# Note de cours CM 1 MIA

Marius THORRE, César BLANC, Alexandre AGUEDO

16 septembre 2024

## 1 Optimisation

Soit $f : D = \text{dom}(f) \subset R^n \to R$, l'objectif est de trouver le minimum de la fonction $f$ sur un domaine. Pour cela, il faut au préalable connaître la nature du domaine :

- Un domaine est dit **ouvert** si c'est un ensemble de points dans un espace tel que, pour chaque point de cet ensemble, il existe un voisinage entièrement contenu dans l'ensemble.
- Un domaine est dit **fermé** si son complémentaire est ouvert. Cela signifie qu'il inclut ses frontières. Donc, si par exemple $x$ est à la frontière du domaine, il n'existe pas de région autour de $x$ qui soit entièrement dans le domaine.
- Est-ce que le domaine et la fonction sont convexes ?
- Voici un exemple de domaines convexe et non convexe :

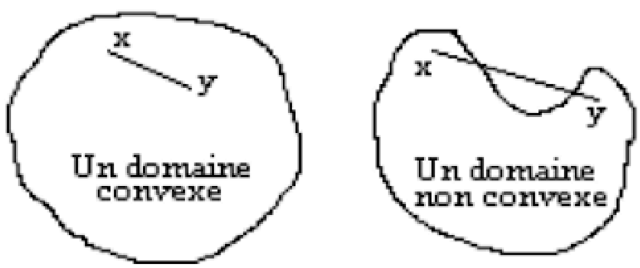



FIGURE 1 – Représentation de domaines convexe et non convexe

- La fonction $f$ est dite **convexe** sur un intervalle $I$, si sa courbe représentative est entièrement située au-dessus de chacune de ses tangentes.
- La fonction $f$ est dite **concave** sur un intervalle $I$, si sa courbe représentative est entièrement située en dessous de chacune de ses tangentes.
- Voici un exemple de ces fonctions à la figure 2 :

Une fonction $f$ est dite lisse si elle est dérivable pour chaque terme de son domaine. Exemple d'une fonction non lisse à la figure 3 :

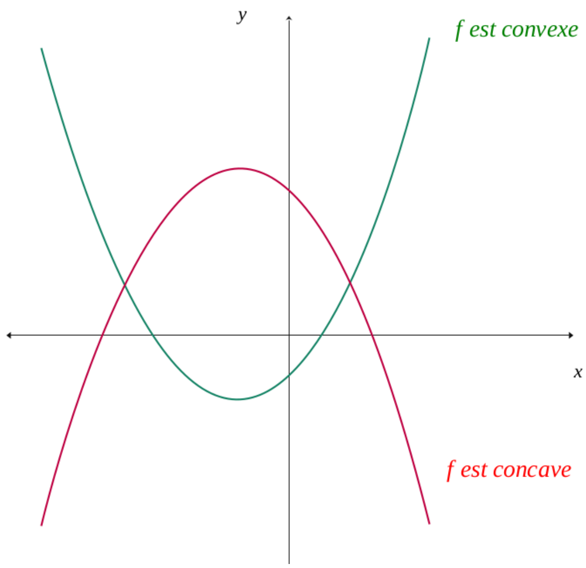



Figure 2 – Représentation de fonctions convexe et concave

## 1.1 Définition

$x^*$ est un minimum global de $f$ si et seulement si, pour tout $x \in \text{dom}(f)$, $f(x^*) \leq f(x)$. Cela signifie que la valeur de $f$ en $x^*$ est la plus petite ou égale à toutes les autres valeurs que prend f dans son domaine.

$x^*$ est un minimum local de $f$ si et seulement s'il existe un ensemble ouvert $U \subset \text{dom}(f)$ contenant $x^*$ tel que, pour tout $x \in U$, $f(x^*) \leq f(x)$. Cela signifie que la valeur de $f$ en $x^*$ est la plus petite ou égale à toutes les autres valeurs dans une région du domaine, mais il peut exister d'autres points en dehors de cette région où la fonction prend une valeur plus petite.

## 1.2 Topologie

Dans $R^n$, la boule fermée $B_f(x, \varepsilon)$, centrée sur $x$ et de rayon $\varepsilon$ associée à la norme euclidienne est l'ensemble des points $y \in R^n$ tels que

$$\|y - x\|_2 \leq \varepsilon.$$

Cela inclut donc tous les points à l'intérieur de la boule ainsi que ceux sur la frontière de la boule.

Dans $R^n$, la boule ouverte $B_o(x, \varepsilon)$, centrée en $x$ et de rayon $\varepsilon$ associée à la norme euclidienne est l'ensemble des points $y \in R^n$ tels que

$$\|y - x\|_2 < \varepsilon.$$

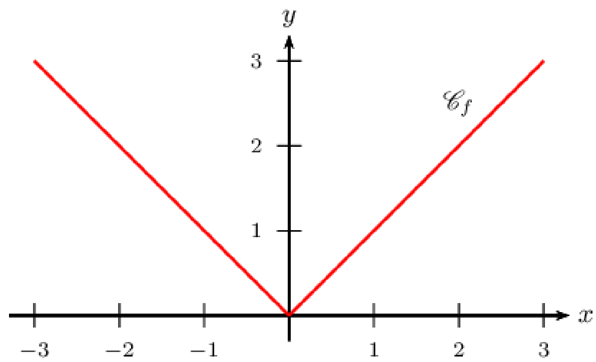



FIGURE 3 – Représentation de la fonction valeur absolue dans l'intervalle $[-3, 3]$

Cela représente uniquement les points strictement à l'intérieur de la boule, sans inclure la frontière.

Un ensemble $U$ est un ouvert de $R^n$ si et seulement si $U \subseteq R^n$ et pour tout $x \in U$, il existe $\varepsilon > 0$ tel que

$$B_f(x, \varepsilon) \subset U.$$

Un ensemble $U$ est donc un fermé de $R^n$ si et seulement si son complément dans $R^n$ est un ouvert de $R^n$.

Exemples sur l'ensemble des réels : l'intervalle $]-1, 3[$ est une boule ouverte de centre 1 et de rayon 2. L'intervalle $[0; +\infty[$ est un fermé pour la bonne raison que son complémentaire $]-\infty, 0[$ est un ouvert.

Dans $R^2$, une boule ouverte est un ensemble de points défini par une inéquation de type $(x-a)^2 + (y-b)^2 < r^2$. Si l'inégalité est large, la boule est fermée.

Pour plus d'informations sur les boules et voisinages, consultez le site suivant : [1]

1. `http://www.jybaudot.fr/Analysesup/voisinage.html`.

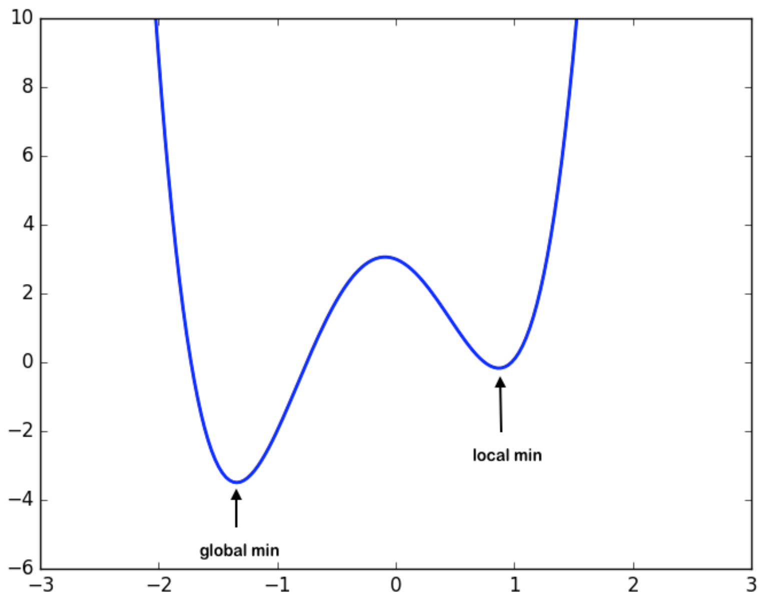



FIGURE 4 – Représentation de minimum local et global dans une fonction $f$

## 1.3 Matrice Jacobienne

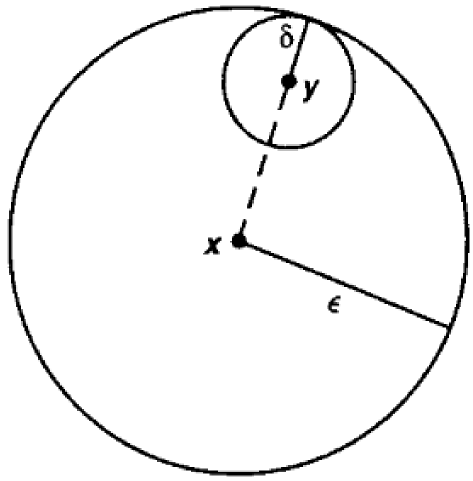
Soit $f : U \subset R^n \to R^m$, de classe $C^1$.
$f$ prend des vecteurs de dimensions $n$ et les transforme en vecteurs de dimension $m$. $f$ est de classe $C^1$, ce qui signifie que ses dérivées partielles existent et sont continues.

Définissons la notion de dérivée partielle d'une fonction :
$f : x \mapsto (f_1(x), f_2(x), \ldots, f_m(x))$
La dérivée partielle de $f_i$ par rapport à $x_j$ est donnée par :

$$\frac{\partial f_i}{\partial x_j}(a_1, \ldots, a_n) \mapsto \lim_{h \to 0} \frac{f_i(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - f_i(a_1, \ldots, a_n)}{h} \tag{1}$$

La matrice Jacobienne est une matrice contenant toutes les dérivées partielles de chaque composante de la fonction $f$, elle est alors :

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$




Si $m = 1$, c'est-à-dire $f : R^n \to R$, donc fonction scalaire, la Jacobienne est un vecteur ligne qui correspond au gradient de la fonction. Ce gradient est la transposée de la matrice Jacobienne dans ce cas.

## 1.4 Chain Rule Jacobienne

Lorsque deux fonctions sont composées, soit $f \circ g$, la Jacobienne de la composition $f \circ g$ est donnée par le produit des Jacobiennes de $f$ et $g$.

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

On doit donc calculer la Jacobienne de $g$, puis l'évaluer en $a$, et ensuite multiplier cela par la Jacobienne de $f$ évaluée en $g(a)$.

## 1.5 Matrice Hessienne

Soit $f : U \subset R^n \to R$. une fonction de classe $C^2$, c'est-à-dire une fonction deux fois continûment différentiable. L'hessienne de $f$ est la matrice carrée $n * n$ qui contient toutes les dérivées secondes partielles de la fonction $f$ et est définie par :

$$\nabla^2 f = H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

## 1.6 Conditions d'optimalité

**Conditions nécessaires, premier ordre**

Si $x^*$ est un minimum local et $f$ est de classe $C^1$ sur un ouvert $U \subset R^n$ contenant $x^*$, alors $\nabla f(x^*) = 0$.

**Conditions nécessaires, second ordre**

Si $x^*$ est un minimum local et $f$ est de classe $C^2$ sur un ouvert $U \subset R^n$ contenant $x^*$, alors :

— $\nabla f(x^*) = 0$

— $\nabla^2 f(x^*)$ est semi-définie positive.

Une matrice Hessienne semi-définie positive signifie que $f$ présente soit une courbure nulle, soit une courbure convexe autour de $x^*$.

**Conditions suffisantes, second ordre**

Si $f$ est de classe $C^2$ sur un ouvert $U \subset R^n$ contenant $x^*$, et si $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est définie positive, alors $x^*$ est un minimum local.

Une matrice Hessienne définie positive signifie que $f$ présente une courbure convexe autour de $x^*$.

## 1.7 Existence de minimum

**Théorème de Weierstrass**

Si $f$ est définie et continue sur un sous-ensemble fermé et borné de $R^n$, alors $f$ admet un minimum global.

**Théorème**

Si $f$ est définie et continue sur $R^n$ et coercive (i.e. $f(x) \to +\infty$ quand $\|x\| \to +\infty$. Autrement dit, plus on s'éloigne de l'origine (plus $\|x\|$ augmente), plus la valeur de la fonction $f(x)$ augmente.), alors $f$ admet un minimum global sur tout sous-ensemble fermé de $R^n$.

**Théorème**

Si $f$ est convexe et minorée, alors $f$ admet un minimum global.

## 1.8 Convexité

L'ensemble $X$ est convexe si et seulement si pour tout $x, y$ dans $X$, le segment $[x, y] := \{tx + (1-t)y \mid t \in [0, 1]\}$ est inclus dans $X$.

Ce qui signifie que le segment qui relie $x$ à $y$ est entièrement contenu dans $X$.

Une fonction $f : X \subset R^n \to R$ est convexe si et seulement si $X$ est convexe et pour tout $x, y$ dans $X$ et pour tout $t \in [0, 1]$,

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

Cela signifie que la valeur de la fonction en un point sur le segment entre $x$ et $y$ est toujours inférieure ou égale à la moyenne pondérée des valeurs de $f$ en $x$ et $y$, ce qui donne à la fonction convexe une forme de bol.

Si l'inégalité est stricte pour tout $t \in ]0, 1[$, la fonction est dite *strictement convexe*, donc aucune parties "plates".

Quelques exemples de fonctions convexes :
- les fonctions linéaires ;
- les combinaisons linéaires positives de fonctions convexes ;
- le maximum (point par point) d'un ensemble de fonctions convexes ;
- ...

Si $f$ est de classe $C^2$, $f$ est convexe si et seulement si $\nabla^2 f$ est semi-définie positive sur l'intérieur de $X$, donc que toutes les valeurs propres de l'Hessienne sont soit positives, soit nulles.

Si $\nabla^2 f$ est définie positive sur l'intérieur de $X$, donc que toutes ses valeurs propres sont strictement positives, alors $f$ est strictement convexe (avec $X = \text{dom } f$).

## 1.9 Unicité du minimum

Si $f$ est une fonction strictement convexe définie sur un ensemble convexe $X \subset R^n$, alors $f$ admet au plus un minimum global. Cela s'explique par le fait que la fonction a une forme telle qu'elle descend vers un seul point minimum, puis remonte immédiatement après, il ne peut y avoir qu'un seul point où $f(x)$ est minimal.

Si $f$ est une fonction convexe définie sur un ensemble convexe $X \subset R^n$, alors tout minimum local de $f$ est aussi un minimum global de $f$. Si $X$ est ouvert, alors $\nabla f(x^*) = 0$ si et seulement si $x^*$ est un minimum global de $f$. Cela s'explique par le fait que lorsqu'un point $x^*$ où $\nabla f(x^*) = 0$, ce point ne peut pas être un minimum local sans être aussi un minimum global. Dans une fonction convexe, toute descente vers un point bas est nécessairement la descente vers le minimum global. Il n'y a pas de "vallées" cachées ou de points plus bas que ce minimum local.