Préparation à l'agrégation externe de Sciences Sociales

Compléments sur l'optimisation des fonctions de plusieurs variables

2021-2022

1 Gradient

Le gradient d'une fonction de plusieurs variables en un certain point est un vecteur qui caractérise la variabilité de cette fonction au voisinage de ce point. Le gradient est la généralisation à plusieurs variables de la dérivée d'une fonction d'une seule variable.

Définition 1. *Soit f une fonction qui admet des dérivées partielles par rapport à toutes ses variables en $x = (x_1, x_2, \ldots, x_n)$. Le gradient de f en x est défini comme le vecteur*

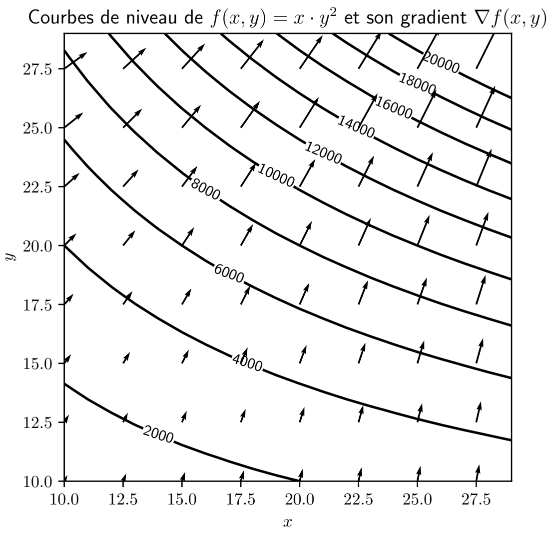
$$\nabla f(x_1, x_2, \ldots, x_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Le vecteur gradient pointe dans la direction où la fonction croît le plus rapidement, et son module est égal au taux de croissance dans cette direction. Dans une représentation graphique, le gradient est orthogonal à la courbe de niveau de la fonction au même point.

Exemple Soit la fonction $f(x, y) = x \cdot y^2$ définie sur $\mathbb{R}^2$. Le gradient de f est

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} y^2 \\ 2xy \end{pmatrix}$$

On le note parfois aussi $\nabla f(x, y) = (y^2, 2xy)$.



2 Fonctions convexes

Définition 2. *On dit qu'une partie A de $\mathbb{R}^n$ est convexe si et seulement si pour tout $x \in A$, pour tout $y \in A$ et pour tout $\lambda \in [0,1]$, alors $\lambda x + (1-\lambda)y \in A$.*

Autrement dit, si x et y appartiennent à A, alors le segment de droite les reliant appartient aussi à A.

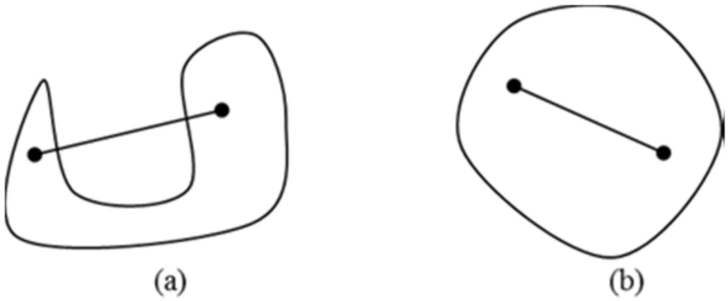


Figure 1: **(a)** ensemble non convexe. **(b)** ensemble convexe.

Définition 3. *Soit A une partie convexe de $\mathbb{R}^n$ et soit $f : A \to \mathbb{R}$ une fonction réelle. On dit que f est convexe si et seulement si pour tout $x \in A$, pour tout $y \in A$ et pour tout $\lambda \in [0,1]$, alors*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

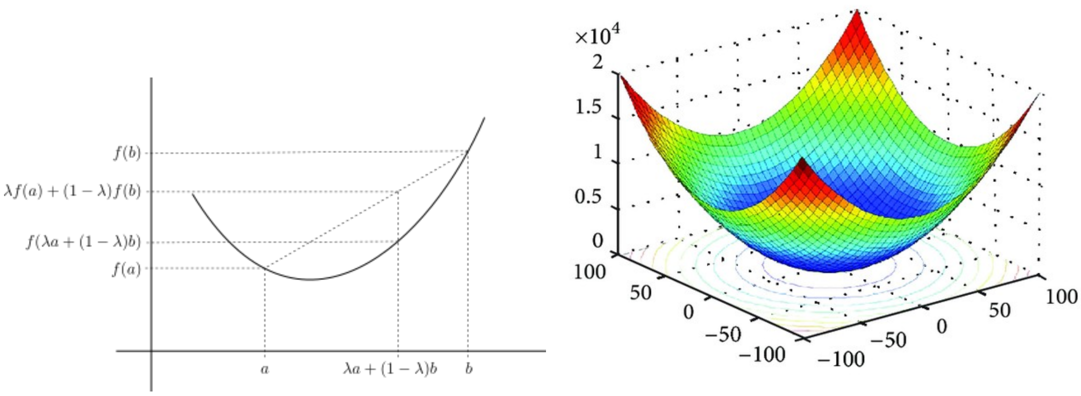


Figure 2: Fonctions convexes.

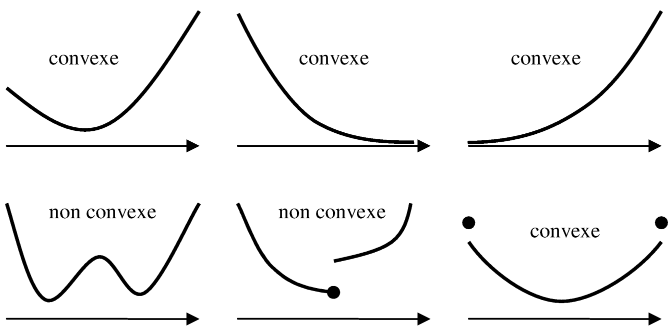


Figure 3: Fonctions convexes et non-convexes.

Définition 4. *Soit A une partie convexe de $\mathbb{R}^n$ et soit $f : A \to \mathbb{R}$ une fonction réelle. On dit que f est concave si $-f$ est convexe.*

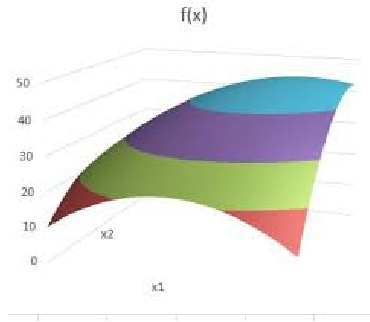


Figure 4: Fonction concave.

Théorème 1. *Soit $A \in \mathbb{R}^2$ un ouvert convexe et soit $f \in \mathscr{C}^2$ une fonction réelle avec matrice hessienne*

$$Hess_f(x, y) = \begin{pmatrix} r(x, y) & s(x, y) \\ s(x, y) & t(x, y) \end{pmatrix}$$

Alors,

- *f est convexe sur A si et seulement si*

$$r(x, y)t(x, y) - s^2(x, y) \geq 0 \qquad \forall (x, y) \in A$$

et

$$r(x, y) \geq 0 \qquad \forall (x, y) \in A;$$

- *f est concave sur A si et seulement si*

$$r(x, y)t(x, y) - s^2(x, y) \geq 0 \qquad \forall (x, y) \in A$$

et

$$r(x, y) \leq 0 \qquad \forall (x, y) \in A.$$

Théorème 2. *Soit $A \in \mathbb{R}^n$ un ouvert convexe et $f : A \to \mathbb{R}$ une fonction convexe. Alors, $a \in A$ est un minimum local de f si et seulement si a est un minimum global de f sur A.*

Théorème 3. *Soit $A \in \mathbb{R}^n$ un ouvert convexe et $f : A \to \mathbb{R}$ une fonction concave. Alors, $a \in A$ est un maximum local de f si et seulement si a est un maximum global de f sur A.*

Exemple La fonction f définie sur $\mathbb{R}^2$ par

$$f(x, y) = x^4 + 4x^2 + 4xy + 4y^2 - 36$$

admet des dérivées partielles premières et secondes par rapport à x et à y en tout point de $\mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 + 8x + 4y \qquad \frac{\partial f}{\partial y}(x, y) = 4x + 8y$$

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 12x^2 + 8 \qquad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 4$$

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = 4 \qquad \frac{\partial^2 f}{\partial y^2}(x, y) = 8$$

et
$$\mathrm{Hess}_f(x,y) = \begin{pmatrix} 12x^2+8 & 4 \\ 4 & 8 \end{pmatrix}$$
Comme,
$$r(x,y)t(x,y) - s^2(x,y) = (12x^2+8)\cdot 8 - 4\cdot 4 \;\geq\; 8\cdot 8 - 4\cdot 4 \;>\; 0$$
et
$$r(x,y) = 12x^2 + 8 \;\geq\; 8 \;>\; 0$$
on peut conclure que f est convexe d'après le théorème 1.

Les points critiques de f sont déterminés par
$$4x^3 + 8x + 4y = 0 \qquad et \qquad 4x + 8y = 0$$
La deuxième équation nous dit qu'au point critique $y = -\frac{x}{2}$ et en remplaçant dans la première équation on obtient
$$4x^3 + 8x - 2x = 4x^3 + 6x = 2x(x^2+3) = 0$$
d'où $x = 0$ et $y = 0$. Le seul point critique est donc $(0,0)$. Comme $r(0,0)t(0,0) - s^2(0,0) = 8^2 - 4^2 > 0$ et $r(0,0) = 8 > 0$ on peut déduire que $(0,0)$ est un minimum local de f.

Comme f est convexe, par le théorème 2, $(0,0)$ est un minimum global de f.

3 Optimisation sur un domaine fermé

Les théorèmes précédents nous indiquent comment trouver des extrema sur des domaines ouverts. Quand on cherche des extrema sur des domaines fermés il faut séparer l'étude en deux parties: chercher les extrema dans le domaine ouvert où on a exclut la frontière, et chercher les extrema dans la frontière du domaine. Finalement, tous les extrema trouvés sont des candidats à évaluer pour trouver les extrema dans le domaine fermé.

Exemple On cherche les extrema de $f(x,y) = x^2 + y^2$ sur un domaine $A = \{(x,y) \in \mathbb{R}^2, 1 \leq x \leq 2,\ 0 \leq y \leq 1\}$.

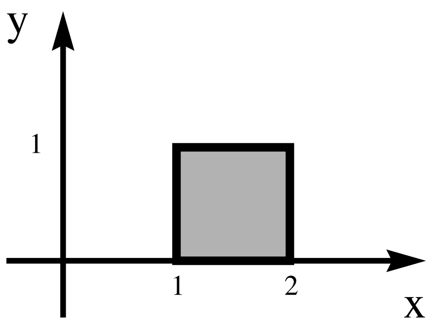


Figure 5: Optimisation sur un domaine fermé.

Comme $\frac{\partial f}{\partial x} = 2x$ et $\frac{\partial f}{\partial y} = 2y$, il n'y a pas de points critiques à l'intérieur du domaine A. Il nous reste à vérifier les frontières. On peut les séparer en quatre cas:

1. Quand $1 \leq x \leq 2$ et $y = 0$, il faut étudier la fonction $b_1(x) = f(x,0) = x^2$. Mais $b_1'(x) = 2x$ et n'a pas de point critique en $1 < x < 2$.

2. Quand $1 \leq x \leq 2$ et $y = 1$, il faut étudier la fonction $b_2(x) = f(x,1) = x^2 + 1$. Mais $b_2'(x) = 2x$ et n'a pas de point critique en $1 < x < 2$.

3. Quand $x = 1$ et $0 \leq y \leq 1$, il faut étudier la fonction $b_3(x) = f(1,y) = 1 + y^2$. Mais $b_3'(y) = 2y$ et n'a pas de point critique en $0 < y < 1$.

4. Quand $x = 2$ et $0 \leq y \leq 1$, il faut étudier la fonction $b_4(x) = f(2, y) = 4 + y^2$. Mais $b_4'(y) = 2y$ et n'a pas de point critique en $0 < y < 1$.

Finalement il nous faut évaluer les quatre points frontières des quatre cas précédents: $(1, 0)$, $(2, 0)$,$(1, 1)$ et $(2, 1)$. On a

$$f(1,0) = 1 \qquad f(2,0) = 4 \qquad f(1,1) = 2 \qquad f(2,1) = 5$$

et le minium de f en A est 1 en $(1, 0)$ et le maximum est 5 en $(2, 1)$.

Théorème 4. *Soit $A \in \mathbb{R}^n$ un ensemble fermé et borné. Soit $f : A \to \mathbb{R}$ une fonction continue. Alors, ils existent $m \in A$ et $M \in A$ tels que*

$$f(m) \leq f(x) \qquad \forall x \in A$$
$$f(M) \geq f(x) \qquad \forall x \in A$$

4 Optimisation sous contrainte

Exemple Soit $Y(K, L) = K \cdot L^2$ définie sur $K \geq 0$ et $L \geq 0$. Quel est le maximum de Y sous la contrainte $K + L = 30$?

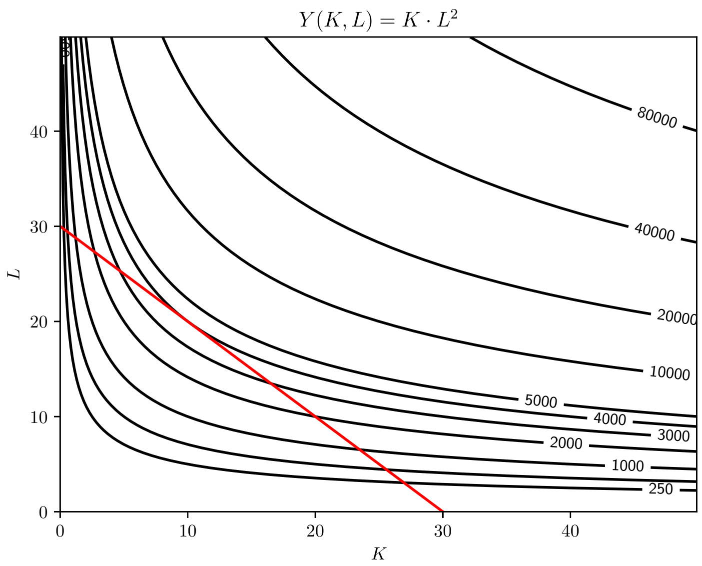


Figure 6: Optimisation sous contrainte.

Définition 5. *Soit A une partie de $\mathbb{R}^n$ et soit $f : A \to \mathbb{R}$ et $g : A \to \mathbb{R}$ deux fonctions réelles. Soit $S = \{x \in A, g(x) = 0\}$ l'ensemble de points de A où g est nulle.*

- *On dit que f admet un minimum local en $a \in A$ lié à la contrainte $g(x) = 0$ si*

$$g(a) = 0 \qquad (s \in S)$$

et

$$\exists\, r > 0, \quad f(x) \geq f(a), \quad \forall\, x \in S \cap B(a, r)$$

- *On dit que f admet un maximum local en $a \in A$ lié à la contrainte $g(x) = 0$ si*

$$g(a) = 0 \qquad (s \in S)$$

et

$$\exists\, r > 0, \quad f(x) \leq f(a), \quad \forall\, x \in S \cap B(a, r)$$

4.1 Méthode de substitution

Supposons qu'on cherche un extremum de $f(x,y)$ en $\mathbb{R}^2$ lié à la contrainte $g(x,y) = 0$. Supposons qu'on peut trouver une fonction h_g tel que $y = h_g(x)$ satisfait

$$g(x, h_g(x)) = 0 \qquad \forall x \in \mathbb{R}.$$

Alors, on peut remplacer h_g dans f et définir la fonction F

$$F(x) = f(x, h_g(x)).$$

L'extremum de F (optimisation sans contrainte) est aussi la solution au problème original.

Exemple Soit $Y(K,L) = K \cdot L^2$ définie sur $K \geq 0$ et $L \geq 0$. Quel est le maximum de Y sous la contrainte $K + L = 30$?

Dans ce cas, la contrainte est donné par $g(K,L) = K + L - 30 = 0$. Mais $K = h_g(L) = 30 - L$. En remplaçant dans Y on obtient

$$F(L) = (30 - L)L^2 \qquad \text{pour} \qquad 0 \leq L \leq 30$$

La dérivée de F est

$$F'(L) = 2(30 - L)L - L^2 = 60L - 2L^2 - L^2 = L(60 - 3L)$$

Il y a deux points critiques $F'(L) = 0$, $L = 0$ et $L = \frac{60}{3} = 20$, mais seulement $L = 20$ est à l'intérieur du domaine de définition. On a donc trois candidats à vérifier, le point critique et les deux points frontières du domaine $L = 0$ et $L = 30$. On a

$$F(0) = 0 \qquad F(20) = 10 \cdot 20^2 = 4000 \qquad F(30) = 0$$

On peut conclure que le maximum de Y sous la contrainte $K + L = 30$ s'obtient pour $L = 20$ et $K = 30 - L = 10$ et vaut $Y(10, 20) = 4000$.

4.2 Multiplicateurs de Lagrange

Théorème 5. *Soit A une partie de $\mathbb{R}^n$ et soit $f : A \to \mathbb{R}$ et $g : A \to \mathbb{R}$ deux fonctions dérivables. Soit $S = \{x \in A, g(x) = 0\}$ l'ensemble de points de A où g est nulle. Supposons que $\nabla f(x) \neq 0$ pour tout $x \in S$. Si f admet un extremum en $a \in S$ lié à la contrainte $g(x) = 0$, alors*

$$\nabla f(a) = \lambda \nabla g(a)$$

où $\lambda \in \mathbb{R}$ est appelé le multiplicateur de Lagrange.

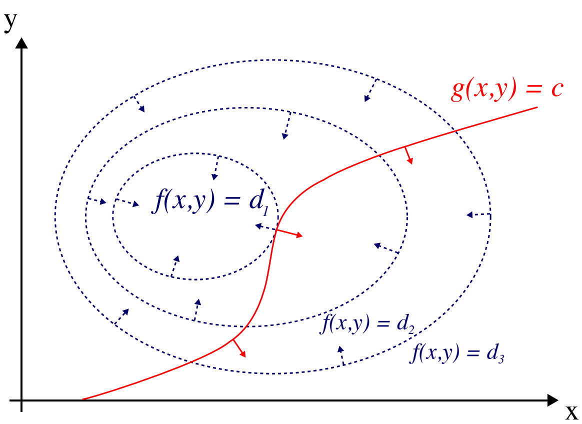


Figure 7: Dans un extremum de f lié à la contrainte $g(x,y) = 0$ on a $\nabla f(a) = \lambda \nabla g(a)$.

Le théorème 5 donne une condition nécessaire mais pas suffisante. Aussi, s'il y a un extremum, on ne sait pas a priori s'il s'agit d'un minimum ou d'un maximum.

Méthode des multiplicateurs de Lagrange ou du Lagrangien Supposons qu'on cherche un extremum de $f(x, y)$ en $\mathbb{R}^2$ lié à la contrainte $g(x, y) = 0$. On appelle Lagrangien la fonction définie sur (x, y) et λ par

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

On cherche des points critiques de L,

$$\nabla L(x, y, \lambda) = 0$$

Autrement dit, on cherche des solutions du système suivant:

$$\begin{cases} \frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial L}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \\ \frac{\partial L}{\partial \lambda} = -g(x, y) = 0 \end{cases}$$

Les deux premières équations correspondent à imposer $\nabla f(x, y) = \lambda \nabla g(x, y)$. La troisième équation impose la contrainte $g(x, y) = 0$. Les extrema de f sous la contrainte g sont à chercher parmi les solutions du système d'équation précédente.

Exemple On cherche le maximum de la fonction $f(x, y) = x \cdot y^2$ pour $x \geq 0$ et $y \geq 0$ lié à la contrainte $x + y = k$ pour $k > 0$.

Dans ces conditions la fonction g est $g(x, y) = x + y - k = 0$. On définie le Lagrangien

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y) = xy^2 - \lambda(x + y - k)$$

La condition $\nabla L(x, y, \lambda) = 0$ nous amène au système suivant:

$$\begin{cases} \frac{\partial L}{\partial x} = y^2 - \lambda = 0 \\ \frac{\partial L}{\partial y} = 2xy - \lambda = 0 \\ \frac{\partial L}{\partial \lambda} = -x - y + k = 0 \end{cases}$$

Les deux premières équations nous dissent que $y^2 = 2xy$. La troisième nous impose $x = k - y$. En remplaçant cela nous donne $y^2 = 2(k - y)y$, ou $y(3y - 2k) = 0$. Le système a donc deux solutions: $y = 0$ avec $x = k$ et $y = \frac{2}{3}k$ avec $x = \frac{k}{3}$. Il faut aussi essayer les frontières du domaine, donc les points $x = 0$, $y = k$ et $x = k$, $y = 0$ (déjà inclus avant). On a

$$f(0, k) = 0 \qquad f(k, 0) = 0 \qquad f\left(\frac{k}{3}, \frac{2}{3}k\right) = \frac{4}{27}k^3.$$

Comme $k > 0$ on peut conclure que le maximum cherché est $\frac{4}{27}k^3$ pour $x = k$ et $y = \frac{2}{3}k$.