \times \mathbb{R} \to \mathbb{R}, (x, y) \mapsto f(x) - y.$$

Remarque : On montre que réciproquement, si $F \subset \mathbb{R}^n$ est fermé, il existe une fonction continue $f : \mathbb{R}^n \to \mathbb{R}$ telle que $F = \{x \in \mathbb{R}^n : f(x) = 0\}$ (voir les exercices).

Exemple : On considère $\mathbb{R}^{n^2}$ comme l'ensemble des matrices $n \times n$. Le sous-ensemble O des matrices inversibles est ouvert. En effet, une matrice A n'est pas inversible si et seulement si son déterminant est nul. Le déterminant est une fonction continue sur $\mathbb{R}^n$, et donc l'ensemble des matrices qui ne sont pas inversibles est fermé.

On peut voir la proposition précédente comme un cas particulier d'une caractérisation plus générale :

Théorème : Soit un ouvert $E \subset \mathbb{R}^n$, et soit une fonction $f : E \to \mathbb{R}^m$. La fonction f est continue sur E si et seulement si, pour tout ouvert $O \subset \mathbb{R}^m$, $f^{-1}(O)$ est un ouvert.

Preuve :

Supposons d'abord que f est continue sur E. Soit O un ouvert de $\mathbb{R}^m$. Si $f^{-1}(O) = \varnothing$, $f^{-1}(O)$ est bien ouvert. Sinon, prenons $a \in f^{-1}(O)$ et montrons qu'il existe $r > 0$ tel que $B(a, r) \subset f^{-1}(O)$. Comme O est ouvert, il existe $\epsilon > 0$ tel que $B(f(a), \epsilon) \subset O$. Ensuite, comme E est ouvert et que f est continue, il existe $\delta > 0$ tel que, pour tout $x \in B(a, \delta)$, on ait $f(x) \in B(f(a), \epsilon)$. Dès lors $B(a, \delta) \subset f^{-1}(O)$ et on peut donc prendre $r = \delta$.

Supposons ensuite que, pour tout ouvert $O \subset \mathbb{R}^m$, $f^{-1}(O)$ est un ouvert. Soit $a \in E$. Fixons $\epsilon > 0$ et considérons l'ouvert $B(f(a), \epsilon)$. L'image inverse de cet ouvert est un ouvert contenant a, et il existe donc $\delta > 0$ tel que l'image de $B(a, \delta)$ soit contenue dans $B(f(a), \epsilon)$. Donc f est continue en a.

Remarque : En général, il n'est pas vrai que l'image d'un ouvert par une fonction continue est encore un ouvert. Par exemple $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ envoie l'ouvert $\mathbb{R}$ sur le fermé $[0, +\infty[$.

Continuité et compacité

Le théorème suivant montre que les fonctions continues préservent la compacité.

Théorème : Soit $E \subset \mathbb{R}^n$ un ensemble compact, et soit une fonction $f : E \to \mathbb{R}^m$ continue sur E. Alors $f(E)$ est un compact de $\mathbb{R}^m$.

Preuve : Montrons que toute suite $(y_k)_{k\geq 1} \subset f(E)$ admet une sous-suite convergente dans $f(E)$. Etant donné $(y_k)_{k\geq 1} \subset f(E)$, il existe une suite $(x_k)_{k\geq 1} \subset E$ telle que $y_k = f(x_k)$. Comme E est compact, $(x_k)_{k\geq 1}$ admet une sous-suite $(x_{g(j)})_{j\geq 1}$ qui converge vers un point $x \in E$. La suite $(f(x_{g(j)}))_{j\geq 1}$ est une sous-suite de $(f(x_k))_{k\geq 1} = (y_k)_{k\geq 1}$ qui converge vers $f(x)$ par continuité de f. Ceci montre que $f(E)$ est compact.

Remarque : La réciproque est fausse. Considérons la fonction $f : [0,3] \subset \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = x1_{[0,1[}(x) + (x-1)1_{[1,2]}(x) + (x-2)1_{]2,3]}(x).$$

L'ensemble $f([0,3]) = [0,1]$ est compact, mais f est discontinue en 1 et en 2.

Remarque : Le résultat est faux si 'compact' est remplacé partout par 'fermé' ou 'borné'.

Ce théorème est très utile en optimisation sous la forme du corollaire suivant :

Corollaire (théorème de Weierstrass) : Soit un compact $E \subset \mathbb{R}^n$ et soit une fonction continue $f : E \to \mathbb{R}$. La fonction f atteint ses bornes sur E.

Preuve : L'ensemble $f(E)$ est un ensemble fermé borné de $\mathbb{R}$. Il existe donc deux nombres $m, M \in f(E)$ tels que $m = \min\{y \in f(E)\}$ et $M = \max\{y \in f(E)\}$. Donc il existe $a, b \in E$ tels que $f(a) = m$ et $f(b) = M$, et donc tels que

$$f(x) \geq f(a) \quad \text{pour tout } x \in E \qquad \text{et} \qquad f(x) \leq f(b) \quad \text{pour tout } x \in E.$$

Exemple : Soit $E \subset \mathbb{R}^n$ un fermé non borné, et soit une fonction continue $f : E \to \mathbb{R}$ telle que

$$\lim_{\|x\|\to\infty} f(x) = +\infty$$

pour une certaine norme $\|\cdot\|$ sur $\mathbb{R}^n$. Une telle fonction est dite coercive. Montrons que la fonction f admet un minimum global.

Soit $a \in E$. Comme f est coercive, il existe $R > 0$ tel que, si $\|x\| > R$ et $x \in E$, on ait $f(x) > f(a)$. Par ailleurs, $\overline{B}(0,R) \cap E$ est un compact non vide, de sorte que f restreinte à cet ensemble admet un minimum global en un certain $a \in \overline{B}(0,R) \cap E$ (on a écrit $\overline{B}(0,R)$ pour $\mathrm{adh}(B(0,R))$). Or, d'après ce qui précède, f définie sur E atteint aussi son minimum global en a.

4. Différentiabilité : définitions

Une fonction est différentiable en un point si elle est bien approchée dans un voisinage de ce point par une fonction affine (sa tangente dans le cas d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$). Dans ce chapitre, nous partons de ce point de vue pour définir la différentiabilité d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$, puis nous donnons des interprétations géométrique et cinématique de la différentielle dans deux cas particuliers.

A partir d'ici, on convient que les points de $\mathbb{R}^n$ sont représentés par des vecteurs colonnes ; en particulier une application linéaire f de $\mathbb{R}^n$ dans $\mathbb{R}^m$ est représentée par une matrice $A \in \mathbb{R}^{m\times n}$ ($f : \mathbb{R}^n \to \mathbb{R}^m, x \mapsto Ax$).

Dérivée d'une fonction réelle d'une variable réelle

Soit $a \in \mathbb{R}$, et soit une fonction $f : \mathbb{R} \to \mathbb{R}$ définie sur un ouvert contenant a. On dit que f est différentiable en a s'il existe un nombre $f'(a)$ tel que

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = f'(a).$$

On peut réexprimer cette définition comme ceci : f est différentiable en a s'il existe un nombre $f'(a)$ et une fonction $\varepsilon : \mathbb{R} \to \mathbb{R}$ définie sur un voisinage de 0, continue en 0 et satisfaisant $\varepsilon(0) = 0$, tels que

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

pour tout h dans un voisinage de 0. C'est cette formulation que nous généraliserons pour définir la différentiabilité de fonctions de plusieurs variables.

Preuve : D'une part, si f est différentiable en a, on peut écrie pour tout $h \neq 0$ dans un voisinage de l'origine

$$\frac{f(a+h) - f(a)}{h} = f'(a) + \tilde{\varepsilon}(h)$$

où $\tilde{\varepsilon}(h) \to 0$ lorsque $h \to 0$. Dès lors, la fonction ε définie dans un voisinage de l'origine par

$$\varepsilon(h) = \frac{h}{|h|}\tilde{\varepsilon}(h) \quad \text{pour} \quad h \neq 0 \qquad \text{et} \qquad \varepsilon(0) = 0,$$

est bien continue en 0, satisfait $\varepsilon(0) = 0$, et on a bien $f(a+h) = f(a)+f'(a)h+|h|\varepsilon(h)$. Pour démontrer la réciproque, on note que $\varepsilon(h) \to 0$ lorsque $h \to 0$ implique $(|h|/h)\varepsilon(h) \to 0$ lorsque $h \to 0$.

Remarque : Si f est différentiable en a et si on note $h = x - a$, on obtient au voisinage de a :

$$f(x) = f(a) + f'(a)(x-a) + |x-a|\varepsilon(x-a).$$

En d'autres termes, f est différentiable en a si elle est "bien approchée" par sa tangente

$$T(x) = f(a) + f'(a)(x-a)$$

dans un voisinage du point a, au sens où $f(x) - T(x) = o(|x-a|)$ quand $x \to a$.

Différentiabilité d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^m$: définition

On part de ce dernier point de vue pour définir la différentiablité d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}^m$. Commençons par donner une définition dans un style un peu informel : on dit que f est différentiable en un point a si elle est "bien approchée" par une application affine aux alentours du point a, c'est-à-dire par une application de la forme

$$T(x) = f(a) + J_f(a)(x - a)$$

pour une certaine matrice $J_f(a) \in \mathbb{R}^{m\times n}$, au sens où $f(x) - T(x) = o(\|x - a\|)$ quand $x \to a$. La matrice $J_f(a)$ est appelée matrice jacobienne de f en a.

Passons ensuite à une définition plus formelle. Considérons un point $a \in \mathbb{R}^n$ et une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ définie sur un ouvert O contenant a. On dit que f est différentiable en a s'il existe une application linéaire

$$df(a) : \mathbb{R}^n \to \mathbb{R}^m, h \mapsto J_f(a)h$$

représenté par la matrice jacobienne $J_f(a) \in \mathbb{R}^{m\times n}$, et une fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}^m$ définie dans un voisinage de 0, continue en 0 et vérifiant $\varepsilon(0) = 0$, tels que

$$f(a + h) = f(a) + J_f(a)h + \|h\|\varepsilon(h)$$

pour tout h dans un voisinage de l'origine. Nous montrons ci-dessous que l'application $df(a)$ est unique et indépendante du choix de la norme ; on l'appelle la différentielle de f en a.

Preuve de l'unicité. Imaginons que pour deux applications linéaires, de matrices respectives A_1 et A_2, on ait, pour tout h dans un voisinage de l'origine,

$$f(a + h) = f(a) + A_1h + \|h\|\varepsilon_1(h),$$
$$f(a + h) = f(a) + A_2h + \|h\|\varepsilon_2(h),$$

où ε_1 et ε_2 sont continues et nulles en 0. Montrons que $A_1 = A_2$. Pour tout h dans un voisinage de l'origine, on a

$$(A_1 - A_2)h = \|h\|\big(\varepsilon_2(h) - \varepsilon_1(h)\big).$$

Prenant $h = te_k$ ($t \in \mathbb{R}$) pour n'importe quel vecteur de base e_k ($1 \le k \le n$), on obtient pour tout $t \neq 0$ assez petit,

$$(A_1 - A_2)e_k = \frac{|t|}{t}\|e_k\|\big(\varepsilon_2(te_k) - \varepsilon_1(te_k)\big).$$

Le membre de droite peut être rendu arbitrairement proche de 0 en prenant $t \to 0$, d'où $(A_1 - A_2)e_k = 0$ pour tout $1 \le k \le n$, d'où $A_1 = A_2$.

Preuve de l'indépendance du choix de la norme. Supposons que f soit différentiable en a pour la norme $\|\cdot\|$ et montrons que f l'est encore pour toute autre norme $\|\cdot\|'$. Pour tout $h \neq 0$ dans un voisinage de l'origine, on a

$$f(a + h) = f(a) + J_f(a)h + \|h\|\varepsilon(h) = f(a) + J_f(a)h + \|h\|'\frac{\|h\|}{\|h\|'}\varepsilon(h)$$

où ε s'annule et est continue en 0. Par équivalence des normes, il existe $M > 0$ tel que

$$\frac{\|h\|}{\|h\|'} \leq M \qquad \text{pour tout} \qquad h \neq 0.$$

Dès lors la fonction ε' définie dans un voisinage de l'origine par

$$\varepsilon'(h) = \frac{\|h\|}{\|h\|'}\varepsilon(h) \qquad \text{et} \qquad \varepsilon'(0) = 0$$

s'annule et est continue en 0.

Exemples :

(1) Une application constante f sur $\mathbb{R}^n$ est différentiable en tout point $a \in \mathbb{R}^n$ et sa différentielle est nulle. En effet $f(a+h) = f(a)$ pour tout $a, h \in \mathbb{R}^n$.

(2) Une application linéaire f sur $\mathbb{R}^n$ est différentiable en tout point $a \in \mathbb{R}^n$ et est égale à sa différentielle. En effet $f(a+h) = f(a) + f(h)$ pour tout $a, h \in \mathbb{R}^n$.

(3) Une forme quadratique f sur $\mathbb{R}^n$ est différentiable en tout point $a \in \mathbb{R}^n$. Munissons $\mathbb{R}^n$ de la norme $\|\cdot\|_2$. Ecrivons $f(x) = \langle x, Ax\rangle$, où A est une matrice symétrique $n \times n$. Alors

$$f(a+h) = \langle a+h, A(a+h)\rangle = f(a) + 2\langle Aa, h\rangle + \langle h, Ah\rangle.$$

On voit que la définition de différentiabilité est satisfaite en prenant

$$df(a) : \mathbb{R}^n \to \mathbb{R}, h \mapsto 2\langle Aa, h\rangle,$$

ainsi que

$$\varepsilon(h) = \frac{\langle h, Ah\rangle}{\|h\|_2} \text{ pour } h \neq 0 \qquad \text{et} \qquad \varepsilon(0) = 0.$$

Cette fonction est bien continue en 0 car, par l'inégalité de Cauchy-Schwarz et le fait que toute application linéaire est bornée,

$$\frac{|\langle h, Ah\rangle|}{\|h\|_2} \leq \|Ah\|_2 \leq M\|h\|_2,$$

pour une certaine constante $M > 0$.

(4) Aucune norme n'est différentiable en l'origine. En effet si c'était le cas pour une norme $\|\cdot\|$, il devrait exister une application linéaire, représentée par le vecteur $v \in \mathbb{R}^n$, telle que, dans le voisinage de l'origine,

$$\|h\| = \langle v, h\rangle + \|h\|\varepsilon(h), \qquad \|h\| = -\langle v, h\rangle + \|h\|\varepsilon(-h)$$

et donc, pour tout $h \neq 0$,

$$2 = \varepsilon(h) + \varepsilon(-h),$$

ce qui est incompatible avec $\varepsilon(h) \to 0$ si $h \to 0$.

Dérivées partielles et calcul de la matrice jacobienne
Soit $a \in \mathbb{R}^n$, et soit une fonction $g : \mathbb{R}^n \to \mathbb{R}$ définie dans un voisinage de a. Etant donné un vecteur e_k de la base canonique de $\mathbb{R}^n$ ($1 \leq k \leq n$), on définit la dérivée partielle de g en a dans la direction k par la limite

$$\frac{\partial g}{\partial x_k}(a) = \lim_{t \to 0} \frac{g(a + te_k) - g(a)}{t}$$

si cette limite existe. On peut lire cette définition comme ceci : la dérivée partielle de g par rapport à x_k est la dérivée de la fonction d'une variable $g(x_1, \ldots, x_{k-1}, \cdot, x_{k+1}, \ldots, x_n)$, obtenue en "gelant" toutes les variables autres que x_k, c'est-à-dire en les voyant comme des paramètres.

Soit à présent une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ différentiable en a. A l'aide de la notion de dérivée partielle, calculons les éléments de la matrice jacobienne $J_f(a)$, en commençant par deux cas particuliers.

Cas d'une f fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. La matrice $J_f(a)$ est une matrice ligne, et peut donc être représentée par un vecteur de $\mathbb{R}^n$. Ce vecteur s'appelle le gradient de f en a, et se note $\nabla f(a)$. Voyons que

$$\nabla f(a) = \Big(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\Big)^t,$$

et donc que

$$J_f(a)h = \langle \nabla f(a), h \rangle = \frac{\partial f}{\partial x_1}(a)h_1 + \cdots + \frac{\partial f}{\partial x_n}(a)h_n$$

pour tout $h \in \mathbb{R}^n$. En effet, prenant $h = te_k$ pour un vecteur de base e_k ($1 \leq k \leq n$) et $t \in \mathbb{R}$ dans un voisinage de l'origine, on déduit de la définition de différentiabilité que

$$f(a + te_k) = f(a) + (\nabla f(a))_k t + |t|\,\|e_k\|\varepsilon(te_k)$$

où ε est continue et s'annule en 0. Dès lors

$$(\nabla f(a))_k = \lim_{t \to 0} \frac{f(a + te_k) - f(a)}{t} = \frac{\partial f}{\partial x_k}(a).$$

Cas d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}^m$. La matrice $J_f(a)$ est une matrice colonne, et peut donc être représentée par un vecteur de $\mathbb{R}^m$. Si $f = (f_1, \ldots, f_m)^t$, on note

$$(J_f(a))_k = f_k'(a),$$

et de plus on écrit souvent $f'(a)$ au lieu de $J_f(a)$, car l'élément k de ce vecteur est simplement la dérivée de l'application f_k en a. En effet, pour tout h dans un voisinage de l'origine et pour tout $1 \leq k \leq n$,

$$f_k(a + h) = f_k(a) + (J_f(a))_k h + |h|\varepsilon_k(h)$$

où ε_k est continue et s'annule en 0. Dès lors, on trouve

$$(J_f(a))_k = \lim_{h \to 0} \frac{f_k(a + h) - f_k(a)}{h} = f_k'(a).$$

Cas général. Si $f : \mathbb{R}^n \to \mathbb{R}^m$, on peut combiner ce qu'on a trouvé dans les deux cas précédents, pour obtenir

$$(J_f(a))_{j,k} = (J_f(a)e_k)_j = \frac{\partial f_j}{\partial x_k}(a) = \lim_{t\to 0} \frac{f_j(a+te_k) - f_j(a)}{t},$$

si $f = (f_1, \ldots, f_m)^t$. Donc

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Remarque. Dans la définition de dérivée partielle en un point donnée plus haut, on ne demande pas que la fonction soit différentiable en ce point, et nous verrons effectivement des cas où les dérivées partielles sont bien définies, alors que la fonction n'est pas différentiable. De même, on peut parler de gradient ou de matrice jacobienne d'une fonction sans demander qu'elle soit différentiable (ce sont alors simplement des matrices contenant les dérivées partielles). Par contre, on ne parle de différentielle d'une fonction que lorsque celle-ci est différentiable.

Propriétés géométriques du gradient

Nous donnons deux propriétés géométriques très similaires du gradient, qui permettent de mieux visualiser ce vecteur. Soit un ouvert $E \subset \mathbb{R}^n$ non vide, et soit une fonction $f : E \to \mathbb{R}$ différentiable sur E. Prenons $a \in E$ et supposons que $\nabla f(a) \neq 0$.

Montrons d'abord que $\nabla f(a)$ est perpendiculaire à toute direction selon laquelle f reste constante dans l'approximation affine. Cette propriété est à l'origine de la méthode des multiplicateurs de Lagrange (voir le dernier chapitre de ce cours).

Etant donné $z \in \mathbb{R}$, la 'surface' de niveau z de la fonction f se définit comme

$$C(z) = \{x \in E : f(x) = z\}.$$

Prenons $z = f(a)$ de sorte que $a \in C(z) = C(f(a))$. Pour x dans un voisinage de a, on a

$$f(x) = f(a) + \langle \nabla f(a), x-a \rangle + \|x-a\|\varepsilon(x-a)$$

où ε est continue et s'annule en 0. L'ensemble des points x tels que f reste constante dans l'approximation affine satisfont donc la relation

$$\langle \nabla f(a), x-a \rangle = 0,$$

qui décrit le 'plan' tangent à la surface de niveau $C(f(a))$ au point a. Le gradient $\nabla f(a)$ est donc perpendiculaire à ce plan.

Remarque : les mots 'surface' et 'plan' sont bien adaptées au cas $n = 3$; on les remplace respectivement par 'courbe' et 'droite' si $n = 2$, et par 'hyper-surface' et 'hyper-plan' si $n > 3$.

Montrons ensuite que $\nabla f(a)$ pointe dans la direction où l'accroissement de f est maximal dans l'approximation affine, et qu'à normalisation près, c'est l'unique vecteur qui jouit de cette propriété. Cette observation est à la base de la méthode de descente de gradient.

Comme

$$f(a+h) = f(a) + \langle \nabla f(a), h\rangle + \|h\|\varepsilon(h),$$

où ε est continue et s'annule en 0, on cherche l'unique vecteur h_* tel que

$$\langle \nabla f(a), h_*\rangle = \max\{\langle \nabla f(a), h\rangle : \|h\|_2 = 1\}.$$

Le vecteur

$$h_* = \frac{\nabla f(a)}{\|\nabla f(a)\|_2}$$

résout ce problème : d'une part

$$\langle \nabla f(a), h_*\rangle = \|\nabla f(a)\|_2$$

et d'autre part, pour tout $h \in \mathbb{R}^n$ tel que $\|h\|_2 = 1$, l'inégalité de Cauchy-Schwarz implique

$$\langle \nabla f(a), h\rangle \leq \|\nabla f(a)\|_2.$$

Considérons ensuite h' qui résout aussi ce problème et montrons $h' = h_*$. On a

$$\langle \nabla f(a), h'\rangle = \langle \nabla f(a), h_*\rangle$$

et donc, divisant $\nabla f(a)$ par sa norme,

$$\langle h_*, h'\rangle = \langle h_*, h_*\rangle.$$

Dès lors

$$0 = \langle h_*, h_* - h'\rangle = \langle h_* - h', h_* - h'\rangle + \langle h', h_* - h'\rangle = \|h_* - h'\|^2 + \|h_*\|^2 - \|h'\|^2 = \|h_* - h'\|^2$$

puisque $\|h_*\|_2 = \|h'\|_2 = 1$. Donc $h' = h_*$.

Interprétation cinématique de la dérivée

Soit $I \subset \mathbb{R}$ un intervalle ouvert non vide, et soit une fonction $f : I \to \mathbb{R}^3$ différentiable sur I. Si on interprète la variable comme un temps $t \in I$ et $f(t)$ comme la position d'un objet ponctuel au temps t dans $\mathbb{R}^3$, le vecteur dérivée $f'(t) = (f_1'(t), f_2'(t), f_3'(t))$ s'interprète comme le vecteur vitesse de cet objet.

La norme $\|\cdot\|_2$ de ce vecteur s'interprète comme le taux de délplacement instantané du mobile :

$$\lim_{s\to 0} \frac{\|f(t+s) - f(t)\|_2}{|s|} = \|f'(t)\|_2$$

pour tout $t \in I$. En effet, comme f est différentiable en tout $t \in I$, il découle de la définition de différentiabilité que

$$\lim_{s\to 0} \frac{f(t+s) - f(t)}{s} = f'(t).$$

Le résultat suit alors de la continuité de la norme : la fonction $g : I \to \mathbb{R}^3$ définie dans un voisinage de 0 par

$$g(s) = \frac{f(t+s) - f(t)}{s} \quad \text{pour} \quad s \neq 0 \qquad \text{et} \qquad g(0) = f'(t)$$

est continue en 0 ; $\|g\|_2$ est donc aussi continue en 0, ce qui implique le résultat.

5. Différentiabilité : propriétés

Nous énonçons et démontrons à présent les propriétés de base des fonctions différentiables.

Différentiabilité et continuité
Soit E un ouvert de $\mathbb{R}^n$, soit $a \in E$ et soit $f : E \to \mathbb{R}$.

Proposition : Si f est différentiable en a, alors f est continue en a.

Preuve : D'après la définition de différentiabilité, on a, pour tout x dans un voisinage de a (prendre $h = x - a$) :

$$f(x) - f(a) = J_f(a)(x-a) + \|x-a\|\varepsilon(x-a),$$

où $\varepsilon(x-a) \to 0$ lorsque $x \to a$. Par ailleurs, on a vu que pour toute application linéaire représentée par une matrice A, il existe une contante $M > 0$ telle que, pour tout $y \in \mathbb{R}^n$, $\|Ay\| \leq M\|y\|$. Donc

$$|f(x) - f(a)| \leq (M + \varepsilon(x-a))\|x-a\| \leq (M+1)\|x-a\|,$$

en prenant x suffisamment proche de a pour que $|\varepsilon(x-a)| \leq 1$. Donc f est continue en a.

Opérations sur les fonctions différentiables
Nous montrons ici l'analogue pour la différentiabilité des propriétés que nous avions obtenues pour la continuité.

Proposition : Supposons $n \geq 2$. Soit $I \subset \mathbb{R}$ un ouvert, et soit $f : I \to \mathbb{R}$ différentiable en un point $a \in I$. Alors la fonction $F : I \times \mathbb{R}^{n-1} \to \mathbb{R}$ définie par $F(x) = f(x_1)$ est différentiable en tout point de la forme $(a, x_2, \ldots, x_n)$. De plus

$$\nabla F(a) = (f'(a), 0, \ldots, 0)^t.$$

Preuve : Il est commode de munir $\mathbb{R}^n$ de la norme $\|\cdot\|_\infty$. Soit $\overline{a} \subset \mathbb{R}^n$ un point de la forme $\overline{a} = (a, x_2, \ldots, x_n)$. Pour tout h dans un voisinage de $\overline{a}$, on a

$$F(\overline{a} + h) - F(\overline{a}) = f(a + h_1) - f(a) = f'(a)h_1 + |h_1|\tilde{\varepsilon}(h_1)$$

où $\tilde{\varepsilon}$ est continue et s'annule en 0. Dès lors

$$F(\overline{a} + h) - F(\overline{a}) = f'(a)h_1 + \|h\|_\infty \varepsilon(h)$$

où ε est définie dans un voisinage de l'origine par

$$\varepsilon(h) = \frac{|h_1|}{\|h\|_\infty}\tilde{\varepsilon}(h_1) \text{ si } h \neq 0, \qquad \varepsilon(0) = 0.$$

La continuité de ε en 0 découle de ce que $|h|/\|h\|_\infty \leq 1$.

Proposition : Soit $E \subset \mathbb{R}^n$ un ouvert et soit $a \in E$. Une fonction $f : E \to \mathbb{R}^m$ est différentiable en a si et seulement si toutes les fonctions $f_j : E \to \mathbb{R}$ sont différentiables en a $(1 \leq j \leq m)$, où on a noté $f = (f_1, \ldots, f_m)^t$.

Preuve : Si f est différentiable en a, on a pour tout $1 \leq j \leq m$, et pour tout h dans un voisinage de l'origine

$$f_j(a+h) - f_j(a) = \langle \nabla f_j(a), h \rangle + \|h\| \varepsilon_j(h)$$

où ε_j est continue et s'annule en 0. Donc toutes les fonctions f_j sont différentiables en a. Réciproquement, si toutes ces relations sont satisfaites, on trouve bien, pour tout h dans un voisinage de l'origine,

$$f(a+h) - f(a) = J_f(a)h + \|h\| \varepsilon(h),$$

où $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)^t$ est continue et s'annule en 0.

Remarque : dans le cas particulier $E \subset \mathbb{R}$ $(n = 1)$, ce résultat implique que f est différentiable en a si et seulement si $f_j'(a)$ est bien définie pour tout $1 \leq j \leq m$.

Proposition : Soit un ouvert $E \subset \mathbb{R}^n$ et soit $a \in E$.

(1) *Si $f, g : E \to \mathbb{R}$ sont différentiables en a, alors $f + g$ l'est aussi et*

$$\nabla(f+g)(a) = \nabla f(a) + \nabla f(a).$$

(2) *Si $f, g : E \to \mathbb{R}$ sont différentiables en a, alors fg l'est aussi et*

$$\nabla(fg)(a) = f(a)\nabla g(a) + g(a)\nabla f(a).$$

(3) *Si $f : E \to \mathbb{R}$ est différentiable en a et si $f(a) \neq 0$, alors $1/f$ est bien définie dans un voisinage de a et est différentiable en a. De plus*

$$\nabla(1/f)(a) = \frac{-1}{f^2(a)} \nabla f(a).$$

(4) *Si $f : E \to \mathbb{R}^m$ est différentiable en a et si $g : \mathbb{R}^m \to \mathbb{R}^p$ est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et*

$$J_{g \circ f}(a) = J_g(f(a)) J_f(a).$$

Preuve :
(1) C'est une application directe de la définition.
(2) Prenons $\|\cdot\| = \|\cdot\|_2$. Pour h dans un voisinage de l'origine, on développe

$$\begin{aligned}
&(fg)(a+h) - (fg)(a) \\
&\quad = \big(f(a+h) - f(a)\big) g(a+h) + f(a)\big(g(a+h) - g(a)\big) \\
&\quad = \big(\langle \nabla f(a), h \rangle + \|h\| \varepsilon_1(h)\big)\big(g(a) + \varepsilon_2(h)\big) + f(a)\big(\langle \nabla g(a), h \rangle + \|h\| \varepsilon_3(h)\big) \\
&\quad = \langle g(a)\nabla f(a) + f(a)\nabla g(a), h \rangle \\
&\qquad + \langle \nabla f(a), h \rangle \varepsilon_2(h) + \|h\| \big(\varepsilon_1(h) g(a) + \varepsilon_1(h)\varepsilon_2(h) + f(a)\varepsilon_3(h)\big)
\end{aligned}$$

où les fonctions ε_1, ε_2 et ε_3 sont continues et s'annulent en 0. Pour obtenir la deuxième égalité, on a utilisé la diférentiabilité de f et g en a, ainsi que la continuité de g en a, qui suit de la différentiabilité. Le résultat suit de ce que

$$\frac{|\langle \nabla f(a), h\rangle|}{\|h\|}|\varepsilon_2(h)| \leq \frac{\|\nabla f(a)\|\,\|h\|}{\|h\|}|\varepsilon_2(h)| = \|\nabla f(a)\||\varepsilon_2(h)| \to 0$$

et

$$\varepsilon_1(h)g(a) + \varepsilon_1(h)\varepsilon_2(h) + f(a)\varepsilon_3(h) \to 0$$

si $h \to 0$.

(3) La fonction f est différentiable, donc continue en a, et donc $1/f$ est bien définie dans un voisinage de a. Pour h dans un voisinage de l'origine, on développe

$$\frac{1}{f(a+h)} - \frac{1}{f(a)} = \frac{f(a) - f(a+h)}{f(a)f(a+h)} = -\frac{\langle \nabla f(a), h\rangle}{f(a)f(a+h)} + \|h\|\frac{\varepsilon_1(h)}{f(a+h)f(a)},$$

où ε_1 est continue et s'annule en 0. Le terme de reste est de la bonne forme : comme f est continue en a, $\varepsilon_1(h)/(f(a)f(a+h)) \to 0$ si $h \to 0$. Pour traiter le premier terme on développe encore

$$\frac{1}{f(a+h)} = \frac{1}{f(a)} + \frac{f(a) - f(a+h)}{f(a)f(a+h)} = \frac{1}{f(a)} + \frac{\varepsilon_2(h)}{f(a+h)f(a)}$$

où ε_2 est continue et s'annule en 0, et donc

$$\frac{\langle \nabla f(a), h\rangle}{f(a)f(a+h)} = \frac{\langle \nabla f(a), h\rangle}{f^2(a)} + \frac{\langle \nabla f(a), h\rangle \varepsilon_2(h)}{f^2(a)f(a+h)}.$$

Comme toute application linéaire est bornée, il existe $M > 0$ tel que, pour tout h dans un voisinage de l'origine,

$$|\langle \nabla f(a), h\rangle \varepsilon_2(h)| \leq M\|h\||\varepsilon_2(h)|,$$

et le terme de reste a donc bien encore une fois la forme voulue.

(4) Comme f est différentiable, et donc continue en a, pour tout voisinage W de $f(a)$, on trouve un voisinage V de a tel que l'image de V par f soit contenue dans W. Dès lors, dans un voisinage de a, on peut développer

$$\begin{aligned}
&g(f(a+h)) - g(f(a)) \\
&\quad = g\big(f(a) + (f(a+h) - f(a))\big) - g(f(a)) \\
&\quad = J_g(f(a))(f(a+h) - f(a)) + \|f(a+h) - f(a)\|\varepsilon_1(f(a+h) - f(a)) \\
&\quad = J_g(f(a))J_f(a)h + \|h\|J_g(f(a))\varepsilon_2(h) + \|f(a+h) - f(a)\|\varepsilon_1(f(a+h) - f(a)),
\end{aligned}$$

où ε_1 et ε_2 sont continues et s'annulent en 0. Montrons que les deux termes de reste ont la forme voulue. Pour le premier, on sait qu'il existe $M_1 > 0$ tel que, pour tout h dans un voisinage de l'origine,

$$\|J_g(f(a))\varepsilon_2(h)\| \leq M_1\|\varepsilon_2(h)\|\,;$$

donc $J_g(f(a))\varepsilon_2(h) \to 0$ quand $h \to 0$. Pour le second, on sait d'une part qu'il existe $M_2 > 0$ tel que

$$\|f(a+h) - f(a)\| \leq \|J_f(a)h\| + \|h\|\varepsilon_2(h) \leq \big(M_2 + \varepsilon_2(h)\big)\|h\|$$

et d'autre part, par continuité de f en a, que $\varepsilon_1(f(a+h) - f(a)) \to 0$ si $h \to 0$. Dès lors, pour tout $h \neq 0$, on peut écrire

$$\|f(a+h) - f(a)\|\varepsilon_1(f(a+h) - f(a)) = \|h\| \frac{\|f(a+h) - f(a)\|\varepsilon_1(f(a+h) - f(a))}{\|h\|}$$

et le second facteur tend vers 0 quand $h \to 0$.

Remarque : Considérons le cas particulier de (4) dans la proposition ci-dessus où $f, g : \mathbb{R}^n \to \mathbb{R}^n$ sont telles que $g \circ f(x) = x$ pour tout $x \in \mathbb{R}^n$, c'est-à-dire le cas où f et g sont réciproques l'une de l'autre. Alors on a

$$J_f(a) = \big(J_g(f(a))\big)^{-1}$$

où l'inverse est l'inverse matriciel.

Exemples :

(1) La fonction

$$f : \mathbb{R}^2 \to \mathbb{R}, (x,y)^t \mapsto (x^2 + y^2)\ln(x^2 + y^2 + 1)$$

est différentiable sur $\mathbb{R}^2$. En effet $f = g \circ h$, où

$$g : \mathbb{R} \to \mathbb{R}, z \to z\ln(z+1)$$

est différentiable sur l'ouvert $]-1, +\infty[$, et où

$$h : \mathbb{R}^2 \to \mathbb{R}_+, (x,y)^t \mapsto x^2 + y^2$$

est différentiable sur $\mathbb{R}^2$ (c'est un polynôme). On calcule

$$\frac{\partial f}{\partial x}(x,y) = g'(h(x,y))\frac{\partial h}{\partial x}(x,y) = \Big(\ln(x^2+y^2+1) + \frac{x^2+y^2}{x^2+y^2+1}\Big)2x,$$
$$\frac{\partial f}{\partial y}(x,y) = g'(h(x,y))\frac{\partial h}{\partial y}(x,y) = \Big(\ln(x^2+y^2+1) + \frac{x^2+y^2}{x^2+y^2+1}\Big)2y.$$

(2) Soient les fonctions

$$g : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \ln x \\ x - y \end{pmatrix}$$

et

$$h : \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 + y^2 \\ 2x \end{pmatrix}.$$

La fonction $f = g \circ h$ est différentiable sur l'ouvert $\mathbb{R}^2 \backslash \{(0,0)\}$, et sa matrice jacobienne est donnée par

$$\begin{aligned} J_{g\circ h}(x,y) &= J_g(h(x,y))J_h(x,y) \\ &= \begin{pmatrix} \frac{1}{x^2+y^2} & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2x & 2y \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} \frac{2x}{x^2+y^2} & \frac{2y}{x^2+y^2} \\ 2x-2 & 2y \end{pmatrix}. \end{aligned}$$

Egalité et inégalité de la moyenne

Commençons par rappeler le théorème de la moyenne pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$ (voir p. 80 du cours de L1 de F. Simenhaus pour une démonstration).

Théorème : Soit $f : [a,b] \to \mathbb{R}$ ($a < b$), continue sur $[a,b]$ et dérivable sur $]a,b[$. Alors il existe $c \in]a,b[$ tel que

$$f(b) - f(a) = f'(c)(b-a).$$

Ce théorème se généralise aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. Si $a, b \in \mathbb{R}^n$, le segment $[a,b] \subset \mathbb{R}^n$ se définit par

$$[a,b] = \{(1-t)a + tb, t \in [0,1]\}.$$

Le segment $]a,b[$ se définit en remplaçant $[0,1]$ par $]0,1[$.

Proposition : Soit $E \subset \mathbb{R}^n$ un ouvert, et soient $a, b \in E$ tels que $[a,b] \subset E$. Soit $f : E \to \mathbb{R}$, continue sur $[a,b]$ et différentiable sur $]a,b[$. Alors il existe $c \in]a,b[$ tel que

$$f(b) - f(a) = \langle \nabla f(c), b-a \rangle.$$

Preuve : Considérons la fonction d'une variable définie par

$$F : [0,1] \to \mathbb{R}, t \mapsto f\big((1-t)a + tb\big).$$

Cette fonction est continue sur $[0,1]$ et différentiable sur $]0,1[$. Il existe donc $\theta \in]0,1[$ tel que $F(1) - F(0) = F'(\theta)$. Or, pour tout $t \in]0,1[$

$$F'(t) = \big\langle \nabla f\big((1-t)a + tb\big), b-a \big\rangle,$$

ce qui donne le résultat avec $c = (1-\theta)a + \theta b$.

Le théorème de la moyenne ne se généralise pas aux fonctions de $\mathbb{R}$ dans $\mathbb{R}^m$ sous la forme d'une égalité comme donné plus haut. Considérons en effet la fonction

$$f : [0, 2\pi] \to \mathbb{R}^2, x \mapsto (\cos x, \sin x)^t.$$

Le vecteur dérivée est donné par $f(x) = (-\sin x, \cos x)^t$ pour tout $x \in]0, 2\pi[$, et n'est jamais égal au vecteur nul. Donc il n'existe pas de $c \in]0, 2\pi[$ tel que $f(2\pi) - f(0) = f'(c)2\pi$. Par contre, on peut donner une généralisation sous forme d'une inégalité :

Théorème : Soit $E \subset \mathbb{R}^n$ un ouvert, et soient $a, b \in E$ tels que $[a, b] \subset E$. Soit $f : E \to \mathbb{R}^m$, continue sur $[a, b]$ et différentiable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que

$$\|f(b) - f(a)\| \leq \|J_f(c)(b - a)\|.$$

En particulier

$$\|f(b) - f(a)\| \leq \sup_{c \in E} \|J_f(c)\| \, \|b - a\|,$$

où $\|J_f(c)\|$ est la norme opérateur de $J_f(c)$.

Preuve : Donnons la preuve dans le cas particulier où $\|\cdot\| = \|\cdot\|_2$ (pour la démonstration générale, voir p. 51 du cours de F. Bolley référencé au début de ce cours). Le résultat est immédiat si $f(b) = f(a)$ et nous supposons $f(b) \neq f(a)$. Etant donné $x \in \mathbb{R}^m$ tel que $\|x\|_2 = 1$, on considère une fonction réelle d'une variable réelle :

$$\phi : [0, 1] \to \mathbb{R} : t \mapsto \phi(t) = \langle x, f((1 - t)a + tb) \rangle.$$

Cette fonction est continue sur $[0, 1]$ et différentiable sur $]0, 1[$. Dès lors, il existe $\theta \in]0, 1[$ tel que

$$\phi(1) - \phi(0) = \phi'(\theta) = \langle x, J_f((1 - \theta)a + \theta b)(b - a) \rangle.$$

Alors, d'une part, comme $\|x\|_2 = 1$, l'inégalité de Cauchy-Schwarz implique

$$\langle x, J_f((1 - \theta)a + \theta b)(b - a) \rangle \leq \|J_f((1 - \theta)a + \theta b)(b - a)\|_2.$$

Et d'autre part, prenant $x = (f(b) - f(a))/\|f(b) - f(a)\|_2$, on a

$$\phi(1) - \phi(0) = \|f(b) - f(a)\|_2.$$

On obtient donc le résultat en prenant $c = (1 - \theta)a + \theta b$.

Dérivées partielles et différentiabilité

Il arrive que toutes les dérivées partielles d'une fonction soient définies en un point, mais que f ne soit pas différentiable en ce point. En effet, prenons par exemple

$$f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto \begin{cases} 0 & \text{si } x = 0 \text{ ou } y = 0, \\ 1 & \text{sinon.} \end{cases}$$

On a $f(0, y) = 0$ pour tout $y \in \mathbb{R}$ et $f(x, 0) = 0$ pour tout $x \in \mathbb{R}$, et dès lors

$$\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial x}(0, 0) = 0.$$

Or f est discontinue en $(0, 0)$ et n'est donc pas différentiable en $(0, 0)$.

Néanomoins, sous des hypothèses supplémentaires, l'existence des dérivées partielles garantit la différentiabilité :

Théorème : Soit $E \subset \mathbb{R}^n$ un ouvert et soit $f : E \to \mathbb{R}$. Si f admet des dérivées partielles en tout point de E, et que toutes les dérivées partielles sont continues sur E, alors f est différentiable en tout point de E.

Preuve : Considérons le cas particulier $n = 2$ (la généralisation au cas n quelconque ne pose aucune difficulté conceptuelle). Prenons $a = (a_1, a_2) \in E$, et montrons que f est différentiable en a. Comme E est ouvert, il existe un voisinage de l'origine tel que, pour tout $h = (h_1, h_2)$ dans ce voisinage, $a + h \in E$. Prenons h dans ce voisinage, et écrivons

$$f(a+h) - f(a) = \big(f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2)\big) + \big(f(a_1 + h_1, a_2) - f(a_1, a_2)\big).$$

On utilise d'abord l'hypothèse que f admet des dérivées partielles en tout point de E. Supposons $h_1, h_2 > 0$ (autres cas analogues). D'après le théorème de la moyenne pour les fonctions d'une variable, il existe $c_1 \in]0, h_1[$ et $c_2(h_1) \in]0, h_2[$ tels que

$$\begin{aligned} f(a_1 + h_1, a_2) - f(a_1, a_2) &= \frac{\partial f}{\partial x_1}(a_1 + c_1, a_2)h_1, \\ f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) &= \frac{\partial f}{\partial x_2}(a_1 + h_1, a_2 + c_2(h_1))h_2. \end{aligned}$$

On utilise ensuite l'hypothèse que les dérivées partielles sont continues :

$$\begin{aligned} \frac{\partial f}{\partial x_1}(a_1 + c_1, a_2) &= \frac{\partial f}{\partial x_1}(a_1, a_2) + \varepsilon_1(c_1), \\ \frac{\partial f}{\partial x_2}(a_1 + h_1, a_2 + c_2(h_1)) &= \frac{\partial f}{\partial x_2}(a_1, a_2) + \varepsilon_2(h_1, c_2(h_2)), \end{aligned}$$

où la fonction ε_1 est définie dans un voisinage de 0, et est continue et s'annule en 0, et où la fonction ε_2 est définie dans un voisinage de $(0,0)$, et est continue et s'annule en $(0,0)$.

Rassemblant les estimations ci-dessus, on trouve

$$f(a+h) - f(a) = \frac{\partial f}{\partial x_1}(a_1, a_2)h_1 + \frac{\partial f}{\partial x_2}(a_1, a_2)h_2 + \big(h_1\varepsilon_1(c_1) + h_2\varepsilon_2(h_1, c_2(h_2))\big).$$

Pour $(h_1, h_2) \neq (0,0)$, écrivons le terme de reste sous la forme

$$h_1\varepsilon_1(c_1) + h_2\varepsilon_2(h_1, c_2(h_2)) = \|h\| \frac{h_1\varepsilon_1(c_1) + h_2\varepsilon_2(h_1, c_2(h_2))}{\|h\|}$$

et montrons que le second facteur du membre de droite tend vers 0 quand $(h_1, h_2) \to (0,0)$. Par commodité, supposons que $\|\cdot\| = \|\cdot\|_\infty$; alors $|h_1|, |h_2| \leq \|h\|_\infty$ d'où

$$\frac{|h_1\varepsilon_1(c_1) + h_2\varepsilon_2(h_1, c_2(h_2))|}{\|h\|} \leq |\varepsilon_1(c_1)| + |\varepsilon_2(h_1, c_2(h_2))|.$$

Comme ε_1 et ε_2 sont continues et s'annulent en l'origine, et que $|c_1| \leq |h_1|$ et $|c_2| \leq |h_2|$, on conclut que le membre de droite de cette dernière expression tend vers 0 quand $(h_1, h_2) \to (0,0)$.

Terminologie : Si $E \subset \mathbb{R}^n$ est ouvert, si $f : E \to \mathbb{R}$ admet des dérivées partielles continues en tout point de E, on dit que f est de classe $\mathcal{C}^1$ sur E.

Remarque : Si $E \subset \mathbb{R}^n$ est un ouvert et que f est diférentiable sur E, il n'est en général pas vrai que ses dérivées partielles sont continues. Considérons par exmple la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = x^2 \sin(1/x) \quad \text{pour} \quad x \neq 0, \qquad f(0) = 0.$$

Cette fonction est différentiable sur $\mathbb{R}$: pour tout $x \neq 0$, on a

$$f'(x) = 2x \sin(1/x) - \cos(1/x),$$

et on a $f'(0) = 0$ car

$$f(h) - f(0) = h^2 \sin(1/h) = |h|\,|h| \sin(1/h)$$

et $||h| \sin(1/h)| \leq |h| \to 0$ si $h \to 0$. Or on voit que f' est discontinue en 0 : la suite $(x_k)_{k\geq 1} = (1/2\pi k)_{k\geq 1}$ tend vers 0, mais $f'(x_k) = -1$ pour tout $k \geq 1$.

Derivée d'une intégrale dépendant d'un paramètre
Nous appliquons plusieurs résultats de ce chapitre pour étudier la dérivée d'une intégrale dépendant d'un paramètre.

Proposition : Soit $f : [a,b] \times E \to \mathbb{R}, (x,y) \mapsto f(x,y)$, où $a < b \in \mathbb{R}$, et où $E \subset \mathbb{R}$ est un ouvert. Supposons que f soit continue sur $[a,b] \times E$, et que $\partial f/\partial y$ existe et soit continue sur $[a,b] \times E$. Alors la fonction

$$g : E \to \mathbb{R}, y \mapsto \int_a^b f(x,y)\,\mathrm{d}x$$

est de classe $\mathcal{C}^1$ sur E et, pour tout $y \in E$, on a

$$g'(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\mathrm{d}x.$$

Preuve : Soit $y \in E$ (dans ce qui suit on ne note pas toujours explicitement la dépendance en y). Pour h dans un voisinage de 0, le théorème de la moyenne garantit l'existence d'un nombre $c(x,h) \in]0,h[$ (on suppose que $h > 0$, l'autre cas est analogue) tel que

$$g(y+h) - g(y) = \int_a^b \big(f(x,y+h) - f(x,y)\big)\,\mathrm{d}x = \int_a^b \frac{\partial f}{\partial y}(x, y + c(x,h))h\,\mathrm{d}x.$$

Comme $\partial f/\partial y$ est continue sur $[a,b] \times E$, on développe pour tout z dans un voisinage de y,

$$\frac{\partial f}{\partial y}(x, y+z) = \frac{\partial f}{\partial y}(x,y) + \varepsilon(x,z)$$

où ε est continue et s'annule sur $[a,b] \times \{0\}$. On a donc aussi

$$g(y+h) - g(y) = \Big(\int_a^b \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x\Big)h + h\int_a^b \varepsilon(x, c(x,h))\,\mathrm{d}x.$$

Pour montrer que g est différentiable, avec la dérivée annoncée, il faut voir que

$$\int_a^b \varepsilon(x, c(x,h))\,\mathrm{d}x \to 0$$

si $h \to 0$. Pour ce faire, montrons que

$$\max\{|\varepsilon(x, c(x,h))|, x \in [a,b]\} \to 0$$

si $h \to 0$. Supposons par l'absurde que ce ne soit pas le cas. Alors il existerait $\delta > 0$, ainsi qu'une suite $z_k = (x_k, h_k)_{k\geq 1}$ avec $x_k \in [a,b]$ et $h_k \to 0$, tels que $|\varepsilon(x_k, c(x_k, h_k))| \geq \delta$. Comme $[a,b]$ est compact, il existerait alors aussi une sous-suite $(x_{\phi(j)})_{j\geq 1}$ de $(x_k)_{k\geq 1}$ qui convergerait vers un point $\overline{x} \in [a,b]$. Mais ε est continue en $(\overline{x}, 0)$ et s'annule en ce point, ce qui mène à une contradiction.

Il reste à voir que g' est continue sur E. Etant donné $y \in E$ et h dans un voisinage de l'origine, on écrit

$$g'(y+h) - g'(y) = \int_a^b \Big(\frac{\partial f}{\partial y}(x, y+h) - \frac{\partial f}{\partial y}(x,y)\Big)\,\mathrm{d}x = \int_a^b \tilde{\varepsilon}(x,h)\,\mathrm{d}x$$

où $\tilde{\varepsilon}$ est définie et continue dans un voisinage de $[a,b] \times \{0\}$, et s'annule sur $[a,b] \times \{0\}$. On conclut comme dans la partie précédente : on montre que $\max\{|\tilde{\varepsilon}(x,h)|, x \in [a,b]\} \to 0$ lorsque $h \to 0$.

Proposition : Soient $[a,b]$, E et f comme dans la proposition précédente, et soit aussi une fonction $\beta : E \to]a,b[$ de classe $\mathcal{C}^1$ sur E. Alors la fonction

$$g : E \to \mathbb{R}, y \mapsto \int_a^{\beta(y)} f(x,y)\,\mathrm{d}x$$

est de classe $\mathcal{C}^1$ sur E et, pour tout $y \in E$, on a

$$g'(y) = f(\beta(y), y)\beta'(y) + \int_a^{\beta(y)} \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x.$$

Preuve : Considérons d'abord la fonction de deux variables

$$G : E \times]a,b[\to \mathbb{R}, (y,z) \mapsto \int_a^z f(x,y)\,\mathrm{d}x.$$

Les dérivées partielles de G existent et sont continues sur $E \times]a,b[$:

$$\frac{\partial G}{\partial y}(y,z) = \int_a^z \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x,$$
$$\frac{\partial G}{\partial z}(y,z) = f(z,y).$$

Donc G est de classe $\mathcal{C}^1$ sur $E \times]a,b[$.

Considérons ensuite la fonction

$$B : E \to E \times]a,b[, y \mapsto (y, \beta(y)),$$

qui est de classe $\mathcal{C}^1$ sur E.

Comme $g = G \circ B$, g est différentiable sur E et, pour tout $y \in E$,

$$g'(y) = \frac{\partial G}{\partial y}(B(y))1 + \frac{\partial G}{\partial z}(B(y))\beta'(y) = \int_a^{\beta(y)} \frac{\partial f}{\partial y}(x,y)\,\mathrm{d}x + f(\beta(y),y)\beta'(y).$$

La continuité de g' suit de la proposition précédente et des résultats sur les opérations sur les fonctions différentiables.

6. Différentiabilité : second ordre

A partir de ce chapitre, on se restreint aux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. Premièrement, cela nous évitera des notations très lourdes (la dérivée d'une fonction à valeurs vectorielles est une application linéaire, et sa dérivée seconde devient donc une application linéaire à valeurs dans les applications linéaires). Ensuite, une fonction à valeurs vectorielles n'est qu'une collection de fonctions à valeurs réelles, et beaucoup de résultats énoncés pour des fonctions à valeurs réelles se généralisent sans problème (mais attention au théorème de la moyenne !). Enfin, ce sont les fonctions qui se révéleront les plus intéressantes pour les applications (au sein de ce cours).

Définition
Soit une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ et soit $a \in \mathbb{R}^n$. On dit que f est deux fois différentiable en a si f est définie et est différentiable dans un voisinage E de a, et si $\nabla f : E \to \mathbb{R}^n$ est différentiable en a.

La matrice jacobienne de ∇f en a s'appelle matrice hessienne de f en a et se note $H_f(a)$:

$$H_f(a) = J_{\nabla f}(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix},$$

où, pour tout $1 \le j \le k \le n$, on a écrit

$$\frac{\partial^2 f}{\partial x_j^2} = \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_j}\Big) \qquad \text{et} \qquad \frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_k}\Big).$$

Exemples :

(1) Soit $a \in \mathbb{R}^n$ et soit $b \in \mathbb{R}$. La fonction affine $f : \mathbb{R}^n \to \mathbb{R}$ définie par $f(x) = \langle a, x\rangle + b$ est deux fois différentiable en tout point $x \in \mathbb{R}^n$, et sa matrice hessienne est partout nulle. En effet, on calcule que $\nabla f(x) = a$ pour tout $x \in \mathbb{R}^n$, puisque pour tout $h \in \mathbb{R}^n$,
$$f(x+h) = \langle a, x\rangle + b + \langle a, h\rangle = f(x) + \langle \nabla f(x), h\rangle.$$
Dès lors ∇f est partout différentiable et de différentielle nulle.

(2) Soit A une matrice symétrique $n \times n$. La forme quadratique $f : \mathbb{R}^n \to \mathbb{R}$ définie par $f(x) = \langle x, Ax\rangle$ est deux fois différentiable en tout point $x \in \mathbb{R}^n$, et sa matrice hessienne vaut $2A$. En effet, on calcule d'abord que $\nabla f(x) = 2Ax$ (voir les exemples du chapitre 4). Donc ∇f est linéaire, et sa matrice jacobienne vaut $2A$ (voir les exemples du chapitre 4).

Symétrie de la matrice hessienne (théorème de Schwarz)
Soit $a \in \mathbb{R}^n$ et soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. On admet le résulat suivant (pour une démonstation, voir par exemple Dieudonné, p. 180).

Théorème : Si f est deux fois différentiable en a, alors sa matrice hessienne est symétrique : pour tout $1 \leq i < j \leq n$,

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

“Contre-exemple” : Considérons l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

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

On vérifie que toutes les dérivées partielles secondes existent en tout point. En particulier

$$\frac{\partial f}{\partial x}(0,y) = -y \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x,0) = x.$$

Dès lors,

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1 \qquad \text{et} \qquad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1.$$

On conclut que f n'est pas deux fois différentiable en $(0,0)$.

Développement limité au second ordre

Au chapitre 4, le développement limité au premier ordre avait servi de point départ pour définir la différentiabilité ; le développement au second ordre apparaît comme une conséquence de la définition des fonctions deux fois différentiables. Soit $a \in \mathbb{R}^n$ et soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$.

Théorème : Si f est deux fois différentiable en a, on a

$$f(x) = f(a) + \langle \nabla f(a), x-a \rangle + \frac{1}{2}\langle x-a, H_f(a)(x-a) \rangle + \|x-a\|^2 \varepsilon(x-a)$$

où ε est continue et s'annule en l'origine.

Preuve : Définissons une application

$$\phi : [0,1] \to \mathbb{R}, t \mapsto \phi(t) = f(tx + (1-t)a).$$

Alors

$$f(x) - f(a) = \phi(1) - \phi(0) = \int_0^1 \phi'(t)\,\mathrm{d}t = \int_0^1 \langle \nabla f(tx + (1-t)a), x-a \rangle\,\mathrm{d}t.$$

Ecrivant $tx + (1-t)a = a + t(x-a)$, on développe alors

$$\nabla f(a + t(x-a)) = \nabla f(a) + tH_f(a)(x-a) + \|x-a\|\varepsilon(t(x-a))$$

et donc

$$f(x) - f(a) = \langle \nabla f(a), (x-a) \rangle + \frac{1}{2}\langle x-a, H_f(a)(x-a)\rangle$$
$$+ \|x-a\| \int_0^1 \langle \varepsilon(t(x-a)), x-a \rangle \,\mathrm{d}t.$$

Il reste à voir que le terme de reste est $o(\|x-a\|^2)$. Par commodité prenons $\|\cdot\| = \|\cdot\|_2$; par l'inégalité de Cauchy-Schwartz, on obtient

$$\|x-a\|_2 \left| \int_0^1 \langle \varepsilon(t(x-a)), x-a \rangle \,\mathrm{d}t \right| \leq \|x-a\|_2^2 \int_0^1 \|\varepsilon(t(x-a))\|_2 \,\mathrm{d}t.$$

Comme $\max_{0\leq t\leq 1} \|\varepsilon(t(x-a))\|_2 \to 0$ si $x \to a$, l'intégrale tend bien vers 0 si $x \to a$.

Remarque sur la démonstration : L'usage de la formule

$$\phi(1) - \phi(0) = \int_0^1 \phi'(t) \,\mathrm{d}t$$

n'est pas évident à justifier, car il n'est pas a priori clair que ϕ' est intégrable. La suite montre que ϕ' est bornée pour $\|x-a\|$ assez petit, et donc intégrable (au sens de Lebesgue). Ceci garantit la validité de la formule utilisée.

Sous des hypothèses un peu plus fortes, on peut donner une expression du terme de reste analogue à celle que fournit le théorème de la moyenne au premier ordre (expression du reste de Lagrange). Rappelons d'abord le résultat pour les fonctions d'une variable (voir p. 85 du cours de L1 de F. Simenhaus) :

Théorème : Soient $a < b \in \mathbb{R}$, et soit une fonction f de classe $\mathcal{C}^1$ sur $[a,b]$ et deux fois dérivable sur $]a,b[$. Alors il existe $c \in]a,b[$ tel que

$$f(b) - f(a) - f'(a)(b-a) = \frac{f''(c)}{2}(b-a)^2.$$

Généralisons ce résultat aux fonctions sur $\mathbb{R}^n$:

Théorème : Soient $a, b \in \mathbb{R}^n$ et soit une fonction $f : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^1$ sur un ouvert contenant le segment $[a,b]$, et deux fois différentiable sur $]a,b[$. Alors il existe $c \in]a,b[$ tel que

$$f(b) - f(a) - \langle \nabla f(a), b-a \rangle = \frac{1}{2}\langle b-a, H_f(c)(b-a)\rangle.$$

Preuve : A nouveau définissons une application

$$\phi : [0,1] \to \mathbb{R}, t \mapsto \phi(t) = f(tb + (1-t)a).$$

Cette fonction satisfait aux hyopthèses du théorème précédent, et il existe donc $\theta \in]0,1[$ tel que

$$\phi(1) - \phi(0) - \phi'(0) = \frac{\phi''(\theta)}{2}.$$

Or

$$\phi'(0) = \langle \nabla f(a), (b-a) \rangle,$$
$$\phi''(\theta) = \langle H_f(\theta b + (1-\theta)a)(b-a), b-a \rangle.$$

On trouve donc le résultat en prenant $c = \theta b + (1-\theta)a$.

Rappel sur les formes quadratiques

Une des premières raisons de considérer un développement au second ordre au voisinage d'un point a, est de décider si, dans un voisinage de ce point, les valeurs de la fonctions sont supérieures ou inférieures au résultat de l'approximation affine. Nous aurons d'abord besoin de quelques rappels sur les formes quadratiques.

Soit A une matrice symétrique $n \times n$.

(1) On dit que A est définie positive si $\langle x, Ax \rangle > 0$ pour tout $\mathbb{R}^n \ni x \neq 0$.

(2) On dit que A est semi-définie positive si $\langle x, Ax \rangle \geq 0$ pour tout $\mathbb{R}^n \ni x \neq 0$.

(3) On dit que A est (semi)-définie négative si $-A$ est (semi)-définie positive.

(4) On dit que A est indéfinie s'il existe $x, y \in \mathbb{R}^n$ tels que $\langle x, Ax \rangle > 0$ et $\langle y, Ay \rangle < 0$.

Il y a principalement deux façons de déterminer le type d'une matrice (voir le cours d'algèbre du premier semestre pour plus de détails) :

1. Par calcul des valeurs propres de A. On obtient un critère simple : A est définie positive si toutes les valeurs propres sont strictement positives, est semi-définie positive si toutes les valeurs propres sont positives ou nulles, est indéfinie s'il existe au moins une valeur propre strictement positive et une valeur propre strictement négative.

Ce critère est simple à démontrer. Soient $\lambda_1, \ldots, \lambda_n$ les n valeurs propres de A, et soient $\psi_1, \ldots, \psi_n$ des vecteurs propres correspondants (on suppose qu'ils forment une base orthonormée de $\mathbb{R}^n$). Soit aussi O une matrice orthogonale telle que $OAO^t = D$ où D est diagonale (les éléments diagonaux de D sont les valeurs propres de A, et les lignes de O les vecteurs propres).

Alors, que la condition soit nécessaire suit de ce que $\langle \psi_i, A\psi_i \rangle = \lambda_i$ pour tout $1 \leq i \leq n$. Ensuite, que la condition soit suffisante résulte de ce que, pour tout $x \in \mathbb{R}^n$,

$$\langle x, Ax \rangle = \langle x, O^t OAO^t Ox \rangle = \langle Ox, DOx \rangle = \sum_{i=1}^{n} \lambda_i (Ox)_i^2.$$

2. Par le procédé de Gauss (complétion des carrés). Le critère est moins simple à exprimer mais il est souvent plus facile à mettre en œuvre que le précédent. A titre d'exemple, prenons le cas particulier $n = 2$. Soit

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} ;$$

la forme quadratique est donnée par

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

Nous affirmons que

(1) si $\det(A) > 0$ et $a, c > 0$, alors A est définie positive,

(2) si $\det(A) > 0$ et $a, c < 0$, alors A est définie négative,

(3) si $\det(A) = 0$ et $a, c \geq 0$, alors A est semi-définie positive,

(4) si $\det(A) = 0$ et $a, c \leq 0$, alors A est semi-définie négative,

(5) si $\det(A) < 0$, alors A est indéfinie

(ces cinq cas couvrent toutes les possibilités). Montrons ces affirmations à l'aide du procédé de Gauss, et considérons le cas particulier $a \neq 0$ (les autres cas sont analogues). On calcule

$$\begin{aligned} f(x_1, x_2) &= a\Big(x_1^2 + 2x_1\frac{b}{a}x_2 + \frac{c}{a}x_2^2\Big) = a\Big(x_1^2 + 2x_1\frac{b}{a}x_2 + \frac{b^2}{a^2}x_2^2 - \frac{b^2}{a^2}x_2^2 + \frac{c}{a}x_2^2\Big) \\ &= a\Big(\Big(x_1 + \frac{b}{a}x_2\Big)^2 + \frac{ac - b^2}{a^2}x_2^2\Big) \\ | &= a\Big(x_1 + \frac{b}{a}x_2\Big)^2 + \frac{\det(A)}{a}x_2^2. \end{aligned}$$

De là on peut analyser les cinq cas donnés plus haut.

Remarque : En plus de ces deux méthodes, la règle de Sylvester permet de déterminer si A est définie positive : A est définie positive si et seulement si les n déterminants des matrices $k \times k$ obtenues en supprimant les $n - k$ dernières lignes et les $n - k$ dernières colonnes sont stristement positifs ($1 \leq k \leq n$).

Position du graphe de f par rapport à l'approximation affine.
Nous donnons deux critères qui permettent de déterminer la position du graphe de f par rapport à l'approximation affine ; le premier est strictement ponctuel et le second local. Soit $a \in \mathbb{R}^n$ et soit une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$. On suppose f différentiable en a, et on note T son approximation affine en a :

$$T(x) = f(a) + \langle \nabla f(a), (x - a) \rangle.$$

Proposition : Supposons que f soit deux fois différentiable en a. Si $H_f(a)$ est définie positive, alors $f(x) > T(x)$ pour tout $x \neq a$ dans un voisinage de a ; si $H_f(a)$ est définie négative, alors $f(x) < T(x)$ pour tout $x \neq a$ dans un voisinage de a ; si $H_f(a)$ est indéfinie, alors, dans tout voisinage de a, il existe $x, y \in \mathbb{R}^n$ tels que $f(x) > T(x)$ et $f(y) < T(y)$.

Preuve : Comme f est deux fois différentiable en a, on a, pour tout x dans un voisinage de l'origine,

$$f(x) = f(a) + \langle \nabla f(a), x - a \rangle + \frac{1}{2}\langle x - a, H_f(a)(x - a)\rangle + \|x - a\|^2 \varepsilon(x - a)$$

où ε est continue et s'annule en l'origine. Donc

$$f(x) - T(x) = \frac{1}{2}\langle x - a, H_f(a)(x - a)\rangle + \|x - a\|^2 \varepsilon(x - a)$$

et on cherche à connaître le signe du membre de droite.

Supposons d'abord que $H_f(a)$ soit définie positive, et voyons qu'en fait il existe $m > 0$ tel que, pour tout $y \in \mathbb{R}^n$,

$$\langle y, H_f(a)y\rangle \geq m\|y\|_2^2.$$

C'est évident pour $y = 0$. Sinon, ça découle de ce qu'on peut prendre

$$m = \min\left\{\frac{\langle y, H_f(a)y\rangle}{\|y\|_2^2} : y \in \mathbb{R}^n, y \neq 0\right\} = \min\{\langle z, H_f(a)z\rangle, \|z\|_2 = 1\}.$$

L'application $z \mapsto \langle z, Az\rangle$ est continue et l'ensemble $\|z\|_2 = 1$ est compact, donc il existe z^* tel que $m = \langle z^*, H_f(a)z^*\rangle > 0$.

Cette propriété permet de conclure dans le cas où $H_f(a)$ est définie positive : prenant $\|\cdot\| = \|\cdot\|_2$, on aura

$$f(x) - T(x) \geq \big(m - |\varepsilon(x - a)|\big)\|x - a\|_2^2 > 0$$

pour $x - a$ assez proche de l'origine.

Le cas où $H_f(a)$ est définie négative est analogue.

Regardons enfin le cas où $H_f(a)$ est indéfinie. Fixons un voisinage V de l'origine. Par hypothèse, il existe $\tilde{x}, \tilde{y} \in \mathbb{R}^n$, dont on peut supposer que $\|\tilde{x}\|_2 = \|\tilde{y}\|_2 = 1$, tels que $\langle \tilde{x}, H_f(a)\tilde{x}\rangle = c > 0$ et $\langle \tilde{y}, H_f(a)\tilde{y}\rangle = -d < 0$. Prenons alors $\lambda > 0$ et posons $x = a + \lambda\tilde{x}$ et $y = a + \lambda\tilde{y}$. Pour λ assez petit, on a $x, y \in V$ et

$$\begin{aligned} f(x) - T(x) &= \frac{\lambda^2}{2}c + \lambda^2\varepsilon(\lambda\tilde{x}), \\ f(y) - T(y) &= -\frac{\lambda^2}{2}d + \lambda^2\varepsilon(\lambda\tilde{y}). \end{aligned}$$

En prenant λ suffisamment petit, la première ligne sera strictement positive, et la seconde strictement négative.

Exemples :

(1) La fonction $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^2 + x^4$ est partout deux fois différentiable et

$$H_f(x, y) = \begin{pmatrix} 2 + 12y^2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Cette matrice est partout définie positive, et dès lors en tout point, le graphe de f se situe au-dessus de celui du plan tangent.

(2) Considérons les fonctions $f, g : \mathbb{R}^2 \to \mathbb{R}$ définies par $f(x,y) = x^2 + y^4$ et $g(x,y) = x^2 - y^4$ pour tout $(x,y) \in \mathbb{R}^2$. En particulier

$$H_f(0,0) = H_g(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Cette matrice est semi-définie, et la proposition ne permet pas de décider si, dans un voisinage de $(0,0)$ le graphe de f, ou de g, est au dessus ou en-dessous du plan tangent $z = 0$.

Néanmoins, on peut ici trouver la réponse autrement. La fonction f est strictement positive en dehors de l'origine, et le graphe de f est donc strictement au-dessus du plan $z = 0$ en tout point autre que l'origine. Pour g, la situation est différente : g est strictement positive sur l'axe $y = 0$ et strictement négative sur l'axe $x = 0$, tous deux privés de l'origine. Dès lors, dans tout voisinage de l'origine, on trouve des points où g est au-dessus du plan $z = 0$, et d'autres où g est en-dessous. Dans ce cas on dit que $(0,0)$ est un point de selle, ou un point col.

(3) Soit $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto xy$. Cette fonction est différentiable sur $\mathbb{R}^2$ et

$$H_f(x,y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Cette matrice est indéfinie : en tout point le graphe de f traverse le plan tangent.

Proposition : Supposons que f soit deux fois différentiable sur un voisinage E de a. Si $H_f(x)$ est définie positive ou semi-définie positive sur E, alors $f(x) \geq T(x)$ sur un voisinage de a ; si $H_f(x)$ est définie négative ou semi-définie négative sur E, alors $f(x) \leq T(x)$ sur un voisinage de a

Preuve : Considérons le cas où $H_f(x)$ est (semi)-définie positive (l'autre cas est analogue). En prenant un développement au second ordre avec reste de Lagrange, on sait que, pour tout x dans un voisinage de a, on trouvera c dans ce voisinage tel que

$$f(x) - T(x) = f(x) - f(a) - \langle \nabla f(a), x - a \rangle = \frac{1}{2} \langle x - a, H_f(c)(x-a) \rangle \geq 0.$$

Remarque : les hypothèses de cette proposition sont plus fortes que celles de la précédente, mais elles permettent d'obtenir un cirtère valide pour les matrices semi-définies positives. Ainsi, on a une autre façon de déduire que $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto x^2 + y^4$ est en tout point au-dessus de son plan tangent en $(0,0)$.

7. Convexité

Nous définissons les ensembles convexes de $\mathbb{R}^n$ ainsi que la convexité des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$; nous en énonçons et démontrons les propriétés de base.

Ensembles convexes

On rappelle qu'étant donné deux points $a, b \in \mathbb{R}^n$, on définit le segment $[a, b]$ par

$$[a, b] = \{x \in \mathbb{R}^n : x = (1 - t)a + tb \text{ pour un certain } t \in [0, 1]\}.$$

Une partie $U \subset \mathbb{R}^n$ est convexe si, pour tous points $a, b \in U$, le segment $[a, b]$ est compris dans U.

Exemples :

(1) Etant donné une norme sur $\mathbb{R}^n$, une boule de rayon $r > 0$ et de centre a est une partie convexe de $\mathbb{R}^n$. Ceci résulte de l'inégalité triangulaire. Prenons $x, y \in \mathbb{R}^n$ tels que

$$\|x - a\| < r, \qquad \|y - a\| < r$$

(le cas de la boule fermée est analogue). Prenons $z = (1 - t)x + ty$ avec $t \in [0, 1]$. Alors

$$\begin{aligned}\|z - a\| &= \|(1 - t)(x - a) + t(y - a)\| \\ &\leq (1 - t)\|x - a\| + t\|y - a\| < (1 - t + t)r = r.\end{aligned}$$

(2) Un sous-espace affine U de $\mathbb{R}^n$ est convexe. On rappelle que $U \subset \mathbb{R}^n$ est un sous-espace affine s'il existe un sous-espace vectoriel $V \subset \mathbb{R}^n$ et un vecteur $v \in \mathbb{R}^n$, tel que $x \in U$ si et seulement si $x - v \in V$. Montrons que U est convexe. Soient $x, y \in U$, et soit $z = (1 - t)x + ty$ avec $t \in [0, 1]$. On a $x - v, y - v \in V$, et donc $z - v = (1 - t)(x - v) + t(y - v) \in V$ et donc $z \in U$.

(3) Une demi-droite dans $\mathbb{R}^n$ est convexe. Une demi-droite est un ensemble de la forme

$$D = \{x = (a + tu), t \geq 0\},$$

où $a \in \mathbb{R}^n$ et $0 \neq u \in \mathbb{R}^n$. Montrons que D est convexe. Si $x, y \in D$, alors $x = a + s_x u$ et $y = a + s_y u$, où $s_x, s_y \geq 0$. Prenons $z = (1 - t)x + ty$; alors $z = (1 - t + t)a + ((1 - t)s_x + ts_y)u$ et on a bien $(1 - t)s_x + ts_y \geq 0$.

(4) Un intersection de parties convexe est convexe. Cela suit directement de la définition.

(5) L'image d'un ensemble convexe par une application linéaire est convexe. Soit $U \subset \mathbb{R}^n$ un convexe et soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une applicaiton linéaire. Soient $x', y' \in f(U)$; ces points s'écrivent $x' = f(x)$ et $y' = f(y)$ pour certains $x, y \in U$. Prenons ensuite $z' = (1 - t)x' + ty'$ avec $t \in [0, 1]$, et voyons que $z' \in f(U)$. Comme U est convexe, $z := (1 - t)x + ty \in U$. De plus $f(z) \in f(U)$, et comme f est linéaire, $f(z) = z'$.

Fonctions convexes
Soit une application $f : U \to \mathbb{R}$, où U est une partie convexe de $\mathbb{R}^n$. On dit que f est convexe si, pour tout $a, b \in U$ et pour tout $t \in]0,1[$,

$$f((1-t)a+tb) \leq (1-t)f(a)+tf(b).$$

De plus, on dit que f est strictement convexe si l'inégalité large $\leq$ est remplacée par une inégalité stricte $<$ dans cette dernière expression. Enfin on dit que f est (strictement) concave si $-f$ est (strictement) convexe.

Exemples :

(1) Une fonction affine est convexe sur $\mathbb{R}^n$, mais non-strictement convexe (une fonction affine est donnée pour tout $x \in \mathbb{R}^n$ par $f(x) = \langle a, x\rangle + b$ pour un certain $a \in \mathbb{R}^n$ et $b \in \mathbb{R}$).

(2) Une norme est convexe sur $\mathbb{R}^n$ mais non strictement convexe. La convexité résulte de l'inégalité triangulaire : pour tout $x, y \in \mathbb{R}^n$ et pour tout $t \in]0,1[$,

$$\|(1-t)x+ty\| \leq (1-t)\|x\| + t\|y\|.$$

Que l'inégalité ne soit pas stricte résulte de ce que, si on prend $y = \lambda x$ pour un certain $\lambda \geq 0$, on a l'égalité dans l'expression précédente.

(3) Une forme quadratique $f : \mathbb{R}^n \to \mathbb{R}$ est donnée pour tout $x \in \mathbb{R}^n$ par $f(x) = \langle x, Ax\rangle$, où A est une matrice symétrique $n \times n$. La forme quadratique f est convexe sur $\mathbb{R}^n$ si et seulement si A est définie positive ou semi-définie positive. En effet, prenons $x, y \in \mathbb{R}^n$, puis $z = (1-t)x + ty$ et calculons

$$\begin{aligned}
&(1-t)f(x) + tf(y) - f(z) \\
&= (1-t)\langle x, Ax\rangle + t\langle y, A, y\rangle - \langle (1-t)x+ty, A((1-t)x+ty)\rangle \\
&= t(1-t)\big(\langle x, Ax\rangle - 2\langle x, Ay\rangle + \langle y, Ay\rangle\big) \\
&= t(1-t)\langle x-y, A(x-y)\rangle.
\end{aligned}$$

La conclusion suit du fait que $t(1-t) > 0$ pour tout $t \in]0,1[$.

Opérations sur les fonctions convexes
Nous donnons quelques propriétés de stabilité des fonctions convexes. Soit $U \subset \mathbb{R}^n$ un ensemble convexe.

(1) La somme de deux fonctions convexes sur un ensemble U est convexe.

(2) Si $f : U \to \mathbb{R}$ est convexe, et si $\lambda \geq 0$, alors λf est convexe sur U.

(3) Le maximum de deux fonctions convexes sur U est convexe.

(4) Si $f : U \to \mathbb{R}$ est convexe, et que $L : \mathbb{R}^m \to \mathbb{R}^n$ est linéaire, alors $f \circ L$ est convexe sur le convexe $V = \{x \in \mathbb{R}^m : L(x) \in U\}$.

(5) Soit $I \subset \mathbb{R}$ un intervalle. Si $g : I \subset \mathbb{R} \to \mathbb{R}$ est convexe et croissante, et si $f : U \to \mathbb{R}$ est convexe et telle que $f(U) \subset I$, alors $g \circ f$ est convexe sur U.

Démontrons la cinquième propriété. On veut montrer que, pour tous $x, y \in U$, et pour tout $t \in [0,1]$,

$$g(f((1-t)x+ty)) \leq (1-t)g(f(x)) + tg(f(y)).$$

Prenons donc $x, y \in U$ et $t \in [0,1]$. D'abord par la convexité de f, on a

$$f((1-t)x+ty) \leq (1-t)f(x) + tf(y).$$

Ensuite comme g est croissante, on a

$$g(f((1-t)x+ty)) \leq g((1-t)f(x)+tf(y)).$$

Enfin comme g est convexe,

$$g((1-t)f(x)+tf(y)) \leq (1-t)g(f(x)) + tg(f(y)).$$

Remarque sur la cinquième propriété. Nous verrons plus loin qu'une fonction deux fois différentiable est convexe si et seulement si la hessienne de f est définie ou semi-définie positive sur U. Supposons f et g deux fois différentiable et, à titre d'exercice, démontrons la propriété par calcul.

Considérons d'abord le cas particulier de $f : \mathbb{R} \to \mathbb{R}$. On a

$$(g \circ f)'' = g'' \circ f \cdot (f')^2 + g' \circ f \cdot f''$$

et on voit que les conditions sur g et f garantissent la positivité de $(g \circ f)''$.

Considérons ensuite le cas plus général $f : \mathbb{R}^n \to \mathbb{R}$. On calcule d'abord

$$\frac{\partial}{\partial x_i}(g \circ f) = g' \circ f \cdot \frac{\partial f}{\partial x_i},$$

puis

$$\frac{\partial^2}{\partial x_i \partial x_j}(g \circ f) = g'' \circ f \cdot \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} + g' \circ f \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

On voit alors que les conditions du théorèmes garantissent la positivité de la hessienne :

$$\langle x, H_{g\circ f}(a)x\rangle = g''(f(a))\langle \nabla f(a), x\rangle^2 + g'(f(a))\langle x, H_f(a)x\rangle.$$

Exemple : Soit la fonction

$$f : (\mathbb{R}_+^*)^2 \to \mathbb{R}, (x,y) \mapsto \frac{\mathrm{e}^{x^2+y^2-3}}{x^2 y}.$$

Cette fonction est convexe. Pour le voir, il suffit de prouver que $h = \ln f$ est convexe, car $f = \mathrm{e}^h$ et $z \to \mathrm{e}^z$ est convexe et croissante sur $\mathbb{R}$. Or

$$\ln f(x,y) = x^2 + y^2 - 3 - 2\ln x - \ln y,$$

et on voit que cette fonction est convexe comme somme de fonctions convexes.

Convexité et continuité

Commençons par énoncer et démontrer une version simple de l'inégalité de Jensen. Soit un ensemble convexe $U \subset \mathbb{R}^n$ et soit une fonction $f : U \to \mathbb{R}^n$.

Proposition : Supposons que f soit convexe sur U, et soient $m \geq 1$ et $p_1, \dots, p_m \geq 0$ tels que $p_1 + \dots + p_m = 1$. Alors

$$f(p_1 x_1 + \dots + p_m x_m) \leq p_1 f(x_1) + \dots + p_m f(x_m)$$

pour tout $x_1, \dots, x_m \in U$.

Preuve : C'est immédiat pour $m = 1$, c'est équivalent à la définition de convexité pour $m = 2$; le résultat s'obtient par récurrence sur m pour $m > 2$.

Théorème : Si U est ouvert et si f est convexe sur U, alors f est continue sur U.

Preuve : Pour simplifier, nous faisons la preuve dans le cas particulier $U = \mathbb{R}^n$ et $n = 2$. De plus, pour alléger les notations, nous démontrons la continuité en l'origine. On munit $\mathbb{R}^2$ de la norme $\| \cdot \|_\infty$. Nous procédons en deux étapes :

1ère étape : Nous montrons qu'il existe $C \in \mathbb{R}$ tel que, pour tout $r > 0$ dans un voisinage de 0, et pour tout x tel que $\|x\| \leq r$, on ait

$$f(x) - f(0) \leq Cr.$$

Pour voir ça, supposons x dans le 1er quadrant (les trois autres cas sont analogues). On a

$$x = p_1 0 + p_2 e_1 + p_3 e_2 \quad \text{où} \quad p_1 = 1 - x_1 - x_2, \quad p_2 = x_1, \quad p_3 = x_2.$$

(et où $0 = (0,0)$, $e_1 = (1,0)$ et $e_2 = (0,1)$). Pour r assez petit on a

$$p_1, p_2, p_3 > 0 \qquad \text{et} \qquad p_1 + p_2 + p_3 = 1.$$

Dès lors, par convexité de f (inégalité de Jensen), on trouve

$$f(x) \leq p_1 f(0) + p_2 f(e_1) + p_3 f(e_2)$$

et donc

$$f(x) - f(0) \leq |p_1 - 1||f(0)| + p_2 |f(e_1)| + p_3 |f(e_2)|.$$

L'affirmation suit alors de ce que

$$|p_1 - 1| \leq 2r, \quad p_2 \leq r, \quad p_3 \leq r.$$

2e étape : Nous montrons qu'il existe $C' \in \mathbb{R}$ tel que, pour tout $r > 0$ dans un voisinage de 0, et pour tout x tel que $\|x\| \leq r$, on ait

$$f(0) - f(x) \leq C' r.$$

Pour voir ça, on écrit cette fois

$$0 = p_1 x + p_2(-e_1) + p_3(-e_2) \quad \text{avec} \quad p_1, p_2, p_3 > 0 \quad \text{et} \quad p_1 + p_2 + p_3 = 1 \;;$$

cette décomposition est rendue possible en prenant

$$p_1 = \frac{1}{1 + x_1 + x_2}, \quad p_2 = \frac{x_1}{1 + x_1 + x_2}, \quad p_3 = \frac{x_2}{1 + x_1 + x_2}$$

(on suppose toujours x dans le 1er quadrant). Par convexité de f, on trouve

$$f(0) - f(x) \leq |p_1 - 1||f(x)| + p_2|f(-e_1)| + p_3|f(-e_2)|.$$

Prenant r dans un voisinage de l'origine, observant qu'encore une fois

$$|p_1 - 1| \leq 2r, \quad p_2 \leq r, \quad p_3 \leq r$$

et se servant de la première étape pour borner $|f(x)|$, on obtient l'affirmation souhaitée.

La continuité en l'origine est une conséquence directe des deux affiramtions ci-dessus.

Remarque : La proposition est fausse si U n'est pas ouvert. En effet, prenons par exemple $U = [0,1] \subset \mathbb{R}$ et considérons $f : U \to \mathbb{R}$ telle que $f(x) = 0$ si $x \in]0,1[$ et $f(x) = 1$ si $x \in \{0,1\}$.

Remarque : En général, il n'est pas vrai qu'une fonction convexe sur un ouvert U est différentiable sur U. C'est le cas de la fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto |x|$, qui est convexe (car c'est une norme) mais pas différenatiable en 0 (car c'est une norme).

Convexité et différentiabilité au premier ordre

Soit $U \subset \mathbb{R}^n$ un ouvert convexe, et soit une application différentiable $f : U \to \mathbb{R}$. Etant donné $a \in U$, on note T_a le plan tangent à f au point a : pour tout $x \in \mathbb{R}^n$,

$$T_a(x) = f(a) + \langle \nabla f(a), (x - a) \rangle.$$

Théorème : La fonction f est convexe si et seulement si, pour tout $a \in U$ et pour tout $x \in U$, $f(x) \geq T_a(x)$.

Preuve : Supposons d'abord f convexe sur U, et montrons que pour tout $a, b \in U$, on a $f(b) - f(a) \geq \langle \nabla f(a), b - a \rangle$. Il existe $r > 0$ telle que la fonction

$$\phi :]-r, 1+r[, t \mapsto f((1-t)a + tb)$$

soit différentiable. Considérons aussi la fonction affine

$$\Phi :]-r, 1+r[, t \mapsto (1-t)f(a) + tf(b).$$

Par convexité, $\Phi(t) \geq \phi(t)$ pour tout $t \in]0,1[$. Un développement au premier ordre autour de $t = 0$ donne

$$\phi(t) = f(a) + t\langle \nabla f(a), (b - a) \rangle + |t|\varepsilon(t)$$

où ε est continue et s'annule en 0. Dès lors pour $t > 0$ on a

$$t(f(b) - f(a)) \geq t\langle \nabla f(a), (b - a) \rangle + |t|\varepsilon(t).$$

Pour que ceci soit vrai dans la limite $t \to 0$, il faut que

$$f(b) - f(a) \geq \langle \nabla f(a), (b - a) \rangle,$$

ce qui est bien ce qu'on souhaitait obtenir.

Supposons ensuite $f(x) \geq T_z(x)$ pour tout $x, z \in U$. Donc, pour tout $a, b \in U$ et pour tout $t \in]0,1[$, on obtient en prenant $z = (1-t)a + tb$,

$$f(x) \geq f((1-t)a + tb) + \langle \nabla f((1-t)a + tb), x - (1-t)a - tb \rangle.$$

Prenant en particulier $x = a$ et $x = b$, on trouve

$$f(a) \geq f((1-t)a + tb) - t\langle \nabla f((1-t)a + tb), b - a \rangle,$$
$$f(b) \geq f((1-t)a + tb) + (1-t)\langle \nabla f((1-t)a + tb), b - a \rangle.$$

Multipliant la première ligne par $1-t$ et la seconde par t, on somme pour obtenir

$$(1-t)f(a) + tf(b) \geq f((1-t)a + tb).$$

Donc f est convexe.

Convexité et différentiabilité au second ordre

Soit $U \subset \mathbb{R}^n$ un ouvert convexe, et soit $f : U \to \mathbb{R}$ une application deux fois différentiable. On note $H_f(a)$ la matrice hessienne de f en un point $a \in U$. Le théorème suivant n'est qu'une réexpression dans le langage de la convexité des deux dernières propositions du chapitre précédent (paragraphe intitulé *Position du graphe de f par rapport à l'approximation affine*).

Théorème : f est convexe sur U si et seulement si, pour tout $a \in U$ la hessienne $H_f(a)$ est définie ou semi-définie positive. De plus, si pour tout $a \in U$, la matrice hessienne $H_f(a)$ est définie positive, alors f est strictement convexe sur U

Preuve : Supposons d'abord f convexe sur U. Comme f est deux fois différentiable sur U, on a en tout $a \in U$, et pour x dans un voisinage de a :

$$f(x) = f(a) + \langle \nabla f(a), (x-a) \rangle + \frac{1}{2}\langle x - a, H_f(a)(x-a) \rangle + \|x - a\|^2 \varepsilon(x - a)$$

où ε est continue e s'annule en 0. Par convexité de f, on doit avoir

$$\frac{1}{2}\langle x - a, H_f(a)(x-a) \rangle + \|x - a\|^2 \varepsilon(x - a) \geq 0$$

pour x dans un voisinage de a. Procédant comme au chaptire précédent, on montre alors que $H_f(a)$ doit être définie ou semi-défine positive.

Supposons ensuite que $H_f(a)$ soit définie ou semi-définie positive pour tout $a \in U$. Alors, pour tout $x \in U$, il existe $c \in]a, x[$ tel que

$$f(x) - f(a) - \langle \nabla f(a), x - a \rangle = \frac{1}{2}\langle (x-a), H_f(c)(x-a) \rangle \geq 0.$$

On conclut que f est convexe. De plus, si $H_f(a)$ est définie positive, on conclut que f est strictement convexe.

Remarque : Il se peut que f soit strictement convexe sur U mais que $H_f(a)$ ne soit pas définie positive pour tout $a \in U$. En effet, $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^4$ est strictement convexe, mais $f''(0) = 0$.

8. Extrema libres

Dans ce chapitre, nous étudions principalement les extrema locaux d'une fonction définie sur un ouvert $U \subset \mathbb{R}^n$. Dans le cadre de ce cours, c'est l'application principale de la théorie développée jusqu'ici.

Définitions

Soit $U \subset \mathbb{R}^n$, et soit une fonction $f : U \to \mathbb{R}$.

On dit que f atteint un maximum global en $a \in U$ si $f(x) \leq f(a)$ pour tout $x \in U$, et un minimum global en a si $f(x) \geq f(a)$ pour tout $x \in U$.

On dit que f atteint un maximum local en $a \in U$ s'il existe un voisinage V de a tel que $f(x) \leq f(a)$ pour tout $x \in U \cap V$, et un minimum local en $a \in U$ s'il existe un voisinage V de a tel que $f(x) \geq f(a)$ pour tout $x \in U \cap V$.

On parle d'extremum pour désigner un minimum ou un maximum.

Points critiques

Soit un ouvert $U \subset \mathbb{R}^n$ et soit une fonction $f : U \to \mathbb{R}$ différentiable sur U. Soit $a \in U$.

Théorème : Si f admet un extremum local en a, alors $\nabla f(a) = 0$

Preuve : Supposons que f admette un maximum local en a (le cas du minimum est analogue). Comme f est différentiable en a, on a pour tout h dans un voisinage de 0,

$$0 \geq f(a+h) - f(a) = \langle \nabla f(a), h \rangle + \|h\|_2 \varepsilon(h)$$

où ε est continue et s'annule en 0. Pour un réel $t > 0$ dans un voisinage de 0, on pose $h = t\nabla f(a)$, et on a donc

$$\begin{aligned}\langle \nabla f(a), h \rangle + \|h\|_2 \varepsilon(h) &= t\|\nabla f(a)\|_2^2 + t\|\nabla f(a)\|_2 \varepsilon(t\nabla f(a)) \\ &= t\|\nabla f(a)\|_2 \big(\|\nabla f(a)\|_2 - \varepsilon(t\nabla f(a))\big) \leq 0.\end{aligned}$$

Comme cette inégalité doit être valide pour tout $t > 0$ assez petit, on a donc $\|\nabla f(a)\|_2 = 0$ et donc $\nabla f(a) = 0$.

Définition : Si $\nabla f(a) = 0$, on dit que a est un point critique de f.

Remarque : En général, il n'est pas vrai que f atteint un extremum local en un point critique. Exemples :

(1) $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ admet un point critique en 0 mais n'admet pas d'extremum.

(2) $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 - y^4$ admet un point critique en $(0, 0)$, mais n'admet pas d'extremum en ce point. En effet, $x \mapsto f(x, 0)$ admet un minimum global en 0 alors que $y \mapsto f(0, y)$ admet un maximum global en 0 ; on dit que $(0, 0)$ est un point selle ou point col. Par contre la fonction $g : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^4$ admet un minimum global en $(0, 0)$, puisque g est positive et $g(0, 0) = 0$.

Extrema locaux
Nous nous servons à présent des résultats obtenus au paragraphe *Position du graphe de f par rapport à l'approximation affine* du chapitre 6, pour obtenir des critères permettant de décider si un point critique donne lieu à un extremum local. Comme là, nous obtenons un critère purement ponctuel et un critère local.

Nous supposons que $U \subset \mathbb{R}^n$ est un ouvert, que $f : U \to \mathbb{R}$ est deux fois différentiable sur U, et que $a \in U$ est un point critique de f.

Proposition : Si $H_f(a)$ est définie positive, alors f atteint un minimum local en a ; si $H_f(a)$ est définie négative, alors f atteint un maximum local en a ; si $H_f(a)$ est indéfinie, alors f n'atteint pas d'extremum en a.

Proposition : Si $H_f(x)$ est définie ou semi-définie positive pour tout x dans un voisinage de a, alors f atteint un minimum local en a ; si $H_f(x)$ est définie ou semi-définie négative pour tout x dans un voisinage de a, alors f atteint un maximum local en a.

Preuve : Ces propositions sont des corollaires directs des deux propositions du paragraphe *Position du graphe de f par rapport à l'approximation affine* du chapitre 6, en notant que $T_a(x) = f(a)$ au point critique a puisque $\nabla f(a) = 0$.

Extrema globaux
La théorie des extrema libres ne permet pas généralement de déterminer si un extremum est global. On a deux cas particuliers intéressants :

(1) Le cas où le domaine de f est compact (ce cas sort en fait du cadre de la théorie des extrema libres qui considère le cas d'un domaine ouvert). Par le théorème de Weierstrass, on sait alors que si f est continue, elle atteint un minimum global et un maximum global sur le domaine.

(2) Le cas où la fonction f est convexe (ou concave) ; dans ce cas aussi l'existence d'un minimum (ou d'un maxium) global est garantie. Voir le théorème ci-dessous.

Soit un ouvert convexe $U \subset \mathbb{R}^n$, soit une fonction différentiable $f : U \to \mathbb{R}$, et soit $a \in U$ un point critique de f.

Proposition : Si f est convexe sur U, alors f atteint un minimum global en a ; si f est concave sur U, alors f atteint un maximum global en a.

Preuve : Considérons le cas où f est convexe ; l'autre est analogue. Par l'avant-dernier théorème du chapitre précédent, on a $f(x) \geq T_a(x) = f(a)$ pour tout $x \in U$.

Exemple 1 (optimisation sur un compact) : Cherchons les extrema de $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 - y^2$ sur le domaine compact $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. Sur $\mathrm{int}(D)$, il y a un unique point critique en $(0, 0)$, mais c'est un point selle ; il n'y a donc pas d'extremum sur $\mathrm{int}(D)$.

Il y a donc des extrema sur le cercle $\partial D = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Paramétrisons ce cercle par

$$x = \cos\theta, \qquad y = \sin\theta \qquad (\theta \in [0, 2\pi[)$$

et définissons

$$F : [0, 2\pi[\to \mathbb{R}, \theta \mapsto f(\cos\theta, \sin\theta) = \cos^2\theta - \sin^2\theta$$

On a $-1 \leq F \leq 1$, et on trouve que F admet un maximum global en $\theta = 0$ et $\theta = \pi$, et un minimum global en $\theta = \pi/2$ et $\theta = 3\pi/2$. Dès lors f admet un maximum global sur U en $(1,0)$ et en $(-1,0)$, et un minimum global en $(0,1)$ et en $(0,-1)$.

Exemple 2 (droite des moindres carrés) : Considérons m points $(x_i, y_i)_{1 \leq i \leq m} \in \mathbb{R}^2$. En général, pour $m > 2$, on ne peut pas trouver de coefficients $a, b \in \mathbb{R}$ tels que

$$y_i = ax_i + b \qquad \text{pour tout } 1 \leq i \leq m.$$

Par contre, on peut trouver $(a,b) \in \mathbb{R}^2$ qui minimise la somme du carré des erreurs $\varepsilon_i = y_i - ax_i - b$, c'est-à-dire qui minimise la fonction convexe $\mathcal{E} : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$\mathcal{E}(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2.$$

Un minimum global est atteint en tout point critique. De plus, il existe un unique point critique (sauf si $x_1 = \cdots = x_n$) : (a,b) est un point critique si

$$\frac{\partial \mathcal{E}}{\partial a}(a,b) = -2\sum_{i=1}^{n} (y_i - ax_i - b)x_i = 0,$$

$$\frac{\partial \mathcal{E}}{\partial b}(a,b) = -2\sum_{i=1}^{n} (y_i - ax_i - b) = 0.$$

On résout ce système :

$$a = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2},$$

$$b = \overline{y} - a\overline{x},$$

où on a noté

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \text{et} \qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

9. Extrema liés : introduction aux multiplicateurs de Lagrange

Au chapitre précédent, nous avons principalement considéré le problème de l'optimisation locale d'une fonction sur un ouvert $U \subset \mathbb{R}^n$. Nous voudrions maintenant étudier des cas plus généraux, incluant en particulier certains ensembles fermés. Ce chapitre constitute une introduction à la méthode des multiplicateurs de Lagrange, qui ne sera vue en détail qu'en L3.

Optimisation sous contrainte d'égalité
Soit un ouvert $U \subset \mathbb{R}^n$ ($n \geq 2$) et soit une application $f : U \to \mathbb{R}$. Soient aussi m applications $g_1, \ldots, g_m : U \to \mathbb{R}$, où $1 \leq m < n$. On cherche à optimiser f sous la contrainte d'égalité $g_1 = \cdots = g_m = 0$, c'est-à-dire à optimiser f sur l'ensemble

$$\mathcal{K} = \{x \in U : g_1(x) = 0, \ldots, g_m(x) = 0\}.$$

Plus précisément, on dit que a est un minimum local de f sous la contrainte $g_1 = \cdots = g_m = 0$, si $a \in \mathcal{K}$ et s'il existe un voisinage $V \subset U$ de a tel que, pour tout $x \in V \cap \mathcal{K}$, on ait $f(x) \geq f(a)$. On définit maximum local et extrema globaux de façon analogue.

Exemples :

(1) Soit une fonction $f : \mathbb{R}^2 \to \mathbb{R}$, et soit $g(x, y) = x^2 + y^2 - 1$. Optimiser f sous la contrainte $g(x, y) = 0$ revient à chercher les extrema de f restreinte au cercle d'équation $x^2 + y^2 = 1$.

(2) Soit une fonction $f : \mathbb{R}^3 \to \mathbb{R}$, et soient les fonctions

$$g_1(x, y, z) = x - y, \qquad g_2(x, y, z) = y - z.$$

Optimiser f sous la contrainte $g_1(x, y, z) = g_2(x, y, z) = 0$ revient à optimiser la fonction f restreinte à la droite de $\mathbb{R}^3$ d'équation $x = y = z$.

Remarque : L'optimisation sous contrainte d'égalité permet aussi de traiter certains cas d'optimisation sous contrainte d'inégalité. Ainsi au chapitre précédent, nous avons optimisé la fonction $f(x, y) = x^2 - y^2$ sous la contrainte d'inégalité $x^2 + y^2 \leq 1$ en cherchant d'abord déventuels extrema sur la boule ouverte définie par l'inégalité stricte $x^2 + y^2 < 1$, puis sur le cercle d'équation $x^2 + y^2 = 1$.

Première méthode de résolution : substitution
Quand la contrainte prend une forme relativement simple, on peut se ramener à un problème d'optimisation sur un ouvert de $\mathbb{R}^{n-m}$ en substituant la contrainte dans l'expression de f. Donnons deux exemples :

(1) Cherchons à optimiser la fonction $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + y$ sous la contrainte $g(x, y) = y - x^2 = 0$. On restreint donc la fonction f aux points de $\mathbb{R}^2$ qui satisfont la contrainte $y = x^2$. Posons

$$F(x) = f(x, x^2) = x + x^2.$$

Optimiser f sous la contrainte $g=0$ se ramène à optimiser F sur $\mathbb{R}$. On trouve que F admet un minimum global en $x=-1/2$. Donc f admet un minimum global sous la contrainte $g=0$ en $(-1/2,1/4)$.

(2) Cherchons à optimiser la fonction $f:\mathbb{R}^2\to\mathbb{R},(x,y)\mapsto x+y$ sous la contrainte $g(x,y)=x^2+y^2-1=0$. La contrainte représente un cercle que l'on peut paramétriser par

$$x=\cos\theta, \qquad y=\sin\theta$$

pour $\theta\in\mathbb{R}$. Posons

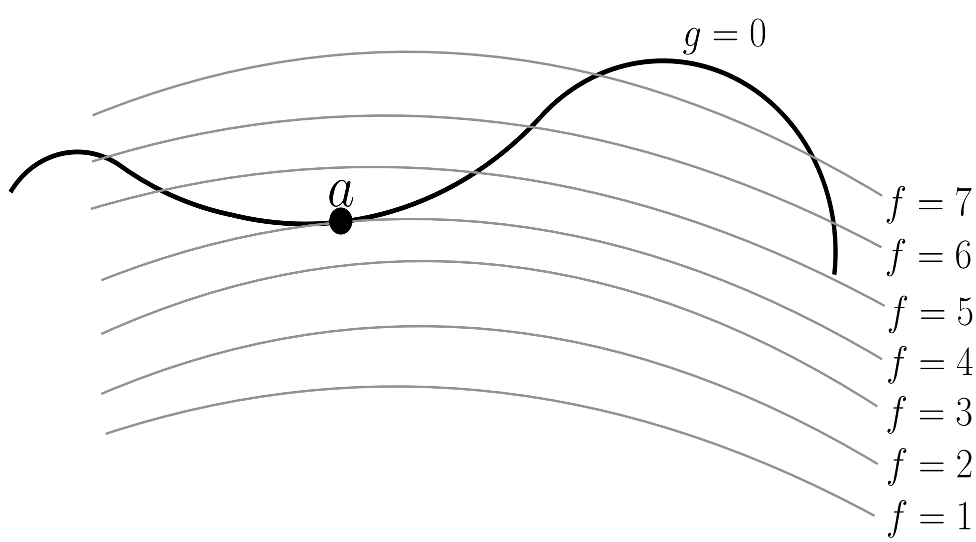
$$F(\theta)=f(\cos\theta,\sin\theta)=\cos\theta+\sin\theta.$$

Optimiser f sous la contrainte $g=0$ revient à optimiser F sur $\mathbb{R}$. On trouve que F admet un maximum global en tous les points de la forme $\theta=\pi/4+2k\pi$ $(k\in\mathbb{Z})$ et un minimum global en tous les points de la forme $\theta=5\pi/4+2k\pi$ $(k\in\mathbb{Z})$. On conclut que f admet un maximum global sous la contrainte $g=0$ en $(1/\sqrt{2},1/\sqrt{2})$ et un minimum global sous la contrainte $g=0$ en $(-1/\sqrt{2},-1/\sqrt{2})$.

Seconde méthode de résolution : multiplicateurs de Lagrange

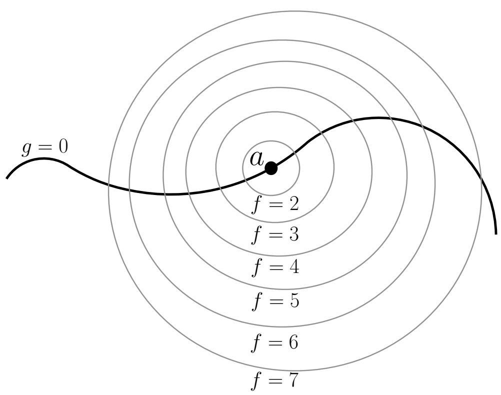
Considérons le cas où $f:\mathbb{R}^2\to\mathbb{R}$ et où on a une seule contrainte $g(x,y)=0$. Nous supposons que les fonctions f et g sont de classe $\mathcal{C}^1$. La méthode des multiplicateurs de Lagrange se base sur l'observation géométrique que si f atteint un extremum sous la contrainte $g=0$ en a (en particulier $g(a)=0$), alors le gradient de f en a est proportionnel au gradient de g en a, c'est-à-dire $\nabla f(a)=\lambda\nabla g(a)$ pour un certain $\lambda\in\mathbb{R}$ que l'on appelle le multiplicateur de Lagrange associé à cet extremum. Ceci est vrai sauf si $\nabla g(a)=0$, cas qui correspond à une dégénérescence qu'on rencotre peu en pratique.

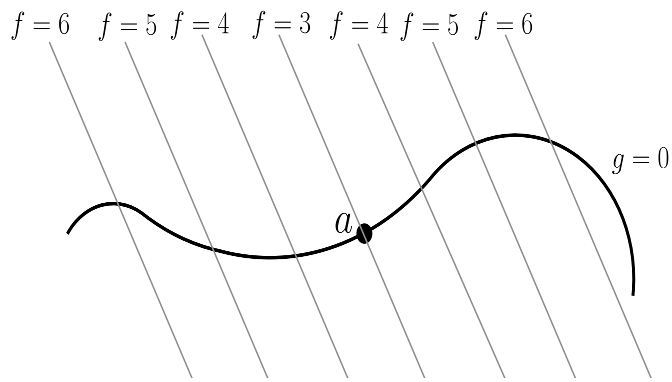
Voyons ceci sur des figures. Le cas le plus courant est représenté ci-dessous :



La contrainte $g = 0$ est représentée en noir et les courbes de niveau de f en gris (de la valeur 1 à la valeur 7). On voit qu'un minimum local de f sous la contrainte $g = 0$ est atteint au point a où la courbe définie par la contrainte $g = 0$ intersecte tangentiellement une courbe de niveau de f. En effet si on longe la courbe d'équation $g(x, y) = 0$ en partant de la gauche, on voit que la valeur de f le long de la courbe diminue d'abord, passant de 6 à 4 au point a, avant d'augmenter à nouveau au fur et à mesure qu'on s'éloigne de a. Or on a vu au chapitre 4 (au paragraphe *Propiétés géométriques du gradient*) que le gradient d'une fonction est perpendiculaire à la tangente de ses courbes de niveau. Donc, en particulier au point a, $\nabla f(a)$ est perpendiculaire à la tangente à la courbe de niveau 4 de f, et $\nabla g(a)$ est perpendiculaire à la tangente à la courbe de niveau 0 de g. Que ces deux courbes de niveau s'intersectent tangentiellement implique que $\nabla f(a)$ et $\nabla g(a)$ soient collinéaires.

Considérons ensuite des cas moins courants, comme les deux cas ci-dessous :





Dans ces deux cas-là, f atteint un minimum local au point a sans contrainte, c'est-à-dire que la courbe déquation $g(x, y) = 0$ passe justement sur un extremum de f. Il en résulte que $\nabla f(a) = 0$, et que $\nabla g(a)$ est trivialement colinéaire à $\nabla f(a)$.

Enfin, il peut arriver que f admette un extremum local en un point a sous la contrainte $g = 0$ et que $\nabla g(a) = 0$. Dans ce cas, les gradients de f et g sont bien collinéaires mais on ne peut en général plus écrire $\nabla f(a) = \lambda \nabla g(a)$ pour un certain $\lambda \in \mathbb{R}$. Ce cas est généralement traité séparément. Notons que ce cas provient parfois d'une mauvaise représentation de la contrainte. Ainsi la contrainte $g(x, y) = 0$ où $g(x, y) = x^2 + y^2 - 1$ et la contrainte $\tilde{g}(x, y) = 0$ où $\tilde{g}(x, y) = (g(x, y))^2$ représentent la même courbe du plan (le cercle de rayon 1 centré en 0). Mais $\nabla g(x, y) \neq (0, 0)$ pour tout point du cercle alors que $\nabla \tilde{g}(x, y) = (0, 0)$ pour tout point du cercle. Donc dans ce cas, ce problème est artificiel.

En fin de compte on voit que f atteint ses extrema sous la contrainte $g = 0$ en des points a qui satisfont soit $\nabla f(a) = \lambda \nabla g(a)$ pour un certain $\lambda \in \mathbb{R}$, soit $\nabla g(a) = 0$ (ce qui ne veut pas dire qu'un extremum est atteint en tout point vérifiant un de ces deux critères, de la même façon qu'on peut avoir un point critique en optimisation libre qui ne correspond à aucun extremum).

Remarque : Il peut être intérssant de reprendre les exemples du paragraphe précédent (résolus par la méthode de substitution), et de vérifier par un dessin que les gradients de f et g sont bien collinéaires aux extrema.

Théorème des multiplicateurs de Lagrange

On peut rendre rigoureux le raisonnement géométrique du paragraphe précédent. Nous considérerons le cas d'une seule contrainte. Soit un ouvert $U \subset \mathbb{R}^n$ et soient deux fonctions $f, g : U \to \mathbb{R}$ de classe $\mathcal{C}^1$.

Théorème : Soit $a \in U$. Si la fonction f admet un extremum sous la contrainte $g = 0$ en a, alors

(1) $g(a) = 0$ *(a est sur la contrainte),*

(2) *Soit $\nabla g(a) = 0$, soit il existe $\lambda \in \mathbb{R}$ tel que $\nabla f(a) = \lambda \nabla g(a)$.*

Remarque sur la terminologie :

(1) Si un point a satisfait aux points (1) et (2) du théorème, on dit que a est un point critique de f sous la contrainte $g = 0$.

(2) Si a est un point critique tel que $\nabla g(a) = 0$, on dit que c'est un point critique de seconde espèce.

(3) Si a est un point critique tel que $\nabla f(a) = \lambda \nabla g(a)$ pour un certain $\lambda \in \mathbb{R}$, on dit que c'est un point critique de première espèce. Dans ce cas le nombre λ est appelé multiplicateur de Lagrange associé à a.

Remarque : Comme en optimisation libre, on peut avoir un point critique qui ne corresponde à aucun extremum. Par exemple, la fonction f définie sur $\mathbb{R}^2$ par $f(x, y) = x^3 + y^3$ admet un point critique en $(0, 0)$ sous la contrainte $y = 0$, mais n'admet pas d'extremum sous cette contrainte.

Preuve partielle du théorème : Pour simplifier les notations, nous allons considérer le cas $U = \mathbb{R}^n$. De plus, nous considérons le cas particulier où g est telle que, pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$g(x_1, \ldots, x_n) = x_n - r(x_1, \ldots, x_{n-1})$$

pour une certaine fonction r de classe $\mathcal{C}^1$ sur $\mathbb{R}^{n-1}$. C'est-à-dire que nous considérons le cas où la contrainte est en fait le graphe d'une fonction des $n-1$ premières variables (bien sûr l'ordre des variables n'a pas d'importance). Cette forme particulière exclut la présence de points critiques de seconde espèce. Le théorème des fonctions implicites permettrait de montrer que cette situation est en fait générale (sauf s'il y a un point critique de seconde espèce), mais nous n'abordons pas ce point ici.

Dans ce cas particulier on peut en fait résoudre le problème d'optimisation par substitution. Définissons une application F sur $\mathbb{R}^{n-1}$ par

$$F(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, r(x_1, \ldots, x_{n-1})).$$

Le problème d'optimisation de f sous la contrainte $g = 0$ est équivalent au problème d'optimisation de F sur $\mathbb{R}^{n-1}$. Donc, si $a = (a_1, \dots, a_n)$ est un extremum de f sous $g = 0$, alors $\tilde{a} = (a_1, \dots, a_{n-1})$ est un point critique de F, et donc

$$0 = \big(\nabla F(\tilde{a})\big)_k = \frac{\partial f}{\partial x_k}(a) + \frac{\partial f}{\partial x_n}(a)\frac{\partial r}{\partial x_k}(\tilde{a})$$

pour tout $1 \leq k \leq n$.

Calculons ensuite séparément $\nabla f(a)$ et $\nabla g(a)$. D'abord, utilisant la relation qu'on vient de trouver,

$$\begin{aligned} \big(\nabla f(a)\big)_k &= \frac{\partial f}{\partial x_k}(a) = -\frac{\partial f}{\partial x_n}(a)\frac{\partial r}{\partial x_k}(\tilde{a}) \quad 1 \leq k \leq n-1, \\ \big(\nabla f(a)\big)_n &= \frac{\partial f}{\partial x_n}(a). \end{aligned}$$

Ensuite, d'après la forme particulière de g,

$$\begin{aligned} \big(\nabla g(a)\big)_k &= \frac{\partial g}{\partial x_k}(a) = -\frac{\partial r}{\partial x_k}(\tilde{a}) \quad 1 \leq k \leq n-1, \\ \big(\nabla g(a)\big)_n &= \frac{\partial g}{\partial x_n}(a) = 1. \end{aligned}$$

Donc

$$\nabla f(a) = \frac{\partial f}{\partial x_n}(a)\nabla g(a),$$

ce qui donne le résultat avec $\lambda = \partial f/\partial x_n(a)$.

Lagrangien

Le résultat précédent ne permet pas de dire si un extremum est effectivement réalisé au point critique. Comme au chapitre précédent, on peut s'appuyer soit sur la compacité soit sur la convexité pour obtenir des informations complémentaires. Quand la contrainte définit un ensemble compact, on sait qu'au moins un point critique correspond à un maximum global et un point critique à un minimum globlal (nous verrons un exemple au paragraphe suivant). Pour exploiter la convexité, nous reformulons le problème de façon un peu différente.

Soit a un point critique de première espèce de f sous la contrainte $g = 0$, et soit λ le multiplicateur de Lagrange associé. On définit le Lagrangien $\mathcal{L}_\lambda$ sur tout l'ouvert U par

$$\mathcal{L}_\lambda(x) = f(x) - \lambda g(x)$$

pour tout $x \in U$. La proposition suivante montre qu'à l'aide de la notion de Lagrangien, on ramène le problème d'optimisation sous contrainte à un problème d'optimisation libre (mais attention, la valeur de λ dépend du point critique).

Proposition : Le point a est un point critique de $\mathcal{L}_\lambda$. De plus, si $\mathcal{L}_\lambda$ admet un extremum local (ou global) en a, alors f admet un extremum local (ou global) en a sous la contrainte $g = 0$.

Preuve : Comme a est un point critique de première espèce sous la contrainte $g = 0$, on a $\nabla f(a) = \lambda \nabla g(a)$, ce qui s'écrit aussi $\nabla \mathcal{L}_\lambda(a) = 0$, ce qui montre que a est un point critique de $\mathcal{L}_\lambda$.

Ensuite supposons que a soit un minimum global de $\mathcal{L}_\lambda$ (les autres cas sont analogues). Donc, pour tout $x \in U$,

$$\mathcal{L}_\lambda(a) \leq \mathcal{L}_\lambda(x),$$

ou donc

$$f(a) - \lambda g(a) \leq f(x) - \lambda g(x),$$

avec en fait $g(a) = 0$ car a est sur la contrainte. Sur la contrainte, $g(x) = 0$ et donc

$$f(a) \leq f(x)$$

pour tout $x \in U$ tel que $g(x) = 0$.

Exemple : Soit f définie sur $(\mathbb{R}_*^+)^n$ par

$$f(x_1, \ldots, x_n) = x_1 \ln x_1 + \cdots + x_n \ln x_n,$$

et soit la contrainte $x_1 + \cdots + x_n = 1$, c'est-à-dire $g = 0$ pour g définie sur $(\mathbb{R}_*^+)^n$ par $g(x_1, \ldots, x_n) = x_1 + \cdots + x_n - 1$.

Trouvons les points critiques. Il n'y a pas de point critique de seconde espèce car $\nabla g(x) = (1, \ldots, 1) \neq (0, \ldots, 0)$ pour tout $x \in (\mathbb{R}_*^+)^n$. Un point critique de première espèce a satisfait $\nabla f(a) = \lambda \nabla g(a)$ pour un certain λ. Donc

$$\ln a_k + 1 = \lambda \qquad \forall 1 \leq k \leq n.$$

On déduit que $a_1 = \cdots = a_n$ et comme a doit appartenir à la contrainte que $a_k = 1/n$, et donc que $\lambda = \ln 1/n + 1 = 1 - \ln n$.

Pour déterminer la nature de l'unique point critique a, écrivons le Lagrangien associé :

$$\mathcal{L}_\lambda(x) = x_1 \ln x_1 + \cdots + x_n \ln x_n - (1 - \ln n)(x_1 + \cdots + x_n - 1).$$

C'est une fonction convexe sur $(\mathbb{R}_*^+)^n$ comme somme de fonctions convexe. Donc $\mathcal{L}_\lambda$ admet un unique minimum global en $(1/n, \ldots, 1/n)$, et donc f admet un unique minimum global sous la contrainte $g = 0$ en $(1/n, \ldots, 1/n)$.

Application : diagonalisation des matrices symétriques

Nous allons utiliser la méthode des multiplicateurs de Lagrange pour donner une preuve du théorème de diagonalisation des matrices réelles symétriques.

Commençons par résoudre un problème d'optimistion sur un compact. Soit une matrice symétrique $A \in \mathbb{R}^{n \times n}$. On cherche à optimiser la fonction $f : \mathbb{R}^n \to \mathbb{R}, x \mapsto \langle x, Ax \rangle$ sous la contrainte $\|x\|_2 = 1$, c'est-à-dire sous la contrainte $g(x) = 0$ où $g : \mathbb{R}^n \to \mathbb{R}, x \mapsto \|x\|_2^2 - 1$.

Calculons les points critiques. Cherchons d'abord les éventuels points critiques de seconde espèce. Comme, pour tout $x \in \mathbb{R}^n$, $\nabla g(x) = 2x$, on aura $\nabla g(x) = 0$ si et seulement si $x = 0$. Or 0 n'est pas sur la contrainte ($g(0) \neq 0$). Il n'y a donc pas de point critique de seconde espèce. Cherchons ensuite les points de première espèce. Comme, pour tout

$x \in \mathbb{R}^n$, $\nabla f(x) = 2Ax$, le point a sera un point critique de première espèce s'il existe $\lambda \in \mathbb{R}$ tel que

$$Aa = \lambda a.$$

Comme la contrainte définit un ensemble compact et que f est continue, on sait qu'elle atteint un minimum et un maximum global. Comme f est de classe $\mathcal{C}^1$, les extrema sont atteints en des points critiques. On a donc montré l'existence de $\lambda_-, \lambda_+ \in \mathbb{R}$ et $a_-, a_+ \in \mathbb{R}^n$, avec $\|a_-\|_2 = \|a_+\|_2 = 1$, tels que $Aa_- = \lambda_- a_-$ et $Aa_+ = \lambda_+ a_+$, c'est-à-dire de deux couples de valeur propre et vecteur propre. De plus, si A n'est pas un multiple de l'indentité, $a_- \neq a_+$ et $\lambda_- \neq \lambda_+$ (car alors f n'est pas constante sur la contrainte). Si par contre $A = r\mathrm{Id}$, alors $\lambda_- = \lambda_+ = r$ et les vecteurs propres sont quelconques.

A l'aide de ce résultat, montrons que A est diagonalisable, c'est à dire qu'il existe une base orthonormée $\{v_1, \ldots, v_n\}$ et des nombres $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que

$$A = \lambda_1 P_{v_1} + \cdots + \lambda_n P_{v_n}$$

où P_{v_i} est le projecteur orthogonal sur v_i :

$$P_{v_i} : x \mapsto \langle x, v_i \rangle v_i$$

pour tout $1 \leq i \leq n$. On vérifie tout de suite que $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A et que $v_1, \ldots, v_n$ sont des vecteurs propres associés.

On procède récursivement. Posons $A_1 = A$. On sait qu'il existe $\lambda_1 \in \mathbb{R}$ et $v_1 \in \mathbb{R}^n$ satisfaisant $\|v_1\|_2 = 1$, tels que $A_1 v_1 = \lambda_1 v_1$. Il peut y avoir d'autres vecteurs propres linéairement indépendants de v_1, associés à la valeur propre λ_1. Notons $\{v_1, \ldots, v_{m_1}\}$ une base orhtonormée du sous-espace propre associé à λ_1. Posons ensuite

$$A_2 = A - \lambda_1 (P_{v_1} + \cdots + P_{v_{m_1}}).$$

La matrice A_2 est symétrique. Cette matrice admet une valeur propre nulle, une partie au moins du sous-espace propre étant engendré par les vecteurs propres $v_1, \ldots, v_{m_1}$. Deux possibilités peuvent se présenter. Soit la matrice A_2 est constante. Alors 0 est l'unique valeur propre de A_2, et A_2 est en fait nulle. Dans ce cas, la question est résolue :

$$A = \lambda_1 (P_{v_1} + \cdots + P_{v_{m_1}}) = \lambda_1 (P_{v_1} + \cdots + P_{v_{m_1}}) + 0(P_{v_{m_1+1}} + \cdots + P_{v_n})$$

où $v_{m_1+1}, \ldots, v_n$ sont des vecteurs tels que $v_1, \ldots, v_n$ soit une base orthonormée de $\mathbb{R}^n$. Soit la matrice A_2 n'est pas constante. Dans ce cas il existe une autre valeur propre $\lambda_2 \neq 0$. Notons $\{v_{m_1+1}, \ldots v_{m_2}\}$ une base orthonormée du sous-espace propre associé. On vérifie que le sous-espace propre associé à λ_2 est orthogonal au sous-espace propre associé à λ_1. Posons ensuite

$$A_3 = A_2 - \lambda_2 (P_{v_{m_1}+1} + \cdots + P_{v_{m_2}}) = A - \lambda_1 (P_{v_1} + \cdots + P_{v_{m_1}}) - \lambda_2 (P_{v_{m_1}+1} + \cdots + P_{v_{m_2}}).$$

A nouveau cette matrice admet une valeur propre nulle et cette fois le sous-espace propre associé contient les vecteurs $v_1, \ldots, v_{m_2}$. On voit que l'on peut itérer ce procédé jusqu'à obtenir une matrice nulle (ce qui sera fait en au plus n étapes), ce qui donne le résultat.