Chapitre 20

Fonctions de plusieurs variables définies sur un ouvert de $\mathbb{R}^n$

Compétences attendues.

- ✓ Reconnaitre la nature topologique d'un ensemble (ouvert, convexe).
- ✓ Montrer qu'une fonction est continue, de classe $\mathscr{C}^1$ ou de classe $\mathscr{C}^2$.
- ✓ Calculer le gradient et la hessienne d'une fonction.
- ✓ Déterminer la nature local d'un point critique à l'aide de la hessienne.
- ✓ Déterminer les extrema locaux d'une fonction sur un ouvert.
- ✓ Déterminer les extrema globaux d'une fonction sur un ouvert convexe.

Mathieu Mansuy

Professeur en ECG deuxième année spécialité mathématiques approfondies au Lycée Louis Pergaud (Besançon)

Page personnelle : `mathieu-mansuy.fr/`

E-mail : `mansuy.mathieu@hotmail.fr`

Motivations

Dans ce chapitre, nous nous intéressons à la recherche d'extremum pour des fonctions de plusieurs variables définies sur un sous-ensemble Ω de $\mathbb{R}^n$.

Exemple. La fonction $f : (x, y) \mapsto \sqrt{y} \cdot \ln(x)$ est définie sur $\Omega = \mathbb{R}_+^* \times \mathbb{R}_+$.

La recherche d'extremum de la fonction dépend de la nature topologique de l'ensemble de définition Ω. On peut déjà le constater sur une fonction d'une variable réelle. Prenons par exemple la fonction f définie sur $\Omega = [2, 5]$ par $f(x) = x^2$. Elle admet un unique maximum en $x = 5$ et un unique minimum en $x = 2$. Or ce ne sont pas des points critiques de f puisque $f'(5) = 10$ et $f'(2) = 4$. Cela provient du fait que $[2, 5]$ n'est pas un intervalle **ouvert**. En effet, vous avez vu l'an dernier qu'un extremum d'une fonction dérivable f est atteint en un point critique de f si f est définie sur un **intervalle ouvert**.

> *Par quoi remplacer « intervalle ouvert » pour que ce résultat reste valable pour des fonctions de plusieurs variables ?*

D'autre part, la nature d'un point critique pour une fonction d'une variable réelle peut être déterminée grâce au signe de la dérivée seconde.

> *Par quoi remplacer le signe de la dérivée seconde pour obtenir une généralisation de ce résultat aux fonctions de plusieurs variables ?*

1 Parties ouvertes de $\mathbb{R}^n$

Dans tout le chapitre, n désignera un entier supérieur ou égal à 1, et $\mathbb{R}^n$ sera muni de sa norme euclidienne définie par :

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad \|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

1.1 Boules ouvertes, boules fermées

Vocabulaire. Soit $a \in \mathbb{R}^n$ et $r > 0$.

- On appelle *boule ouverte* de centre a et de rayon r l'ensemble :

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

- On appelle *boule fermée* de centre a et de rayon r l'ensemble :

$$\mathscr{B}_f(a, r) = \{x \in \mathbb{R}^n,\ \|x - a\| \leq r\}$$

Exemple. Représentons $\mathscr{B}_o(a, r)$ et $\mathscr{B}_f(a, r)$ lorsque $n = 1, 2$.

1.2 Partie ouverte

Définition.

- On appelle *ouvert élémentaire* de $\mathbb{R}^n$ tout sous-ensemble de $\mathbb{R}^n$ de l'une des formes suivantes :

$$\{x \in \mathbb{R}^n,\ \varphi(x) < a\} \quad \text{ou} \quad \{x \in \mathbb{R}^n,\ \varphi(x) > a\}$$

où $a \in \mathbb{R}$ et $\varphi : \mathbb{R}^n \to \mathbb{R}$ est une fonction continue sur $\mathbb{R}^n$.

- On dit qu'un sous-ensemble de $\mathbb{R}^n$ est un *ouvert* de $\mathbb{R}^n$ s'il est la réunion (finie ou infinie) d'ouverts élémentaires.

Remarque. Il faut retenir qu'un ensemble défini par une fonction continue sur $\mathbb{R}^n$ et des inégalités strictes est un ouvert de $\mathbb{R}^n$.

Exemples.

- La fonction $\varphi : x \in \mathbb{R}^n \mapsto \|x - a\| = \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2}$ est continue sur $\mathbb{R}^n$ comme composée de la fonction $x \mapsto (x_1 - a_1)^2 + \cdots + (x_n - a_n)^2$, polynomiale donc continue sur $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}_+$, par la fonction continue $u \in \mathbb{R}_+ \mapsto \sqrt{u}$ continue sur $\mathbb{R}_+$. Et on a :

$$\mathscr{B}_o(a, r) = \{x \in \mathbb{R}^n,\ \varphi(x) < r\}.$$

Ainsi $\mathscr{B}_o(a, r)$ est une partie ouverte de $\mathbb{R}^n$.

- Pour $n = 1$, on a en particulier que pour tout $a < b$, l'intervalle $]a, b[= \mathscr{B}_o\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$ est un ouvert de $\mathbb{R}$.

Pour tout $a \in \mathbb{R}$, $]a, +\infty[= \bigcup_{n=1}^{+\infty}]a, a + n[$ et $] - \infty, a[= \bigcup_{n=1}^{+\infty}]a - n, a[$ sont des ouverts de $\mathbb{R}$.

- $\mathbb{R}^* =] - \infty, 0[\bigcup]0, +\infty[$ est un ouvert de $\mathbb{R}$.

Exercice. Représenter le sous-ensemble $A = \{(x, y) \in \mathbb{R}^2,\ x + y \neq 0\}$ de $\mathbb{R}^2$, puis montrer qu'il est ouvert.

Propriété 1

Soit A une partie de $\mathbb{R}^n$. A est un ouvert de $\mathbb{R}^n$ si et seulement si pour tout $x \in A$, il existe $r > 0$ tel que :

$$\mathscr{B}_o(x, r) \subset A.$$

Remarque. Une partie A est un ouvert de $\mathbb{R}^n$ si pour tout $x \in A$, les points suffisamment proches de x sont aussi dans A. Graphiquement, un ouvert est une partie qui ne contient pas son bord.

Preuve.

$\Rightarrow$ Puisque tout ouvert est la réunion d'ouverts élémentaires, il suffit de montrer cette propriété pour un ouvert élémentaire. Montrons le pour l'ouvert $A = \{x \in \mathbb{R}^n,\ \varphi(x) < a\}$ où $a \in \mathbb{R}$ et $\varphi : \mathbb{R}^n \to \mathbb{R}$ est continue (on procèderait de même pour l'autre « type » d'ouverts élémentaires). Soit donc $x_0 \in A$, de sorte que $\varphi(x_0) < a$. Puisque φ est continue en x_0, on a :

$$\forall \varepsilon > 0,\ \exists r > 0,\ \forall x \in \mathbb{R},\ \|x - x_0\| < r\ \Rightarrow\ |\varphi(x) - \varphi(x_0)| < \varepsilon,$$

ce qui se réécrit :

$$\forall \varepsilon > 0,\ \exists r > 0,\ \forall x \in \mathbb{R},\ x \in \mathscr{B}_o(x_0, r)\ \Rightarrow\ \varphi(x) \in]\varphi(x_0) - \varepsilon, \varphi(x_0) + \varepsilon[.$$

Puisque $\varphi(x_0) < a$, on peut choisir $0 < \varepsilon < a - \varphi(x_0)$ de sorte que $\varphi(x_0) + \varepsilon < a$. Pour un tel ε, on a donc l'existence de $r > 0$ tel que :

$$\forall x \in \mathscr{B}_o(x_0, r),\ \varphi(x) < \varphi(x_0) + \varepsilon < a.$$

Ainsi on a :

$$\mathscr{B}_o(x_0, r) \subset A.$$

D'où le résultat pour A.

$\Leftarrow$ Réciproquement, soit A une partie de $\mathbb{R}^n$ telle que pour tout $a \in A$, il existe $r_a > 0$ tel que :

$$\mathscr{B}_o(a, r_a) \subset A.$$

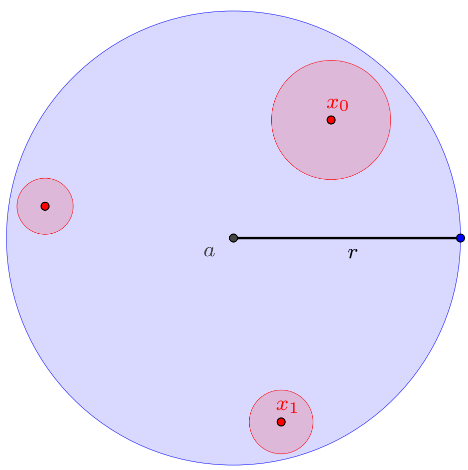
Alors $\bigcup_{a \in A} \mathscr{B}_o(a, r_a)$ est inclus dans A par définition des r_a. Et réciproquement, on a pour tout $a_0 \in A$:

$$a_0 \in \mathscr{B}_o(a_0, r_{a_0}) \subset \bigcup_{a \in A} \mathscr{B}_o(a, r_a).$$

Ainsi on a $A = \bigcup_{a \in A} \mathscr{B}_o(a, r_a)$, et A est bien un ouvert en tant qu'union des ouverts élémentaires $\mathscr{B}_o(a, r_a)$.

$\square$

Exemple. On vérifie graphiquement que cette propriété est bien satisfaite par la boule ouverte $\mathscr{B}_o(a, r)$ de $\mathbb{R}^2$ de centre a et de rayon $r > 0$.



Une boule ouverte est un ouvert.

Exemples.

- Les ensembles $\mathbb{R}^n$ et $\emptyset$ sont ouverts.
- L'intervalle $[0,1]$ n'est pas un ouvert de $\mathbb{R}$. En effet, on a pour tout $r>0$ que :
$$\mathscr{B}_o(0,r) =]-r,r[\not\subseteq [0,1].$$
Plus généralement, on montre que tout intervalle fermé en l'une de ses extrémités n'est pas un ouvert de $\mathbb{R}$. Ainsi **un intervalle est un ouvert de $\mathbb{R}$ si et seulement s'il est de la forme $]a,b[$ avec $-\infty \leq a < b \leq +\infty$.** Par exemple, les intervalles suivants sont des ouverts de $\mathbb{R}$:
$$]-1,1[, \quad]0,+\infty[\quad \text{ou} \quad]-\infty,3[.$$
A contrario, les intervalles suivants ne sont pas des ouverts de $\mathbb{R}$:
$$[-1,1], \quad [0,+\infty[\quad \text{ou} \quad]0,3].$$

Propriété 2

- Une intersection **finie** d'ouverts[a] de $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$.
- Un produit cartésien d'ensembles ouverts est un ouvert.

[a] Attention, c'est faux en général pour une intersection infinie d'ouverts.

Exemple. $]0,1[\times]-2,2[= \{(x,y)\in\mathbb{R}^2,\ 0<x<1 \text{ et } -2<y<2\}$ est un ouvert de $\mathbb{R}^2$.

Exercice. Montrer que l'ensemble $A = \{(x,y)\in\mathbb{R}^2,\ xy>1 \text{ et } x+y<2\}$ est un ouvert de $\mathbb{R}^2$.

2 Fonctions continues et $\mathscr{C}^1$ sur une partie de $\mathbb{R}^n$

On s'est jusqu'à présent intéressé aux fonctions définies sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. On étend ici les notions étudiées aux fonctions f définies sur un sous-ensemble Ω de $\mathbb{R}^n$.

Exemples.

- La fonction $f:(x,y)\mapsto\sqrt{x+y}$ est définie sur $\Omega=\{(x,y)\in\mathbb{R}^2,\ x+y\geq 0\}$.
- La fonction $f:(x_1,\dots,x_n)\mapsto \dfrac{1}{\sum_{k=1}^{n} x_k^2}$ est définie sur $\Omega=\mathbb{R}^n\setminus\{(0,\dots,0)\}$.

2.1 Fonctions continues

Définition.

Soit $\Omega\subset\mathbb{R}^n$. Une fonction $f:\Omega\to\mathbb{R}$ est dite *continue en* $x_0\in\Omega$ si :
$$\forall\varepsilon>0,\ \exists\nu>0,\ \forall x\in\Omega,\ \|x-x_0\|\leq\nu \ \Rightarrow\ |f(x)-f(x_0)|\leq\varepsilon.$$
f est *continue sur* Ω si elle est continue en tout point de Ω.

Tous les résultats obtenus sur les fonctions continues de $\mathbb{R}^n$ dans $\mathbb{R}$ se transposent aux fonctions continues de Ω dans $\mathbb{R}$. En particulier, les fonctions polynomiales sont continues, la somme de fonctions continues est continue, ...

Exercice. Montrer la continuité de la fonction $f : (x, y) \in]0, +\infty[\times]0, +\infty[\mapsto \ln(x^2 + y)$.

2.2 Fonctions de classe $\mathscr{C}^1$

On étend ici la notion de fonctions de classe $\mathscr{C}^1$ aux fonctions définies sur un **ouvert** Ω de $\mathbb{R}^n$.
Soit $a = (a_1, \dots, a_n) \in \Omega$. Ω étant un ouvert de $\mathbb{R}^n$, il existe $r > 0$ tel que $\mathscr{B}_o(a, r) \subset \Omega$. Pour tout $i \in [\![1, n]\!]$, on définit la *i-ème fonction partielle de f en a* en posant :

$$f_{a,i} : \left\{ \begin{array}{ccc}]-r, r[& \to & \mathbb{R} \\ t & \mapsto & f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) \end{array} \right.$$

Définition.

Soit $\Omega \subset \mathbb{R}^n$ un ouvert, $f : \Omega \to \mathbb{R}$ une fonction, et $a = (a_1, \dots, a_n) \in \Omega$.

- On dit que f *admet une dérivée partielle d'ordre 1 par rapport à la i-ème variable en a* lorsque la fonction $f_{a,i}$ est dérivable en 0, c'est-à-dire si

$$\lim_{t \to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{t} \quad \text{existe.}$$

  On appelle alors *dérivée partielle d'ordre 1 de f par rapport à la i-ème variable en a* et on note $\partial_i(f)(a)$ ce réel.

- Lorsque toutes les dérivées partielles de f en a existent, on appelle *gradient de f en a*, et on note $\nabla f(a)$, le vecteur de $\mathbb{R}^n$ défini par :

$$\nabla f(a) = (\partial_1 f(a), \dots, \partial_n f(a)).$$

Définition.

Soit $\Omega \subset \mathbb{R}^n$ un ouvert, et $f : \Omega \to \mathbb{R}$. On dit que f *est de classe $\mathscr{C}^1$ sur* Ω si toutes les applications $\partial_i f$ sont définies et continues sur Ω.

Tous les théorèmes sur les fonctions de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ restent valables sur un ouvert Ω de $\mathbb{R}^n$. En particulier, les fonctions polynomiales sont encore de classe $\mathscr{C}^1$, la somme de fonctions $\mathscr{C}^1$ est $\mathscr{C}^1$, ...
La formule de Taylor à l'ordre 1 est également valable sur un ouvert de $\mathbb{R}^n$.

3 Fonctions de classe $\mathscr{C}^2$

3.1 Dérivées partielles d'ordre 2

Définition.

Soit f une fonction de classe $\mathscr{C}^1$ sur un ouvert $\Omega \subset \mathbb{R}^n$.
Soit $(i, j) \in [\![1, n]\!]^2$. La fonction $\partial_j f$ est une fonction continue de Ω dans $\mathbb{R}$. Si elle admet une dérivée partielle par rapport à la i-ème variable, on note :

$$\partial^2_{i,j} f = \partial_i(\partial_j(f))$$

cette dérivée partielle. On l'appelle la *dérivée partielle d'ordre 2 d'indice (i, j) de f*.

Notation.

La notation $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ est aussi communément utilisée pour la dérivée partielle d'ordre 2 d'indice (i,j). Le programme officiel privilégie cependant la notation $\partial_{i,j} f$.

Exercices. Déterminer les dérivées partielles secondes, si elles existent, de la fonction f définie sur $]0,+\infty[\times]0,+\infty[$ par $f(x,y) = \ln(x) + \ln(y) - xy^2$.

Définition.

On suppose que f admet une dérivée partielle d'ordre 2 par rapport à chaque couple de variables sur Ω. On appelle *matrice hessienne de f au point $a \in \Omega$* (du nom du mathématicien allemand Ludwig Otto Hesse (1811 - 1874)) et on note $\nabla^2(f)(a)$ la matrice de $\mathscr{M}_n(\mathbb{R})$ définie par :

$$\nabla^2(f)(a) = \left(\partial^2_{i,j}(a)\right)_{(i,j)\in[\![1,n]\!]^2} = \begin{pmatrix} \partial^2_{1,1}(f)(a) & \partial^2_{1,2}(f)(a) & \cdots & \partial^2_{1,n}(f)(a) \\ \partial^2_{2,1}(f)(a) & \partial^2_{2,2}(f)(a) & \cdots & \partial^2_{2,n}(f)(a) \\ \vdots & \vdots & & \vdots \\ \partial^2_{n,1}(f)(a) & \partial^2_{2,2}(f)(a) & \cdots & \partial^2_{n,n}(f)(a) \end{pmatrix}.$$

Mise en garde.

Attention aux confusions :

- Le gradient $\nabla(f)(a)$ est **un vecteur** contenant les dérivées partielles d'ordre 1 en a.
- La hessienne $\nabla^2(f)(a)$ est **une matrice** contenant les dérivées partielles d'ordre 2 en a.

Attention aussi à la notation $\nabla^2(f)(a)$: malgré la présence de l'exposant 2, elle ne désigne pas un carré. En particulier, $(\nabla(f)(a))^2$ n'a aucun sens (un vecteur de $\mathbb{R}^n$ au carré... ☠). Quant à $\nabla(f^2)(a)$, il désigne le gradient de la fonction $f \times f$ en a.

Exercice. Soit $g : \mathbb{R}^n \to \mathbb{R}$ la fonction définie par :

$$g(x_1,\ldots,x_n) = \sum_{k=1}^n x_k^2 + \left(\sum_{k=1}^n x_k\right)^2 - \sum_{k=1}^n x_k.$$

Écrire, si elle existe, la matrice hessienne de g en $(x_1,\ldots,x_n) \in \mathbb{R}^n$.

3.2 Fonctions de classe $\mathscr{C}^2$

Définition.

Soit f une fonction de classe $\mathscr{C}^1$ sur un ouvert Ω de $\mathbb{R}^n$.
On dit que f *est une fonction de classe* $\mathscr{C}^2$ *sur* Ω si elle admet des dérivées partielles d'ordre 2 d'indice (i,j) pour tout $(i,j) \in [\![1,n]\!]^2$, et si ses dérivées partielles sont des fonctions continues de Ω dans $\mathbb{R}$.

Remarque. On a le schéma suivant :

$$\begin{array}{ccccc}
f \text{ de classe } \mathscr{C}^2 \text{ sur } \Omega & \Longrightarrow & f \text{ de classe } \mathscr{C}^1 \text{ sur } \Omega & \Longrightarrow & f \text{ continue sur } \Omega \\
\Updownarrow & & \Updownarrow & & \\
\text{Dérivées partielles d'ordre 2} & & \text{Dérivées partielles d'ordre 1} & & \\
\text{continues sur } \Omega & & \text{continues sur } \Omega & &
\end{array}$$

Exemple. Les fonctions f et g de la section précédente sont de classe $\mathscr{C}^2$ sur leur ensemble de définition.

Propriété 3

Les fonctions polynomiales sont de classe $\mathscr{C}^2$ sur $\mathbb{R}^n$.

Preuve. Soit f une fonction polynomiale sur $\mathbb{R}^n$. On sait déjà que f est de classe $\mathscr{C}^1$. De plus pour tout $j \in [\![1,n]\!]$, $\partial_j f$ est encore une fonction polynomiale sur $\mathbb{R}^n$. Elle est donc aussi $\mathscr{C}^1$. Ainsi $\partial_i(\partial_j f)$ existe et est continue pour tout $i \in [\![1,n]\!]$. On en déduit que f admet des dérivées partielles secondes d'indice (i,j) continues sur $\mathbb{R}^n$ pour tout $(i,j) \in [\![1,n]\!]$. Donc f est de classe $\mathscr{C}^2$ sur $\mathbb{R}^n$. $\square$

Propriété 4

Soient f et g des fonctions de classe $\mathscr{C}^2$ sur Ω, et $\lambda \in \mathbb{R}$.

- Les fonctions $\lambda \cdot f$, $f+g$ et $f \times g$ sont de classe $\mathscr{C}^2$ sur Ω.
- Si de plus f ne s'annule pas sur Ω, alors les fonctions $\dfrac{1}{f}$ et $\dfrac{g}{f}$ sont de classe $\mathscr{C}^2$ sur Ω.

Propriété 5

Soit $I \subset \mathbb{R}$ un intervalle. Soit $\varphi : I \to \mathbb{R}$. Supposons que :

- f est de classe $\mathscr{C}^2$ sur un ouvert Ω, à valeurs dans I ;
- φ est de classe $\mathscr{C}^2$ sur I.

Alors $\varphi \circ f$ est de classe $\mathscr{C}^2$ sur Ω.

Exercice. Montrer que la fonction $f : (x_1, \dots, x_n) \mapsto e^{-(x_1^2+\cdots+x_n^2)}$ est de classe $\mathscr{C}^2$ sur $\mathbb{R}^n$.

3.3 Théorème de Schwarz

Théorème 6 (*de Schwarz* (1843 - 1921))

Si f est de classe $\mathscr{C}^2$ sur Ω, alors pour tout $(i,j) \in [\![1,n]\!]^2$ et pour tout point $a \in \Omega$, on a :

$$\partial^2_{j,i}(f)(a) = \partial^2_{i,j}(f)(a).$$

Conséquences. Soit f est une fonction de classe $\mathscr{C}^2$ sur un ouvert $\Omega \subset \mathbb{R}^n$. Alors :

- l'ordre dans lequel on dérive (par rapport à la i-ème puis la j-ème variable ou l'inverse) n'a pas d'importance ;
- pour tout $a \in \Omega$, la matrice hessienne $\nabla^2(f)(a)$ est symétrique réelle. En particulier :
  - la matrice $\nabla^2(f)(a)$ diagonalise dans une base orthonormée de vecteurs propres ;
  - on peut s'intéresser à la forme quadratique q_a associée à la matrice $\nabla^2(f)(a)$ définie pour tout $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ par :

$$q_a(h) = (h_1 \dots h_n)\nabla^2(f)(a) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = (h_1 \dots h_n) \begin{pmatrix} \displaystyle\sum_{j=1}^n \partial_{1,j} f(a) h_j \\ \vdots \\ \displaystyle\sum_{j=1}^n \partial_{n,j} f(a) h_j \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n \partial_{i,j} f(a) h_i h_j.$$

3.4 Développement limité d'ordre 2

Théorème 7 (*Formule de Taylor* (1685 - 1731) *d'ordre 2*)

Soit f une fonction de classe $\mathscr{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$, et soit a un point de Ω.
Alors il existe une fonction ε continue en $(0, \dots, 0)$, vérifiant $\varepsilon(0, \dots, 0) = 0$ et telle que pour tout $h \in \mathbb{R}^n$ « suffisamment petit[a] », on a :

$$f(a+h) = \underbrace{f(a) + \langle \nabla(f)(a), h \rangle + \frac{1}{2} q_a(h)}_{\text{partie principale du DL}} + \underbrace{\|h\|^2 \varepsilon(h)}_{\text{reste du DL}},$$

où q_a désigne la forme quadratique associée à la matrice symétrique $\nabla^2(f)(a)$.

Une telle écriture est unique, et est appelée *développement limité de f d'ordre 2 au point a* (ou *formule de Taylor à l'ordre 2 en a*).

[a] Ω étant ouvert, il existe $r > 0$ tel que $\mathscr{B}_o(a,r) \subset \Omega$. Par h « petit », on entend $\|h\| < r$ de sorte que $a + h \in \Omega$.

Remarque. Si on note $a = (a_1, \dots, a_n)$ et $h = (h_1, \dots, h_n)$, ce développement limité s'écrit :

$$f(a+h) = f(a) + \underbrace{\sum_{i=1}^n \partial_i(f)(a) h_i}_{\text{terme de degré 1}} + \underbrace{\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \partial^2_{i,j} f(a) h_i h_j}_{\text{terme de degré 2}} + (h_1^2 + \dots + h_n^2)\varepsilon(h).$$

En particulier pour $n = 1$, on retrouve le développement limité à l'ordre 2 en a d'une fonction d'une variable réelle :
$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}f''(a)h^2 + h^2\varepsilon(h).$$

Exercice. Écrire de développement limité d'ordre 2 de $f : (x, y) \mapsto \ln(x) + \ln(y) - xy^2$ au point $a = (1, 1)$.

3.5 Dérivées directionnelles

Propriété 8

Soit f une fonction de classe $\mathscr{C}^2$ sur un ouvert Ω, soit $a \in \Omega$ et $h \in \mathbb{R}^n$ non nul.
Alors la fonction $g : t \mapsto f(a + th)$ est définie au voisinage de 0, y est deux fois dérivable et on a :
$$g'(t) = \langle \nabla f(a + th), h\rangle \quad \text{et} \quad g''(t) = q_{a+th}(h).$$
En particulier, on a $g''(0) = q_a(h)$.

Preuve. Puisque Ω est ouvert, il existe $r > 0$ tel que $\mathscr{B}_o\left(a, r \times \|h\|\right) \subset \Omega$, de sorte que $a + th \in \Omega$ pour tout $t \in]-r, r[$. Ainsi g est définie sur l'intervalle $] - r, r[$, et nous avons déjà établi que g y est $\mathscr{C}^1$ et que sa dérivée est :
$$g' : t \mapsto \langle \nabla f(a + th), h\rangle = \sum_{i=1}^{n} h_i \partial_i f(a + th).$$
Mais $\partial_i f$ est de classe $\mathscr{C}^1$ (car f est $\mathscr{C}^2$) et donc $t \mapsto \partial_i f(a + th)$ est dérivable, de dérivée :
$$t \mapsto \sum_{j=1}^{n} \partial_j(\partial_i(f))(a + th)h_j.$$
On en déduit que g' est dérivable et on a :
$$\forall t \in]-r, r[, \quad g''(t) = \sum_{i=1}^{n} h_i \sum_{j=1}^{n} \partial^2_{j,i} f(a + th)h_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \partial^2_{i,j} f(a + th)h_i h_j = q_{a+th}(h).$$

□

4 Recherche d'extremum

4.1 Extrema locaux, extrema globaux

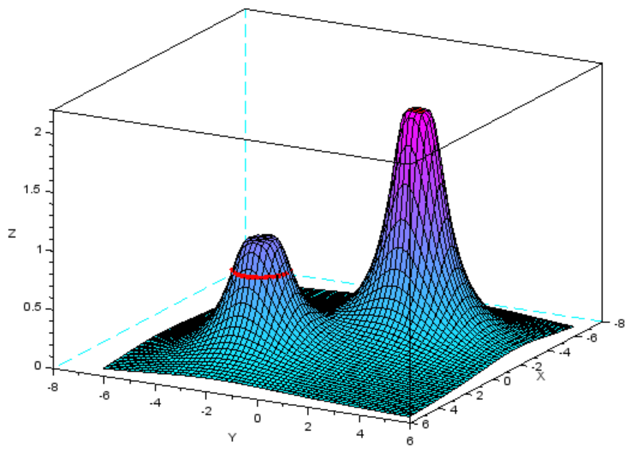
Définition.

Soit A une partie de $\mathbb{R}^n$, $f : A \to \mathbb{R}$ une fonction, et $a \in A$.

- On dit que f admet un *minimum global en* a (resp. *maximum global en* a) lorsque :
$$\forall x \in A, \quad f(x) \geq f(a) \quad (\text{resp. } f(x) \leq f(a)).$$
- On dit que f admet un *extremum global en* a lorsque f admet un minimum global ou un maximum global en a.
- On dit que f admet un *minimum local en* a (resp. *maximum local en* a) lorsqu'il existe $r > 0$ tel que :
$$\forall x \in A, \quad x \in \mathscr{B}_o(a, r) \quad \Rightarrow \quad f(x) \geq f(a) \quad (\text{resp. } f(x) \leq f(a)).$$
- On dit que f admet un *extremum local en* a lorsque f admet un minimum local ou un maximum local en a.

Remarque. Un extremum global est un extremum local.

Exemple. Considérons la fonction f dont le graphe $\mathscr{G}_f$ est représenté ci-dessous.



f admet un maximum global en $a = (-2, 2)$, et un maximum local en $b = (2, -2)$: en effet, pour tout $x \in \mathscr{B}_o(b, 1)$, c'est-à-dire graphiquement pour tout $x \in \mathbb{R}^2$ à l'intérieur du disque rouge, on a $f(x) \leq f(b)$.

4.2 Condition nécessaire d'extremum local

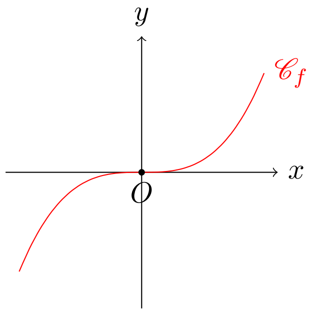
Théorème 9 (*Condition nécessaire d'extremum local*)

On suppose que f est de classe $\mathscr{C}^1$ sur un **ouvert** Ω de $\mathbb{R}^n$.

Si f admet un extremum local en $a \in \Omega$, **alors** a est un point critique de f, c'est-à-dire $\nabla(f)(a) = (0, \ldots, 0)$.

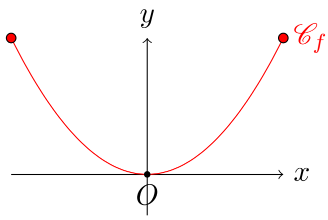
Mise en garde.

- **La réciproque est fausse :** si a est un point critique de f, alors a n'est pas nécessairement un extremum local. Par exemple pour la fonction $f : x \in \mathbb{R} \mapsto x^3$, 0 est un point critique de f puisque $f'(0) = 0$, mais f n'admet pas d'extremum local en 0.



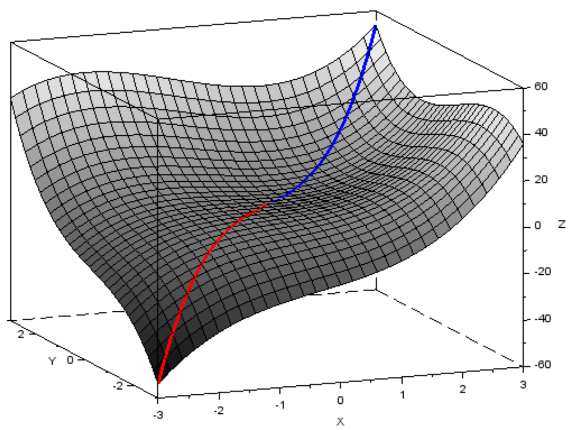
Courbe représentative de la fonction $x \mapsto x^3$.

- **L'hypothèse Ω ouvert est essentielle :** par exemple la fonction $f : x \in [-1,1] \mapsto x^2$ admet deux maximum globaux sur l'intervalle fermé $[-1,1]$ en -1 et en 1, et pourtant ce ne sont pas des points critiques puisque $f'(-1) \neq 0$ et $f'(1) \neq 0$.



Courbe représentative de la fonction $x \mapsto x^2$ sur $[-1,1]$.

Exercice. Déterminer les extrema éventuels de la fonction $f : (x,y) \in \mathbb{R}^2 \mapsto (x-y)^2 + x^3 + y^3$.



Toute boule de centre $(0,0)$ contient des points de la droite $y = x$ d'image positive et d'autres d'image négative.

4.3 Condition suffisante d'extremum local

Cas des fonctions d'une variable réelle

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$, et $a \in \mathbb{R}$ un point critique de f (c'est-à-dire tel que $f'(a) = 0$). Pour déterminer si a est un extremum local ou non, on peut dresser le tableau de variation de f. On peut aussi le déduire du signe de $f''(a)$ comme suit. Par la formule de Taylor Young, on a :

$$f(a+h) - f(a) = \underbrace{f'(a)}_{=0} h + \frac{f''(a)}{2} h^2 + o(h^2) = \frac{f''(a)}{2} h^2 + o(h^2).$$

On en déduit que :

- si $f''(a) > 0$, alors on a
$$f(a+h) - f(a) \underset{h \to 0}{\sim} \frac{f''(a)}{2} h^2.$$
$f(a+h) - f(a)$ est donc du signe de $\dfrac{f''(a)}{2} h^2$, c'est-à-dire positif, lorsque h est au voisinage de 0. Ainsi f admet un minimum local en a. Plus précisément, on peut montrer que ce minimum local est strict : dans un voisinage de a, on a $f(x) > f(a)$ pour tout $x \neq a$.
- si $f''(a) < 0$, alors de même f admet un maximum local strict en a.
- si $f''(a) = 0$, on ne peut pas conclure[1].

Cas des fonctions de plusieurs variables

On généralise à présent la discussion précédente aux fonctions de plusieurs variables afin de déterminer la nature locale des points critiques.

Théorème 10 (*Condition suffisante d'extremum local*)

Soit f une fonction de classe $\mathscr{C}^2$ sur un ouvert Ω, et soit $a = (a_1, \ldots, a_n) \in \Omega$ un point critique de f. Alors :

- si $\mathrm{Spec}(\nabla^2(f)(a)) \subset \mathbb{R}_+^*$, alors f admet un minimum local en a ;
- si $\mathrm{Spec}(\nabla^2(f)(a)) \subset \mathbb{R}_-^*$, alors f admet un maximum local en a ;
- si $\nabla^2(f)(a)$ admet une valeur propre strictement positive et une valeur propre strictement négative, alors f n'admet ni maximum local, ni minimum local en a.

Preuve.

□

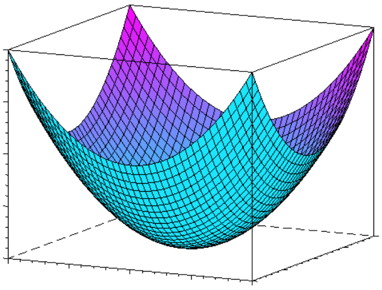
[1] Il faudrait utiliser la formule de Taylor à un ordre supérieur jusqu'à obtenir une dérivée non nulle pour pouvoir conclure.

Remarque. Si $\mathrm{Spec}(\nabla^2(f)(a)) \subset \mathbb{R}_+$ (ou $\mathbb{R}_-$) et que l'une des valeurs propres est nulle, on ne peut pas conclure directement. On cherchera alors d'autres moyens pour déterminer la nature du point critique (étude du signe de $f(a+h) - f(a)$ pour h petit, ...).

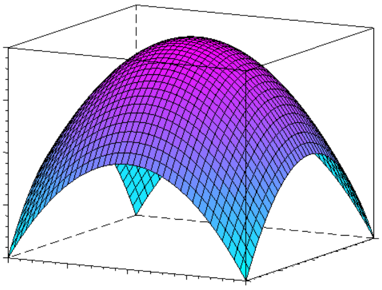
Remarque. Au voisinage d'un point critique a, on a donc d'après la formule de Taylor à l'ordre 2 :

$$f(x) \approx f(a) + \frac{1}{2}q_a(x-a).$$

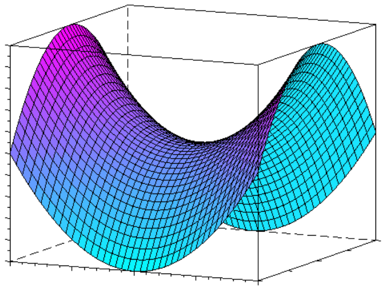
Lorsque $n = 2$, le graphe de f est donc semblable à celui de $x \mapsto f(a) + \frac{1}{2}q_a(x-a)$ au voisinage de a, et possède l'allure suivante, suivant le signe des valeurs propres λ_1 et λ_2 de $\nabla^2 f(a)$:



$\lambda_1 > 0$ et $\lambda_2 > 0$.



$\lambda_1 < 0$ et $\lambda_2 < 0$.



$\lambda_1 > 0$ et $\lambda_2 < 0$.

Définition.

Dans le cas où les deux valeurs propres de la hessienne sont non nulles de signes contraires, on dit qu'on a un *point selle* ou *point col*.

Exemple d'un col en montagne.

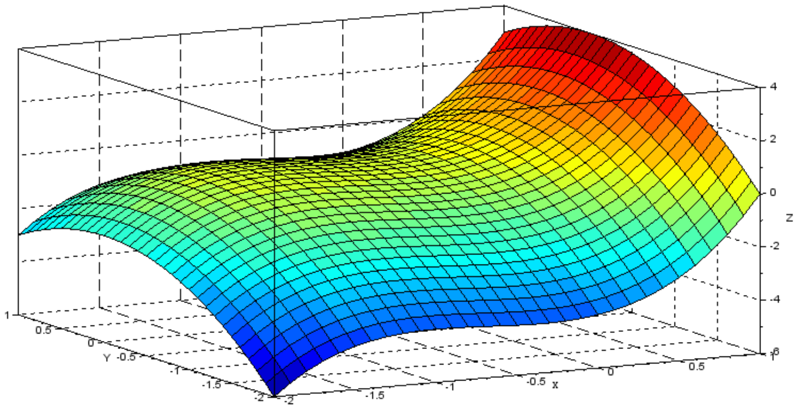
4.4 Méthode et exemples de recherche d'extrema locaux sur un ouvert

Méthode. Recherche d'extrema locaux.

Pour déterminer les extrema locaux d'une fonction f sur un ouvert Ω, on procèdera comme suit :

- *on justifie, si ce n'est pas dit dans l'énoncé, que Ω est ouvert et que f y est de classe $\mathscr{C}^2$;*
- *on calcule le gradient de f, puis on cherche les points critiques ;*
- *on calcule la hessienne de f en le (ou les) points critiques, puis on détermine ses valeurs propres ;*
- *on identifie la nature locale du point critique a à l'aide du signe des valeurs propres de $\nabla^2 f(a)$.*

Exercice. Déterminer les extrema locaux de la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto x(x+1)^2 - y^2$.



Exercice. Déterminer les extrema locaux de la fonction

$$g : (x_1, \ldots, x_n) \in \mathbb{R}^n \mapsto \sum_{k=1}^{n} x_k^2 + \left(\sum_{k=1}^{n} x_k\right)^2 - \sum_{k=1}^{n} x_k.$$

4.5 Condition suffisante d'extremum global sur un ouvert convexe

Rappels sur les fonctions convexes d'une variable réelle

Soit f une fonction de classe $\mathscr{C}^2$ sur un intervalle I de $\mathbb{R}$. Rappelons que f est *convexe* sur I si et seulement si l'une des conditions équivalentes suivantes est satisfaite :

(i) La courbe représentative $\mathscr{C}_f$ de f est en dessous de ses cordes sur I, ce qui s'exprime ainsi :

$$\forall (x_1, x_2) \in I^2, \quad \forall \lambda \in [0,1], \quad f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

(ii) La courbe représentative $\mathscr{C}_f$ de f est au dessus de ses tangentes sur I.

(iii) f' est croissante sur I.

(iv) Pour tout $x \in I$, $f''(x) \geq 0$.

Rappelons également que f est *concave* sur I si $-f$ est convexe sur I.

On dispose du résultat suivant, vu en première année.

Propriété 11 (Condition suffisante d'extremum global)

Soit f une fonction de classe $\mathscr{C}^2$ sur un intervalle ouvert I, et $a \in I$ un point critique de f.

- Si f est convexe sur I, alors f admet un minimum global en a.
- Si f est concave sur I, alors f admet un maximum global en a.

Preuve.

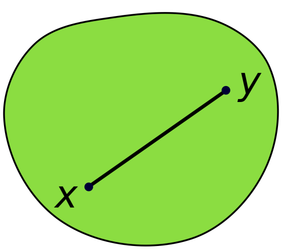
□

Généralisation au cas des fonctions de plusieurs variables

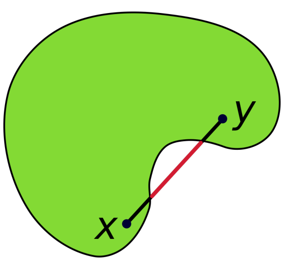
Définition.

On dit qu'un sous-ensemble E de $\mathbb{R}^n$ est convexe lorsque, pour tout couple (x, y) de E^2, le segment $[x, y]$ est inclus dans E, c'est-à-dire :

$$\forall (x,y) \in E^2, \ \forall \lambda \in [0,1], \ \lambda x + (1-\lambda) y \in E.$$



Exemple d'une partie convexe de $\mathbb{R}^2$.



Exemple d'une partie non convexe de $\mathbb{R}^2$.

Exemples.

- Toute boule ouverte ou fermée de $\mathbb{R}^n$ est convexe. $\mathbb{R}^n \setminus \{(0,\dots,0)\}$ n'est pas convexe.
- Dans $\mathbb{R}$, les parties convexes sont exactement les intervalles. Ainsi, $[0,1]$ et $]-2,+\infty[$ sont convexes, mais pas $[-2,-1]\cup]0,1]$ par exemple.

On peut à présent généraliser la condition suffisante d'extremum global au cas des fonctions de plusieurs variables comme suit.

Théorème 12 (Condition suffisante d'extremum global)

Soit f une fonction de classe $\mathscr{C}^2$ sur un ouvert convexe Ω de $\mathbb{R}^n$, et a un point critique de f.

- Si, **pour tout** $x \in \Omega$, $\mathrm{Spec}(\nabla^2(f)(x)) \subset \mathbb{R}_+$, alors f admet un minimum global en a.
- Si, **pour tout** $x \in \Omega$, $\mathrm{Spec}(\nabla^2(f)(x)) \subset \mathbb{R}_-$, alors f admet un maximum global en a.

Preuve.

$\square$

Exercice. Déterminer les extrema globaux de la fonction

$$g : (x_1,\dots,x_n) \in \mathbb{R}^n \mapsto \sum_{k=1}^n x_k^2 + \left(\sum_{k=1}^n x_k\right)^2 - \sum_{k=1}^n x_k.$$