Optimisation d'une fonction réelle de plusieurs variables

Étude au premier ordre

1. Point critique d'une application différentiable.
2. Condition nécessaire d'existence d'un extremum local sur un ouvert
3. Théorème d'optimisation sous une contrainte (théorème de Lagrange).

Étude au second ordre

4. Si f est une fonction de classe $\mathcal{C}^2$ sur un ouvert de $\mathbb{R}^n$ et si f admet un minimum local en x, alors x est point critique de f et $H_f(x) \in \mathcal{S}_n^+(\mathbb{R})$. Adaptation au cas d'un maximum local.
5. Si f est une fonction de classe $\mathcal{C}^2$ sur un ouvert de $\mathbb{R}^n$, si x est point critique de f et si $H_f(x) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f atteint un minimum local strict en x. Adaptation au cas d'un maximum local.
6. Explicitation pour $n = 2$ (trace et déterminant).

Dans ce chapitre, E est un espace vectoriel de dimension finie égale à n (très souvent, $E = \mathbb{R}^n$), et f est une fonction définie sur une partie A de E et à valeurs dans $\mathbb{R}$.

1 Généralités sur les extrema

Définition 1

Soit $a \in A$.
f admet un minimum (global) en a lorsque $\quad \forall x \in A,\ \ f(a) \leqslant f(x)$.
f admet un maximum (global) en a lorsque $\quad \forall x \in A,\ \ f(x) \leqslant f(a)$.
f admet un extremum (global) en a si f admet un minimum global ou un maximum global en a.
f admet un minimum local en a s'il existe $r > 0$ tel que pour tout $x \in A \cap B(a, r)$, $f(a) \leqslant f(x)$.
f admet un maximum local en a s'il existe $r > 0$ tel que pour tout $x \in A \cap B(a, r)$, $f(x) \leqslant f(a)$.

Si f admet un extremum (global) en a, alors elle admet un extremum local en a.

Exercice 1 :

1. Montrer que $f : (x, y) \mapsto (x-1)^2 + y^4 + 7$ admet un minimum sur $\mathbb{R}^2$ et déterminer le ou les points en lesquels ce minimum est atteint.
2. Montrer que cette fonction n'admet pas de maximum global.

Le théorème suivant pourra nous assurer l'existence d'un minimum et d'un maximum de f.

Théorème 1 – Théorème des bornes atteintes

Toute fonction continue sur un compact et à valeurs dans $\mathbb{R}$, est bornée et atteint ses bornes.

2 Recherche des extrema de f sur un ouvert : les extrema libres

2.1 étude à l'ordre 1 sur un ouvert

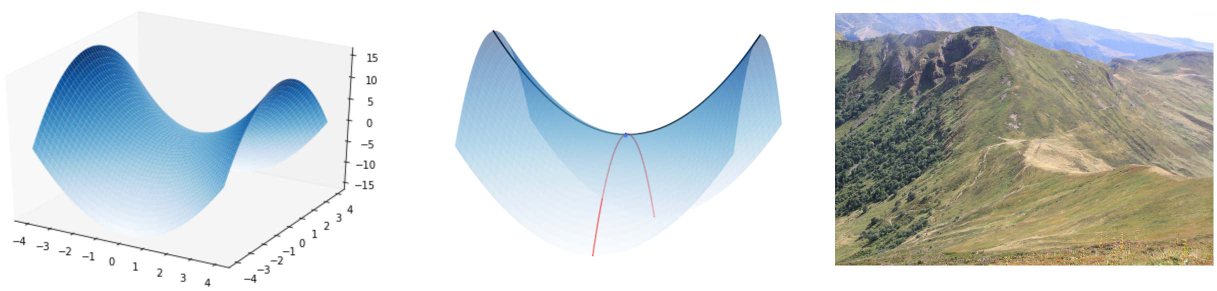
Propriété 1

Soit f de classe $\mathcal{C}^1$ sur un OUVERT $\mathcal{U}$. Si f admet un extremum local en $a \in \mathcal{U}$, alors

$$\mathrm{d}f(a) = 0 \quad \text{ou encore} \quad \nabla f(a) = 0$$

Cela revient à demander la nullité de toutes les dérivées partielles de f en a.
Un point a de $\mathcal{U}$ en lequel $\mathrm{d}f(a) = 0$ est appelé *point critique* de f.
Une fonction n'admet pas nécessairement d'extremum en un point critique, comme l'illustrent les graphes suivants.

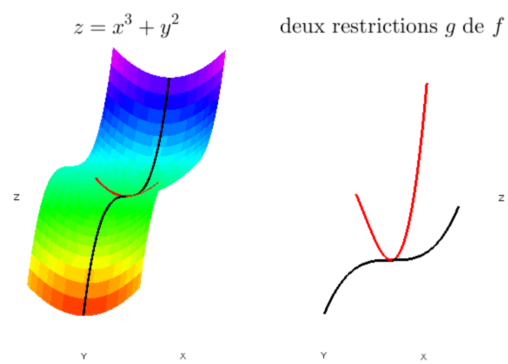
Point col



Gradient nul en un point sans extremum

Exercice 2 : Étudier les extrema locaux de

$$f : (x, y) \mapsto x^3 + y^2$$



Exercice 3 : On considère $f(x, y) = x^4 + y^3 - 3y - 2$.

1. Montrer que f n'admet pas de minimum global.
2. Montrer que f admet deux points critiques et les déterminer.
3. En effectuant à chaque fois une translation, déterminer si ces points critiques correspondent à des extrema locaux.

2.2 étude à l'ordre 2

Lorsque f est de classe $\mathcal{C}^2$ sur l'ouvert $\mathcal{U}$, le développement limité à l'ordre 2 de f en a est

$$f(a+h) = f(a) + \langle \nabla f(a) | h \rangle + \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2)$$

En un point critique a : $f(a+h) - f(a) = \frac{1}{2} h^\top H_f(a) h + \|h\|^2 \varepsilon(h)$.

Propriété 2 – ♡

Soit f est une fonction réelle de classe $\mathcal{C}^2$ sur un ouvert de $\mathbb{R}^n$.

- **condition nécessaire pour un minimum local**
  Si f admet un minimum local en a, alors $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$.
- **condition suffisante pour un minimum local**
  Si a est point critique de f et si $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f atteint un minimum local strict en a.

Remarques importantes :

— On a les adaptations pour un maximum local : si f admet un maximum local en a, alors $H_f(a) \in \mathcal{S}_n^-(\mathbb{R})$. Et si a est point critique de f et si $H_f(a) \in \mathcal{S}_n^{--}(\mathbb{R})$, alors f atteint un maximum local strict en a.

— Si $H_f(a)$ admet une valeur propre strictement positive et une valeur propre strictement négative, alors

— Il y a des situations où ce théorème n'apporte aucune conclusion. C'est le cas quand on sait seulement que toutes les valeurs propres sont positives ou nulles, ou quand toutes les valeurs propres sont négatives ou nulles.

Exercice 4 : $f(x,y,z) = (x+y+z)e^{-x^2-y^2-z^2}$ sur $\mathbb{R}^3$.

1. Montrer que f admet deux points critiques a et b, tels que $b = -a$.
2. Montrer que f admet des extrema locaux en a et b et donner leur nature.

Propriété 3 – Cas $n = 2$, théorème de Monge

Soit f une fonction réelle de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^2$ et soit a un point critique de f.

- Si $\det(H_f(a)) > 0$, f admet un extremum local en a.
  Il s'agit d'un minimum local si $\mathrm{Tr}(H_f(a)) > 0$, et d'un maximum local sinon.
- Si $\det(H_f(a)) < 0$, f admet un point col en a.
- Si $\det(H_f(a)) = 0$, on ne peut pas conclure par cette approche.

Exercice 5 (CCINP 2023) : On définit la fonction $f : (x, y) \mapsto x^2 - 2xy + 2y^2 + e^{-x}$ sur $\mathbb{R}^2$.

1. Établir que $e^{-x} = x$ admet une unique solution sur $\mathbb{R}$.
2. Démontrer que f admet un unique point critique (x_0, y_0) sur $\mathbb{R}^2$.
3. À l'aide de la matrice hessienne, démontrer que f admet un extremum local en (x_0, y_0). Est-ce un minimum ou un maximum ?

3 Recherche des extrema de f sous contrainte : les extrema liés

3.1 distinction intérieur - frontière

Méthode – distinction intérieur - frontière

Soit D une partie de E (non nécessairement ouverte). Les extrema globaux d'une fonction f à valeurs réelles sur D sont :

- soit atteints sur $\mathring{D}$, auquel cas ce sont des extrema locaux sur un ouvert, ce qui ramène à la section précédente,
- soit atteints sur $D \setminus \mathring{D}$.

Exercice 6 (B.E.O.) :

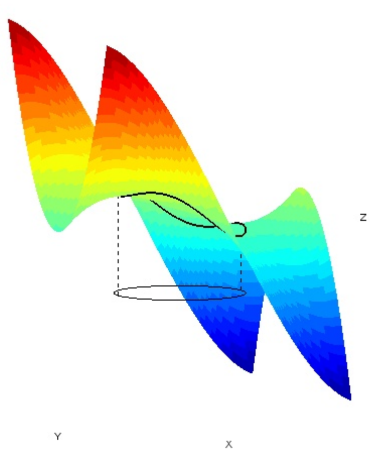
$$\forall (x, y) \in \mathbb{R}^2,\ f(x, y) = 2x^3 + 6xy - 3y^2 + 2$$

1. f admet-elle des extrema locaux sur $\mathbb{R}^2$? Si oui, les déterminer.
2. f admet-elle des extrema globaux sur $\mathbb{R}^2$? Justifier.
3. On pose $K = [0, 1] \times [0, 1]$. Justifier que f admet un maximum global sur K puis le déterminer.

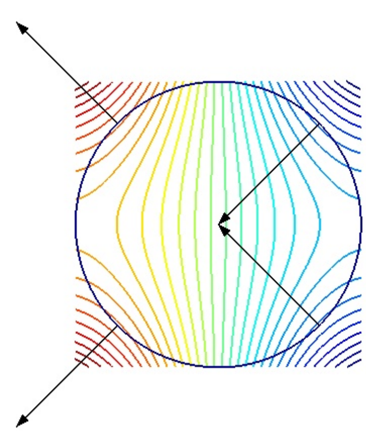
3.2 méthode de substitution

Exercice 7 : Maximiser et minimiser $f : (x, y) \mapsto x^3 - 3x(1 + y^2)$ sous la contrainte $x^2 + y^2 - 1 = 0$, par la méthode de substitution.

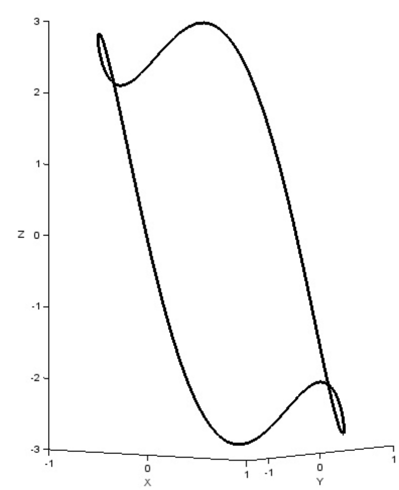
$f(x,y) = x^3 - 3x(1+y^2)$



contrainte $x^2 + y^2 = 1$



restriction de f à la contrainte



Exercice 8 :
On considère une boîte parallélépipédique de côtés x, y et z (en centimètres). Par la méthode de substitution, maximiser le volume de la boîte sous la contrainte : la longueur totale des arêtes vaut 96 cm.

> Comme on ne peut pas toujours exprimer une variable en fonction des autres, nous allons avoir recours à la méthode de Lagrange, ou méthode du Lagrangien, pour déterminer une liste de points « candidats » à être extrema de f sous la contrainte.

3.3 méthode de Lagrange

Soit $f : \mathcal{U} \subset E \to \mathbb{R}$ et $g : \mathcal{U} \to \mathbb{R}$ deux fonctions de classe $\mathcal{C}^1$. On cherche à étudier les extrema de f sous la contrainte $g(x) = 0$. Dans la propriété suivante, X n'est rien d'autre que la partie de E constituée des points satisfaisant la contrainte. Dans les deux exercices précédents, X était respectivement un cercle, un plan.

Théorème 2 – Théorème de Lagrange

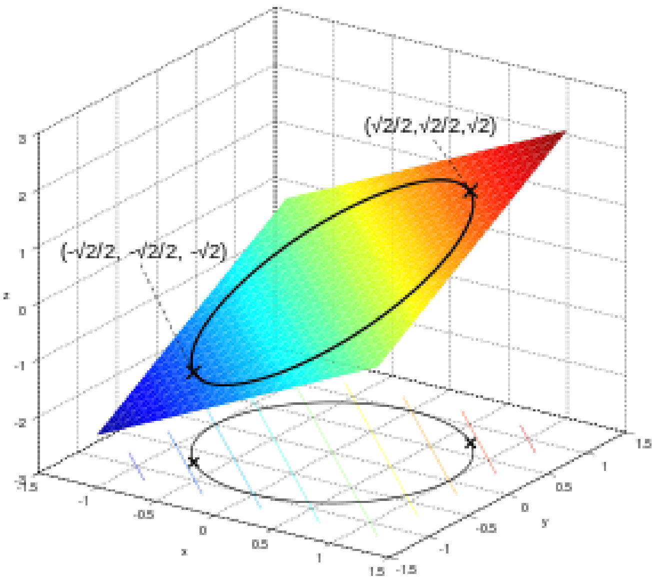
Soit X l'ensemble des zéros de g et $a \in X$. On suppose que $\mathrm{d}g(a) \neq 0$.
Si la restriction de f à X admet un extremum local en a, alors :

- $\mathrm{d}f(a)$ est nulle sur T_aX,
- $\mathrm{d}f(a)$ est colinéaire à $\mathrm{d}g(a)$.

> Si E est euclidien, la formulation de ce théorème revient à : soit X l'ensemble des zéros de g et $a \in X$. On suppose que $\nabla g(a) \neq 0$.
> Si la restriction de f à X admet un extremum local en a, alors $\nabla f(a)$ est colinéaire à $\nabla g(a)$.

Exercice 9 : Dans $E = \mathbb{R}^2$, on considère $f(x,y) = x + y$, sous la contrainte $x^2 + y^2 = 1$ et on cherche à optimiser f sous cette contrainte.

1. Montrer que f admet un minimum et un maximum sous la contrainte.
2. Déterminer, par la méthode de Lagrange, la liste des points candidats à être extrêmum de f sous la contrainte.
3. Conclure, et vérifier vos résultats sur la représentation graphique donnée ci-dessous.



Exercice 10 : $f(x,y,z) = x + yz$ et $\mathcal{S} = \{(x,y,z) \in \mathbb{R}^3,\ x^2 + y^2 + z^2 = 1\}$.

1. Montrer que f admet un minimum et un maximum sur $\mathcal{S}$.
2. Déterminer les points candidats à être extrêmum de f sous la contrainte.
3. Donner le maximum et le minimum de f sous la contrainte.

4 Annexe : quelques éléments de démonstrations

Propriété 1

• Supposons que f admette un extremum local en a. Soit $(e_1, \dots, e_n)$ une base de E.
Comme $\mathcal{U}$ est ouvert, la fonction de la variable réelle $g : t \mapsto f(a + te_i)$ est définie sur un voisinage de 0, de classe $\mathcal{C}^1$ et admet un extremum en 0. On a donc $g'(0) = 0$. Donc $\frac{\partial f}{\partial x_i}(a) = 0$.
Ainsi $\mathrm{d}f(a) = 0$.

• Soit $f(x, y) = x^2 - y^2$. $\nabla f(x, y) = (2x, 2y)$ et le seul point critique de f est $a = (0, 0)$.
$f(x, 0) - f(0, 0) = x^2$ et $f(0, y) - f(0, 0) = -y^2$.

Propriété 2

Soit f est une fonction de classe $\mathcal{C}^2$ sur un ouvert de $\mathbb{R}^n$. $H_f(a)$ est une matrice symétrique réelle. Nous allons utiliser le théorème de caractérisation spectrale,

$$H_f(a) \in \mathcal{S}_n^+(\mathbb{R}) \Leftrightarrow \forall h \in \mathbb{R}^n,\ h^\top H_f(a) h \geqslant 0 \Leftrightarrow \mathrm{Sp}(H_f(a)) \subset \mathbb{R}^+$$

$$H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R}) \Leftrightarrow \forall h \in \mathbb{R}^n \setminus \{0\},\ h^\top H_f(a) h > 0 \Leftrightarrow \mathrm{Sp}(H_f(a)) \subset \mathbb{R}^{++}$$

et la formule de Taylor-Young à l'ordre 2 en a :

$$f(a + h) - f(a) = \langle \nabla f(a) | h \rangle + \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2)$$

• Supposons que f admette un minimum local en a.
On a vu plus haut que $\nabla f(a) = 0$. On a

$$f(a + h) - f(a) = \frac{1}{2} h^\top H_f(a) h + o(\|h\|^2)$$

Soit λ une valeur propre de $H_f(a)$ et h un vecteur propre de norme 1 associé. On a pour t réel au voisinage de 0 :

$$f(a + th) - f(a) \underset{t \to 0}{=} \frac{t^2}{2} h^\top H_f(a) h + o(t^2 \|h\|^2)$$

Or $h^\top H_f(a) h = h^\top \lambda h = \lambda$, donc $f(a + th) - f(a) \underset{t \to 0}{=} \frac{t^2}{2} \lambda + o(t^2)$.
Par hypothèse de minimum en a, $\frac{t^2}{2} \lambda + o(t^2) \geqslant 0$, puis $\frac{\lambda}{2} + o(1) \geqslant 0$, et enfin $\lambda \geqslant 0$.
Donc $\mathrm{Sp}(H_f(a)) \subset \mathbb{R}^+$ et donc $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$.

• Supposons que a est point critique de f et que $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$. Il existe une fonction de limite 0 en 0 telle que

$$f(a + h) - f(a) \underset{h \to 0}{=} \frac{1}{2} h^\top H_f(a) h + \varepsilon(h) \|h\|^2$$

Notons $\lambda_{\min}$ la plus petite valeur propre de $H_f(a)$. Par la caractérisation spectrale, $\lambda_{\min} > 0$. Par le théorème spectral, il existe une matrice orthogonale P et une matrice D diagonale de diagonale comportant les valeurs propres de $H_f(a)$ telles que

$$\begin{aligned} h^\top H_f(a) h &= h^\top P D P^\top h = y^\top D y \text{ où } y = P^\top h \\ &= \sum_{i=1}^n \lambda_i y_i^2 \geqslant \lambda_{\min} \underbrace{\sum_{i=1}^n y_i^2}_{= \|y\|^2 = \|h\|^2} \end{aligned}$$

Comme $\lim\limits_{h \to 0} \varepsilon(h) = 0$, il existe $r > 0$ tel que $\|h\| < r \Rightarrow \varepsilon(h) > -\frac{\lambda_{\min}}{4}$.
Pour $h \in B(0, r)$, on a $\varepsilon(h) \|h\|^2 \geqslant -\frac{\lambda_{\min}}{4} \|h\|^2$, et pour $h \neq 0$,

$$f(a + h) - f(a) \geqslant \frac{\lambda_{\min}}{4} \|h\|^2 > 0$$

et f atteint un minimum local strict en a.

• Pour adapter les résultats au cas d'un maximum, on utilise : f admet un maximum si et seulement si $-f$ admet un minimum.

Propriété 3

Soit f une fonction réelle de classe $\mathcal{C}^2$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^2$ et soit a un point critique de f. On sait que la trace et le déterminant d'une matrice sont la somme et le produit des valeurs propres complexes comptées avec multiplicité. Ici $H_f(a)$ est symétrique réelle, donc n'admet que des valeurs propres réelles. Si on note λ, μ ses valeurs propres (éventuellement $\lambda = \mu$) alors $\begin{cases} \lambda + \mu = \mathrm{Tr}(H_f(a)) \\ \lambda \mu = \det(H_f(a)) \end{cases}$.

- Si $\det(H_f(a)) > 0$, alors λ et μ ont même signe (et sont non nulles). Par la condition suffisante de la propriété précédente, f admet un extremum local en a. Il s'agit d'un minimum local si les valeurs propres sont positives, ce qui est le cas si et seulement si $\text{Tr}(H_f(a)) > 0$.
- Si $\det(H_f(a)) < 0$, alors les valeurs propres sont de signe contraire, et par la condition nécessaire de la propriété précédente, f n'admet pas d'extremum en a.

  Soit h un vecteur propre de norme 1 associé à la valeur propre λ. Comme rédigé au cours de la démonstration précédente :
  $$f(a+th) - f(a) \underset{t\to 0}{\sim} \frac{\lambda}{2}t^2$$

  > Donc une valeur propre strictement négative conduit à un maximum dans la direction d'un vecteur propre associé, et une valeur propre strictement positive conduit à un minimum dans la direction d'un vecteur propre associé.

  Ici f admet un point col en a.

- Si $\det(H_f(a)) = 0$, on ne peut pas conclure par cette approche.

Théorème 2 – Théorème de Lagrange

Soit X l'ensemble des zéros de g et $a \in X$. On suppose que $\mathrm{d}g(a) \neq 0$. On va montrer le théorème suivant, énoncé ainsi dans le programme :
Si la restriction de f à X admet un extremum local en a, alors :

— $\mathrm{d}f(a)$ est nulle sur T_aX,

— $\mathrm{d}f(a)$ est colinéaire à $\mathrm{d}g(a)$.

On suppose donc que la restriction de f à X admet un extremum local en a.
• Montrons que $\mathrm{d}f(a)$ est nulle sur T_aX.

Soit $x \in T_aX$. Il existe $\gamma :]-\varepsilon, \varepsilon[\to X$, dérivable en 0, telle que $\begin{cases} \gamma(0) & = a \\ \gamma'(0) & = x \end{cases}$

$f_{|X}$ admet un extremum en a, donc $f \circ \gamma$ admet un extremum local en 0, et sa dérivée en 0 vaut 0. Donc
$$\mathrm{d}f(\gamma(0))(\gamma'(0)) = 0 \text{ soit } \mathrm{d}f(a)(x) = 0$$
Ainsi $\mathrm{d}f(a)$ est nulle sur T_aX.

• Au chapitre différentiel, on a vu que, puisque $\mathrm{d}g(a) \neq 0$, $T_aX = \ker \mathrm{d}g(a)$.
On vient donc de montrer l'inclusion : $\ker \mathrm{d}g(a) \subset \ker \mathrm{d}f(a)$.
Si $\mathrm{d}f(a) = 0$, alors il existe $\lambda = 0$ tel que $\mathrm{d}f(a) = \lambda \mathrm{d}g(a)$.
Si $\mathrm{d}f(a) \neq 0$, alors $\mathrm{d}f(a)$ est une forme linéaire non nulle, et son noyau est de dimension $n-1$, et donc $\dim \ker \mathrm{d}f(a) = \dim \ker \mathrm{d}g(a)$. On a donc $\ker \mathrm{d}g(a) = \ker \mathrm{d}f(a)$. On termine avec le lemme d'algèbre :

Lemme 1

Deux formes linéaires non nulles φ et ψ ont même noyau si et seulement s'il existe $\lambda \in \mathbb{R}$ tel que $\varphi = \lambda\psi$.

Démonstration du lemme :
($\Leftarrow$) facile et ne nous intéresse pas dans notre démonstration globale.
($\Rightarrow$) Si φ et ψ ont même noyau, comme elles sont non nulles, le noyau est de dimension $n-1$ et il existe $a \in E$ tel que $\text{Vect}(a)$ soit supplémentaire dans E de $\ker \phi = \ker \psi$. On pose $\lambda = \frac{\varphi(a)}{\psi(a)}$ (existe car si $\psi(a) = 0$, ψ est nulle sur E et c'est exclu). On vérifie que $\varphi = \lambda\psi$.