Optimisation

Chapitre 0: Introduction

Optimisation classique

Master IASD

Tarik AMTOUT

tarik.amtout@gmail.com

Introduction

Soit $f(\boldsymbol{x}) = f(x_1, \ldots, x_n)$ une fonction différentiable de n variables à valeurs réelles. Nous désirons trouver $\boldsymbol{x}^*$ tel que $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$, pour tout $\boldsymbol{x}$. Dans ce cas, on dit que f possède un **minimum global** en $\boldsymbol{x}^*$. Les variables notées en caractères gras représentent des vecteurs à n composantes.

On dit que la fonction $f(\boldsymbol{x})$ possède un **minimum local** en $\boldsymbol{x}_0$ s'il existe un ε, $\varepsilon > 0$, tel que pour tout $\boldsymbol{x}$ au ε-voisinage de $\boldsymbol{x}_0$, $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.

La méthode d'optimisation classique nous permet de déterminer des minima locaux en utilisant les dérivées partielles de f. Du calcul différentiel, nous trouvons que si f est différentiable et possède un minimum local au point $\boldsymbol{x}_0$, alors $\boldsymbol{x}_0$ doit être une solution à l'ensemble des n équations :

$$\frac{\partial f(\boldsymbol{x})}{\partial x_j} = 0, \quad j = 1, \ldots, n$$

Exemple

Si $f(\boldsymbol{x})$ est une fonction de deux variables x_1 et x_2, le minimum local $\boldsymbol{x}_0$, s'il existe, est une solution des deux équations :

$$\frac{\partial f(\boldsymbol{x})}{\partial x_1} = 0 \quad et \quad \frac{\partial f(\boldsymbol{x})}{\partial x_2} = 0$$

Dans ce cas, $j = 1, 2$.

Le résultat précèdent est également vrai pour une définition analogue du maximum local. Cette méthode est connue sous le nom de méthode d'optimisation sans contrainte.

Lorsqu'il y a des contraintes, il faut résoudre le problème à l'aide **des multiplicateurs de Lagrange**. Supposons qu'il y ait un ensemble de contraintes de la forme: $b_i(\mathbf{x}) = b_i$, $i = 1, \ldots, m$

Nous pouvons alors écrire la fonction de Lagrange comme suit :

$$F(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i [b_i - g_i(\boldsymbol{x})]$$

où $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ est un vecteur de nombres réels.

Si $f(\boldsymbol{x})$ atteint un minimum local au point $\boldsymbol{x}_0$, alors il existe un vecteur $\boldsymbol{\lambda}$ tel que :

$$\frac{\partial f(\boldsymbol{x}_0)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\boldsymbol{x}_0)}{\partial x_j} = 0, \quad j = 1, \ldots, n$$

et

$$g_i(\boldsymbol{x}_0) = b_i, \quad i = 1, \ldots, m$$

L’optimisation classique se scinde en deux types de problèmes : **l’optimisation sans contrainte** et **l’optimisation avec contraintes**. Dans les deux cas, *le but consiste à trouver les valeurs qui maximisent ou minimisent une fonction*. Toutefois, dans l’optimisation avec contraintes, les solutions sont soumises à des restrictions (contraintes).

Optimisation classique sans contrainte

Soit f(x) une fonction d'une variable réelle.

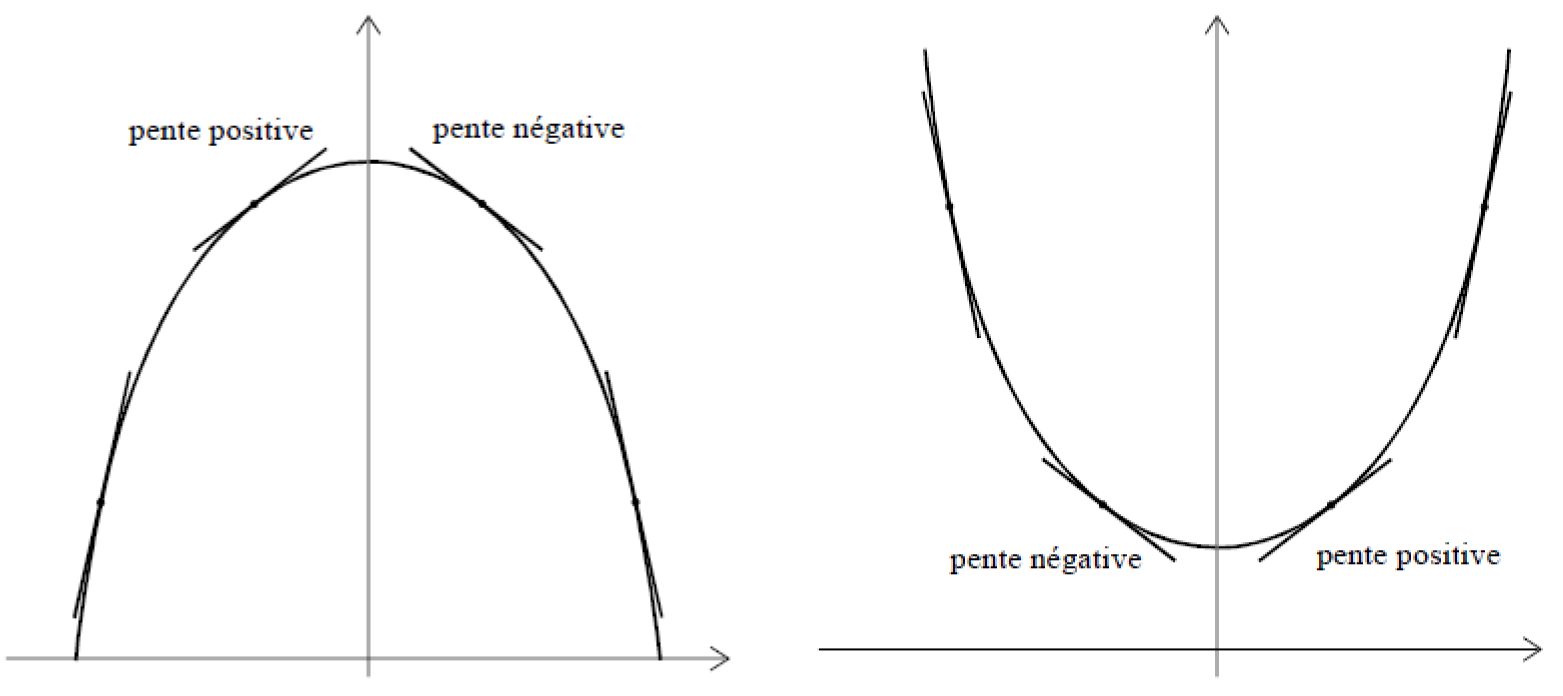
Si la fonction f(x) et sa dérivée f'(x) sont continues en un point où de croissante la fonction devient décroissante, alors elle possède un **maximum**. En d'autres termes, la pente de la tangente passe du positif au négatif.

Le raisonnement contraire est valable pour **un minimum**.

Dans les deux cas, il s'agit d'un point où la pente de la tangente est nulle. Comme la pente de la tangente est donnée par la première dérivée de la fonction, il faut annuler la première dérivée de f pour obtenir **les points candidats** (points où la fonction est susceptible de présenter un minimum ou un maximum). Pour chaque point candidat, il faut déterminer s'il s'agit d'un minimum ou d'un maximum. Il existe deux critères pour déterminer si un point candidat est un extremum (minimum ou maximum).

1^er critère

- Si le signe de la dérivée est positif puis devient négatif quand x croît, alors le point candidat est un maximum de la fonction.
- Si le signe de la dérivée est négatif puis devient positif quand x croît, alors le point candidat est un minimum de la fonction.



$2^{ème}$ critère

Le second critère fait appel à la dérivée seconde de la fonction.

$$Soit\ P(x_0; f(x_0))\ le\ point\ en\ lequel\ :\ f'(x_0) = 0.$$

Alors si en ce point :

1. $f''(x_0) < 0$, *il s'agit d'un maximum.*
2. $f''(x_0) > 0$, *il s'agit d'un minimum.*

Exemple 1

Soit la fonction $y = f(x) = x^3 - 3x^2 + 5.$

Calculons sa première dérivée :

$$f'(x) = \frac{dy}{dx} = 3x^2 - 6x = 3x(x-2)$$

et sa deuxième dérivée :

$$f''(x) = \frac{d^2y}{dx^2} = 6x - 6 = 6(x-1)$$

On obtient les points candidats de la fonction en résolvant l'équation $f'(x) = 0$:

$$3x(x-2) = 0 \Rightarrow x_1 = 0 \quad \text{et} \quad x_2 = 2$$

Les valeurs correspondantes de y sont $y_1 = 5$ et $y_2 = 1$.

Exemple 1

Déterminons la nature des points candidats en utilisant le 2ème critère

$$f''(0) = -6 < 0 \text{ et } f''(2) = 6 > 0.$$

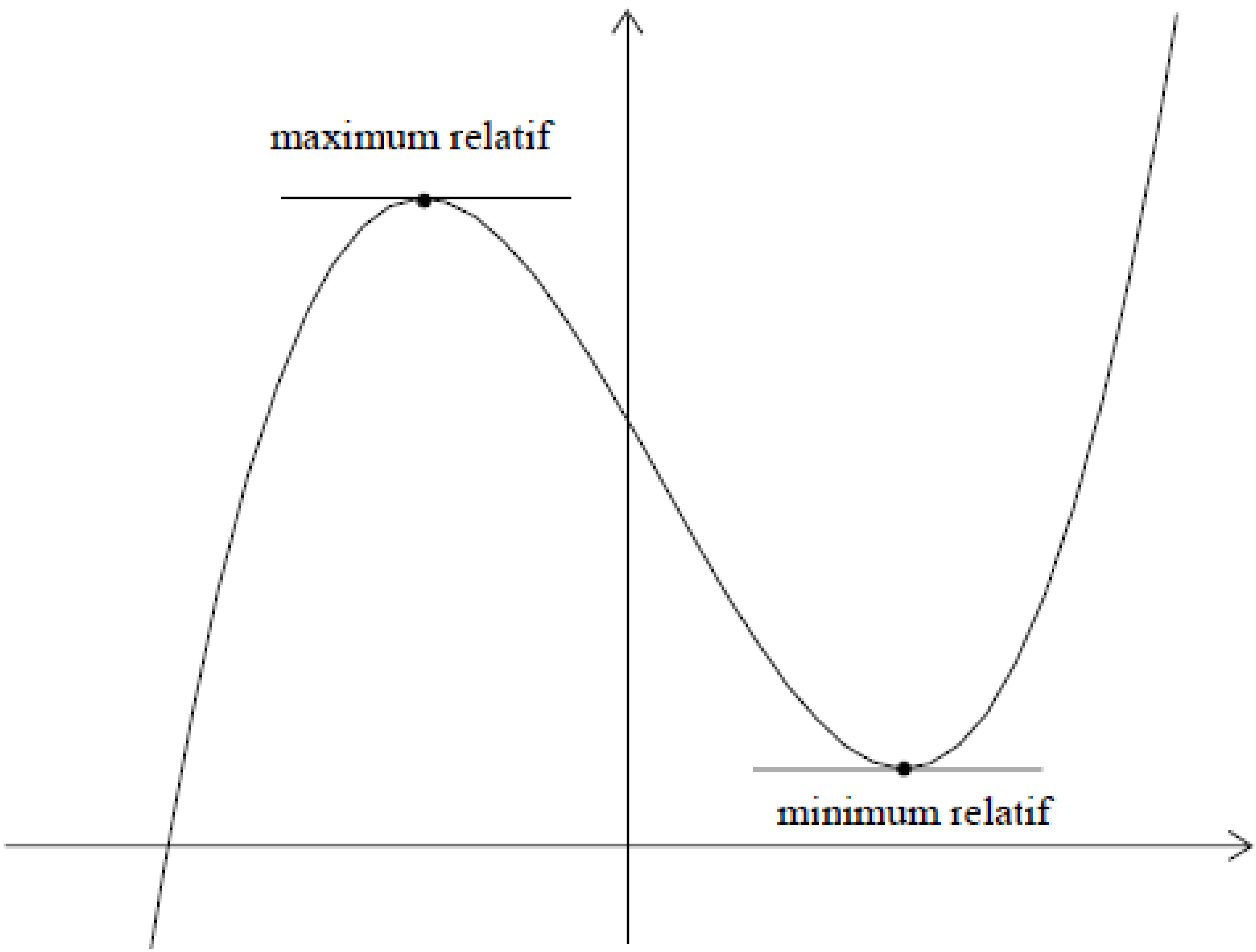
La fonction a donc un **maximum** en $x_1 = 0$ et un **minimum** en $x_2 = 2$.

Les coordonnées du **maximum** sont (0 ,5) et celles du **minimum** (2 ,1).

Remarque

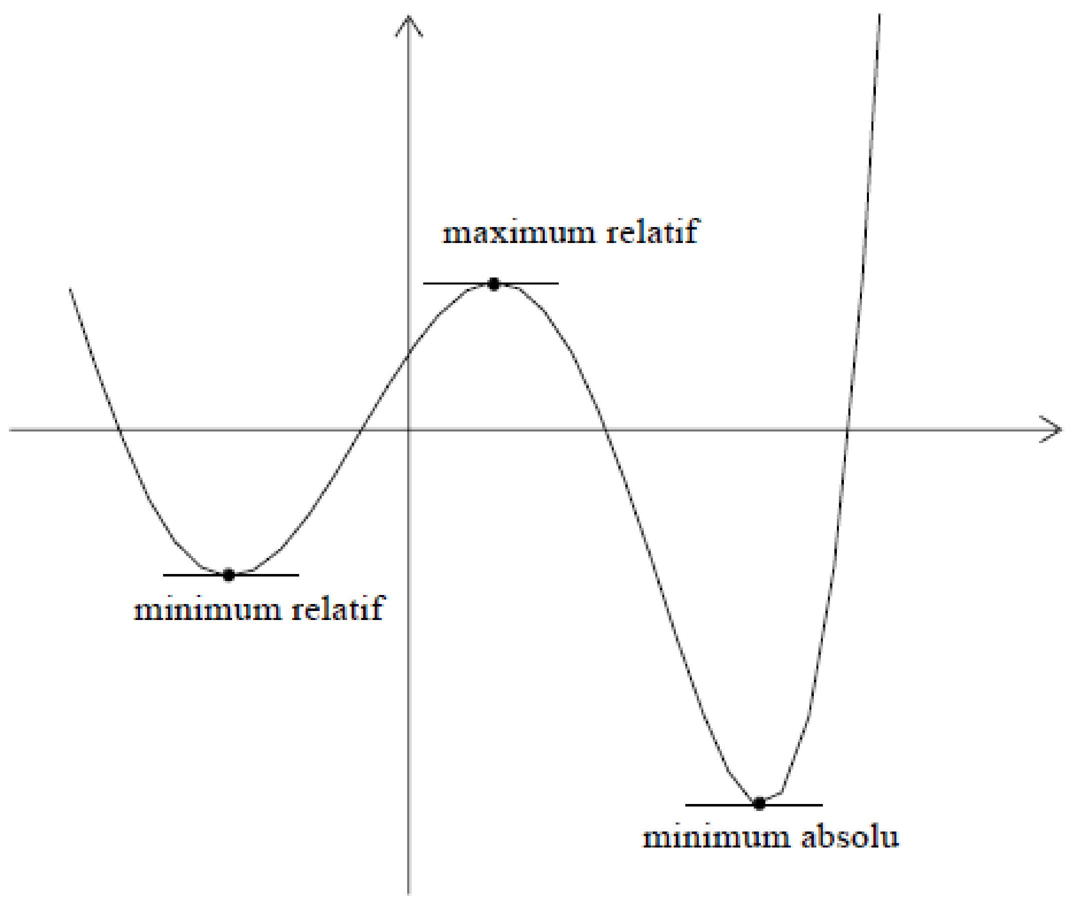
Par conséquent, dans un voisinage convenablement choisi du point $x_1 = 0$, la valeur de la fonction est plus petite que f(x). On dit que la fonction possède un maximum local ou relatif au point $x_1 = 0$. Au point $x_2 = 2$, on dit que la fonction possède un minimum local ou relatif .

Ces deux valeurs, maximum et minimum, sont appelées des extrema locaux ou relatifs, parce qu'il existe des valeurs de x pour lesquelles la fonction prend des valeurs plus grandes que 5 ou plus petites que 1, comme l'illustre la figure suivante


maximum relatif
minimum relatif

Remarque

On peut définir ce que l'on appelle un maximum ou un minimum global ou absolu. En ce point, la fonction prend la valeur la plus grande (ou la plus petite) sur un intervalle donné ; la figure ci-dessous donne un exemple d'extrema relatifs et absolus.



Exemple 2

Un fabricant de postes de télévision produit q postes par semaine à un coût total de C = $6q^2+80q + 5000$. C'est un monopoleur et son prix s'exprime par la relation p = 1080-4q.

Montrons que le bénéfice net maximum est atteint lorsque la production est de 50 postes par semaine.

On sait que le revenu de la firme est égal au prix multiplié par la quantité :

$$\begin{aligned} R = pq &= (1080 - 4q)q \\ &= 1080q - 4q^2 \end{aligned}$$

Son bénéfice s'exprime par :

$$\begin{aligned} B = R - C &= 1080q - 4q^2 - (6q^2 + 80q + 5000) \\ &= -10q^2 + 1000q - 5000 \end{aligned}$$

La dérivée de cette fonction est : $B^{'} = -20q + 1000$

La solution de l'équation $B' = 0$ est: $q = 50$

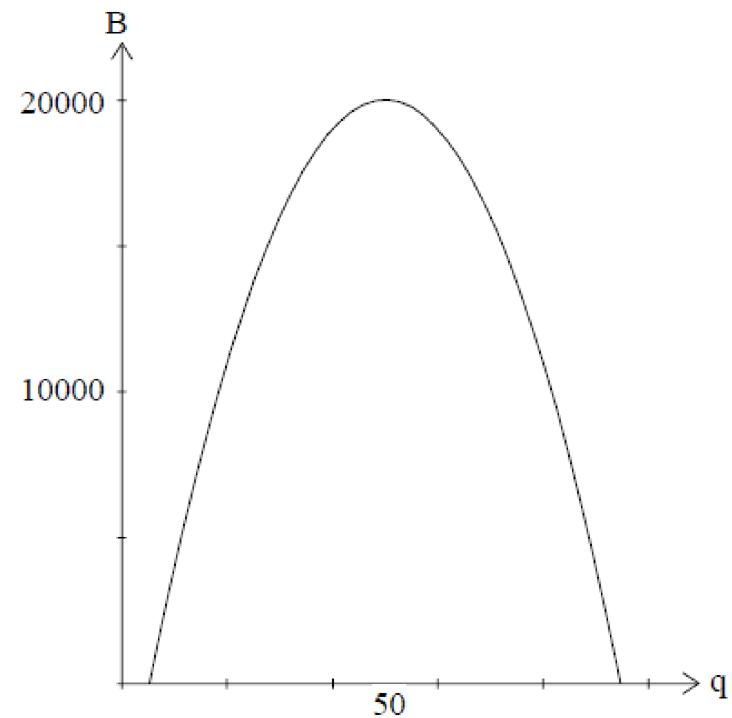
Sa dérivée seconde est : $B'' = -20 < 0$

Par conséquent, la fonction de bénéfice présente un maximum en $x = 50$. Ce maximum est un **maximum global** car la fonction est une parabole, comme l'indique la figure suivante. Pour ce volume de production, le bénéfice réalisé s'élève à:

$B = -10(50)^2 + 1000(50) - 5000 = 20000$

Le prix du poste de télévision est de :

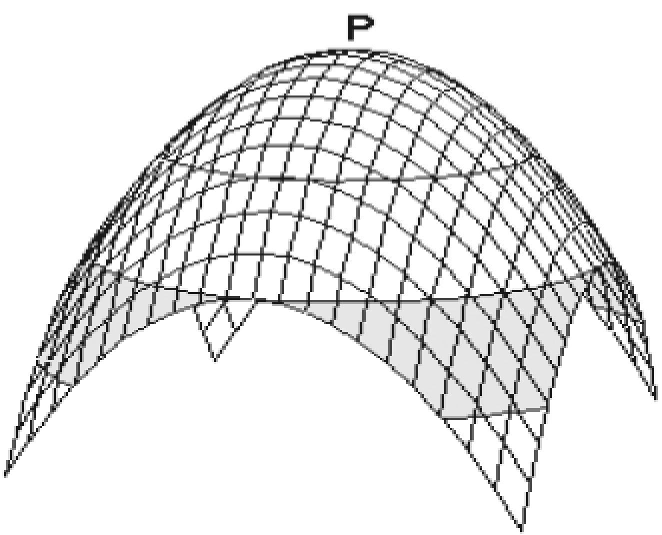
$p = 1080 - 4(50) = 880$

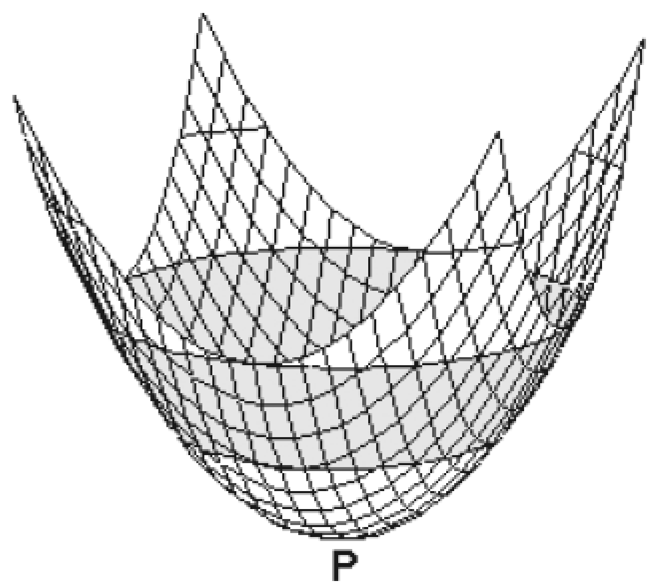


Cas des fonctions à plusieurs variables

Pour les fonctions à deux variables, $z = f(x, y)$, le graphe est une surface dans l'espace à trois dimensions.

- Une telle fonction présente un maximum au point $P(x_0, y_0, f(x_0, y_0))$, si $f(x_0, y_0)$ atteint une valeur supérieure à toutes celles que prend $f(x, y)$ au voisinage de $x = x_0$ et $y = y_0$.
- De même, $f(x, y)$ possède un minimum au point $P(x_0, y_0, f(x_0, y_0))$, si $f(x_0, y_0)$ atteint une valeur inférieure à toutes celles que prend $f(x, y)$ au voisinage de $x = x_0$ et $y = y_0$.





Il en résulte qu'au point $P(x_0, y_0, f(x_0, y_0))$, il existe un plan **tangent horizontal**

Ce plan tangent est engendré par deux tangentes, elles-mêmes déterminées par :

$$\frac{\partial f}{\partial x} \text{ et } \frac{\partial f}{\partial y}$$

Ainsi, la condition nécessaire à l'existence d'un extremum est la suivante :

$$\frac{\partial f}{\partial x} = 0 \text{ et } \frac{\partial f}{\partial y} = 0$$

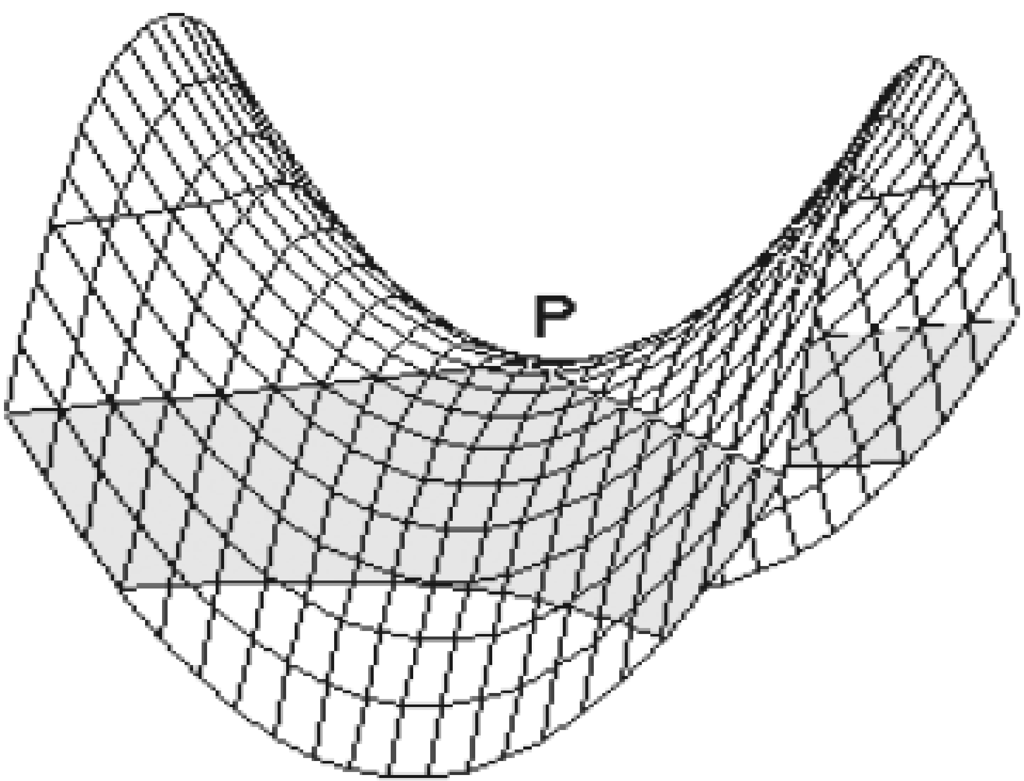
Cette condition est nécessaire mais pas suffisante. En effet, il existe

des fonctions pour lesquelles $\frac{\partial f}{\partial x} = 0 \text{ et } \frac{\partial f}{\partial y} = 0$ sans qu'il existe un extremum en ce point.

Dans ce cas, on parle de point-selle.

Un **point-selle** la fonction présente un **minimum** pour l'une des variables et un **maximum** pour l'autre variable.

La figure ci-dessous illustre cette situation.



Il faut donc remplir une condition suffisante qui est la suivante :

$$\alpha = f_{xx} f_{yy} - f_{xy}^2 > 0$$

où :

- $f_{xx} = \partial^2 f / \partial x^2$, c'est-à-dire que la fonction a été dérivée deux fois par rapport à x,
- $f_{yy} = \partial^2 f / \partial y^2$, ce qui signifie que la fonction a été dérivée deux fois par rapport à y,
- $f_{xy}^2 = (\partial^2 f / \partial x \partial y)^2$, c'est-à-dire que la première dérivée se fait par rapport à y et la deuxième par rapport à x ; cette expression est ensuite élevée au carré.

Condition suffisantepour les fonctions à deux variables.

Soit $P(x_0; y_0; f(x_0, y_0))$ *le point en lequel :*

$$\partial f/\partial x = \partial f/\partial y = 0$$

Alors si en ce point :

1. $f_{xx} > 0$ *et* $\alpha > 0$, f *possède un minimum au point* P.
2. $f_{xx} < 0$ *et* $\alpha > 0$, f *possède un maximum au point* P.
3. $\alpha < 0$, f *ne possède ni minimum ni maximum au point* P, *mais un point-selle.*
4. $\alpha = 0$, *on ne peut pas conclure.*

Notons que la condition suffisante évoquée ci-dessus provient d'un résultat plus général concernant les fonctions à n variables $f(x_1, x_2, \ldots, x_n)$.

Avant d'énoncer ce résultat général, introduisons la matrice des secondes dérivées partielles. Celle-ci joue un rôle clé dans la détermination des extrema d'une fonction à plusieurs variables. Cette matrice est appelée **matrice hessienne** et se présente sous la forme :

$$\boldsymbol{H} = \begin{pmatrix} fx_1^2 & fx_1x_2 & \dots & fx_1x_n \\ fx_2x_1 & fx_2^2 & \dots & fx_2x_n \\ \vdots & \vdots & \ddots & \vdots \\ fx_{n-1}x_1 & fx_{n-1}x_2 & \dots & fx_{n-1}x_n \\ fx_nx_1 & fx_nx_2 & \dots & fx_n^2 \end{pmatrix}$$

où $f_{x_1^2} = \partial^2 f / \partial x_1^2$, $f_{x_1x_2} = \partial^2 f / \partial x_1 \partial x_2$, ..., et $fx_n^2 = \partial^2 f / \partial x_n^2$

On appelle mineurs principaux de la matrice **H**, notés $\mathbf{\Delta}_i$, les déterminants des sous-matrices de **H** obtenues en lui retirant ses n-i dernières lignes et colonnes (i = 1, . . . , n). Dans le cas général, la recherche des extrema d'une fonction à plusieurs variables est basée sur le critère dans le slide suivant

Critère

Soit P le point en lequel :

$$\partial f/\partial x_1 = \partial f/\partial x_2 = \ldots = \partial f/\partial x_n = 0.$$

Alors:

1) Si les mineurs principaux de la matrice hessienne au point P sont tous strictement positifs, il s'agit d'un minimum.
2) Si les mineurs principaux de la matrice hessienne au point P sont de signes alternés, le premier étant strictement négatif, il s'agit d'un maximum.
3) Si les mineurs principaux ne vérifient pas l'une des conditions ci-dessus prises au sens large (c'est-à-dire respectivement ”positif ou nul” et ”négatif ou nul”), il ne s'agit ni d'un minimum ni d'un maximum, mais d'un point-selle.
4) Si les conditions (1) et (2) se vérifient au sens large, alors on ne peut pas conclure.

Cas particulier : 2 variables

En effet, dans ce cas, tout comme $f_{xy} = f_{yx}$, on a :

$$\begin{aligned} \triangle_1 &= f_{xx} \\ \triangle_2 &= f_{xx} f_{yy} - (f_{xy})^2 \end{aligned}$$

Ainsi

$$\begin{cases} \triangle_1 > 0 \text{ et } \triangle_2 > 0 \Longrightarrow \text{minimum.} \\ \triangle_1 < 0 \text{ et } \triangle_2 > 0 \Longrightarrow \text{maximum.} \\ \triangle_1 \text{ quelconque et } \triangle_2 < 0 \Longrightarrow \text{point-selle.} \\ \triangle_1 \text{ quelconque et } \triangle_2 = 0 \Longrightarrow \text{on ne peut pas conclure.} \end{cases}$$

Exemple

Soit la fonction $z = f(x, y) = x^2 + y^2$.

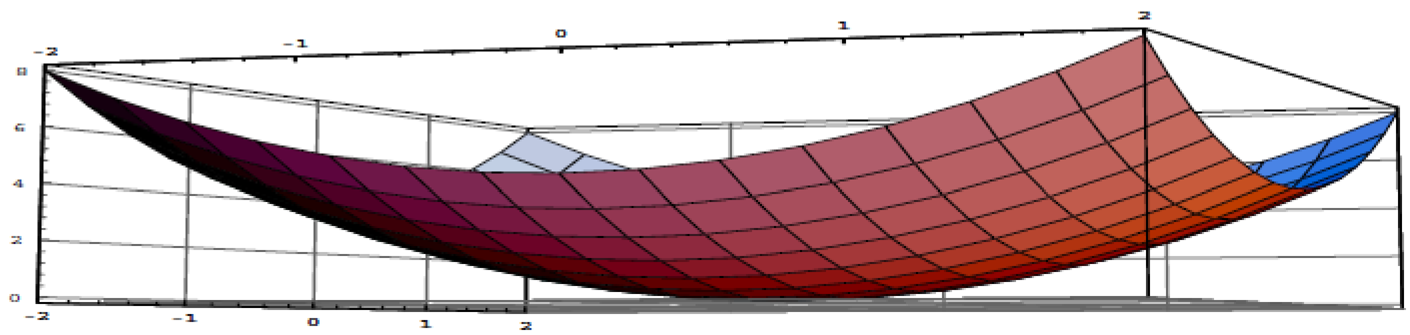
Les candidats aux extrema s'obtiennent en résolvant le système d'équations : $\partial f/\partial x = 0$ et $\partial f/\partial y = 0$.

$$\frac{\partial f}{\partial x} = 2x = 0 \Longrightarrow x = 0$$

$$\frac{\partial f}{\partial y} = 2y = 0 \Longrightarrow y = 0$$

Il existe donc un point candidat en $x_0 = y_0 = z_0 = 0$; ce point est forcément un minimum puisque $f(x, y) = x^2 + y^2 > 0, \forall x \neq 0, \forall y \neq 0$,

comme la figure en témoigne



En effet, en appliquant le résultat précèdent, on trouve :

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = 0$$

La matrice hessienne est donc définie par :

$$\boldsymbol{H} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Ici, $\triangle_1 = 2$ et $\triangle_2 = 4$. Par conséquent, f possède un minimum en $(0; 0; 0)$.

Exemple

Soit $f(x,y) = 4x^2 - xy + y^2 - x^3$.

On a :

$$\frac{\partial f}{\partial x} = 8x - y - 3x^2 \quad \frac{\partial f}{\partial y} = -x + 2y$$

$$\frac{\partial^2 f}{\partial x^2} = 8 - 6x \qquad \frac{\partial^2 f}{\partial y^2} = 2 \qquad \frac{\partial^2 f}{\partial x \partial y} = -1$$

Les points candidats s'obtiennent en résolvant les deux équations :

$$\begin{aligned} 8x - y - 3x^2 &= 0 \\ -x + 2y &= 0 \end{aligned}$$

Après substitution de x = 2y, on trouve :

$$-12y^2 + 15y = 0 \Rightarrow \quad \begin{array}{l} y_1 = 0 \\ y_2 = \dfrac{5}{4} \end{array}$$

Comme $x = 2y$, on trouve $x_1 = 0$ et $x_2 = \dfrac{5}{2}$

Il existe donc deux points candidats : $P_1(0;0;0)$ et $P_2(\dfrac{5}{2};\dfrac{5}{4};\dfrac{125}{16})$.

La matrice hessienne est définie par :

$$\boldsymbol{H} = \begin{pmatrix} 8-6x & -1 \\ -1 & 2 \end{pmatrix}$$

Évaluons la matrice hessienne pour le premier point candidat $x_1 = 0$, $y_1 = 0$:

$$\boldsymbol{H} = \begin{pmatrix} 8 & -1 \\ -1 & 2 \end{pmatrix}$$

Comme $\triangle_1 = 8 > 0$ et $\triangle_2 = 15 > 0$, il s'agit d'un minimum.

Pour le second point candidat $x = \dfrac{5}{2}$ et $y = \dfrac{5}{4}$, on obtient :

$$\boldsymbol{H} = \begin{pmatrix} -7 & -1 \\ -1 & 2 \end{pmatrix}$$

Comme $\triangle_1 = -7 < 0$ et $\triangle_2 = -15 < 0$, il ne s'agit ni d'un minimum ni d'un maximum, mais d'un point-selle de la fonction.

Exemple

Soit $f(x,y,z) = x^4 - 17x^2 + 2y^2 + z^2 - 2xy - 2yz + 81.$

Annulons les premières dérivées partielles :

$$\begin{aligned} \frac{\partial f}{\partial x} &= 4x^3 - 34x - 2y = 0 \\ \frac{\partial f}{\partial y} &= 4y - 2x - 2z = 0 \\ \frac{\partial f}{\partial z} &= 2z - 2y = 0 \end{aligned}$$

Par simplification, de la 3ème équation on tire : $y = z$

Substituons la dernière équation dans la 2ème équation :

$$\begin{aligned} 4y - 2x - 2y &= 0 \\ 2y - 2x &= 0 \\ x &= y \end{aligned}$$

En introduisant les deux dernière équation dans la 1^er équation , on trouve :

$$\begin{aligned} 4x^3 - 34x - 2x &= 0 \\ 4x^3 - 36x &= 0 \\ 4x(x^2 - 9) &= 0 \end{aligned}$$

Cette dernière équation a trois solutions :

$$x_1 = 0, \;\; x_2 = -3, \;\; x_3 = 3$$

Les valeurs correspondantes de y et de z sont :

$$\begin{aligned} y_1 &= 0 \quad y_2 = -3 \quad y_3 = 3 \\ z_1 &= 0 \quad z_2 = -3 \quad z_3 = 3 \end{aligned}$$

Les deuxièmes dérivées partielles sont :

$$\begin{array}{lll} \partial^2 f/\partial x^2 = 12x^2 - 34 & \partial^2 f/\partial x \partial y = -2 & \partial^2 f/\partial x \partial z = 0 \\ \partial^2 f/\partial y \partial x = -2 & \partial^2 f/\partial y^2 = 4 & \partial^2 f/\partial y \partial z = -2 \\ \partial^2 f/\partial z \partial x = 0 & \partial^2 f/\partial z \partial y = -2 & \partial^2 f/\partial z^2 = 2 \end{array}$$

Pour le point $x_1 = y_1 = z_1 = 0$, la matrice hessienne est définie par :

$$\boldsymbol{H} = \begin{pmatrix} -34 & -2 & 0 \\ -2 & 4 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

$$\begin{aligned} \triangle_1 &= -34 \\ \triangle_2 &= -140 \\ \triangle_3 &= -144 \end{aligned}$$

D’après le critère, ces trois mineurs principaux ne vérifient ni la condition 1 ni la condition 2, prises au sens large.

La fonction présente donc un point-selle en $x_1 = y_1 = z_1 = 0$.

Pour les points $x_2 = y_2 = z_2 = -3$ et $x_3 = y_3 = z_3 = 3$, la matrice hessienne est la même :

$$\boldsymbol{H} = \begin{pmatrix} 74 & -2 & 0 \\ -2 & 4 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

On a alors :

$$\begin{aligned} \triangle_1 &= 74 \\ \triangle_2 &= 292 \\ \triangle_3 &= 288 \end{aligned}$$

Ainsi, pour chacun de ces deux points, la fonction présente un minimum.

Optimisation classique avec contraintes

Dans de nombreuses applications pratiques, les variables d'une **fonction donnée** sont soumises à certaines **conditions ou contraintes.** Ces contraintes peuvent être formulées sous forme **d'égalités** ou **d'inégalités**.

- Par exemple, Si un producteur fabrique deux biens, il peut vouloir minimiser le coût total tout en étant obligé de fabriquer une quantité totale minimum spécifiée.
- Une compagnie peut désirer maximiser ses ventes résultant de deux publicités alors qu'elle doit observer la contrainte du budget de publicité.
- Le consommateur désirant maximiser la fonction d'utilité provenant de la consommation de certains biens est restreint par son budget.

Dans le cas où les contraintes s'expriment sous forme d'égalités, l'optimisation de la fonction peut être obtenue grâce à la **méthode des multiplicateurs de Lagrange** pour trouver les extrema d'une fonction soumise à des contraintes d'égalité.

Dans un premier temps, introduisons cette méthode dans le cas simple où la fonction à optimiser (fonction objectif) est une fonction à deux variables f(x, y) soumise à une seule contrainte de la forme g(x, y) = 0.

La méthode des multiplicateurs de Lagrange consiste à construire une fonction auxiliaire F(x, y, λ), appelée Lagrangien, définie ainsi :

$$F(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

où λ (appelé multiplicateur de Lagrange) est une inconnue. Il faut ensuite annuler ses premières dérivées partielles (condition nécessaire) :

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0$$

$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0$$

$$\frac{\partial F}{\partial \lambda} = g(x, y) = 0$$

Les points candidats s'obtiennent en résolvant ce système de trois équations à trois inconnues (x, y, λ).

Mentionnons que la troisième équation de ce système $\partial F/\partial\lambda = g(x, y) = 0$ n'est rien d'autre que la contrainte ! Les points candidats satisfont par conséquent cette contrainte.

La solution des trois équations ci-dessus fournit les points candidats de la fonction sous contrainte. Ces points candidats satisfont la contrainte mais il reste à déterminer leur nature ; pour cela, introduisons la **matrice hessienne bordée :**

$$\boldsymbol{H} = \begin{pmatrix} \frac{\partial^2 F}{\partial\lambda^2} & \frac{\partial^2 F}{\partial\lambda\partial x} & \frac{\partial^2 F}{\partial\lambda\partial y} \\ \frac{\partial^2 F}{\partial x\partial\lambda} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x\partial y} \\ \frac{\partial^2 F}{\partial y\partial\lambda} & \frac{\partial^2 F}{\partial y\partial x} & \frac{\partial^2 F}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x\partial y} \\ \frac{\partial g}{\partial y} & \frac{\partial^2 F}{\partial y\partial x} & \frac{\partial^2 F}{\partial y^2} \end{pmatrix}$$

dont le déterminant sera noté $\mid \boldsymbol{H} \mid$.

La condition suffisante pour l'existence d'un extremum

$$Soit\ P(x_0; y_0; f(x_0, y_0))\ le\ point\ en\ lequel$$

$$\partial F/\partial x = \partial F/\partial y = \partial F/\partial \lambda = 0$$

$Alors,\ si\ en\ ce\ point$

$$\mid \boldsymbol{H} \mid < 0 \Longrightarrow minimum\ au\ point\ P.$$

$$\mid \boldsymbol{H} \mid > 0 \Longrightarrow maximum\ au\ point\ P.$$

Généralisations de la méthode des multiplicateurs de Lagrange

La méthode des multiplicateurs de Lagrange peut se généraliser à l'optimisation d'une fonction à n variables $f(x_1, \ldots, x_n)$ soumise à k contraintes :

$$g_j(x_1, \ldots, x_n) = 0, \;\; j = 1, \ldots, k \text{ où } 1 \leq k \leq n$$

Dans ce cas, le Lagrangien s'écrit :

$$F(x_1, x_2, \ldots, x_n, \lambda_1, \ldots, \lambda_k) = f(x_1, \ldots, x_n) + \sum_{j=1}^{k} \lambda_j g_j(x_1, \ldots, x_n)$$

L'annulation des premières dérivées partielles fournit un système de $n+k$ équations à $n+k$ inconnues :

$$\begin{aligned}
\frac{\partial F}{\partial x_1} &= \frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1} + \ldots + \lambda_k \frac{\partial g_k}{\partial x_1} = 0 \\
\frac{\partial F}{\partial x_2} &= \frac{\partial f}{\partial x_2} + \lambda_1 \frac{\partial g_1}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2} + \ldots + \lambda_k \frac{\partial g_k}{\partial x_2} = 0 \\
&\vdots \\
\frac{\partial F}{\partial x_n} &= \frac{\partial f}{\partial x_n} + \lambda_1 \frac{\partial g_1}{\partial x_n} + \lambda_2 \frac{\partial g_2}{\partial x_n} + \ldots + \lambda_k \frac{\partial g_k}{\partial x_n} = 0 \\
\frac{\partial F}{\partial \lambda_1} &= g_1(x_1, \ldots, x_n) = 0 \\
&\vdots \\
\frac{\partial F}{\partial \lambda_k} &= g_k(x_1, \ldots, x_n) = 0
\end{aligned}$$

Les conditions du deuxième ordre permettant de déterminer s'il s'agit d'un maximum ou d'un minimum reposent sur le calcul des mineurs de la matrice hessienne bordée suivante :

$$
\mathbf{H} = \left(\begin{array}{ccc|ccc}
 & \cdots & 0 & \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\
0 & \cdots & 0 & \dfrac{\partial g_2}{\partial x_1} & \cdots & \dfrac{\partial g_2}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \dfrac{\partial g_k}{\partial x_1} & \cdots & \dfrac{\partial g_k}{\partial x_n} \\
\hline
\dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_k}{\partial x_1} & \dfrac{\partial^2 F}{\partial x_1^2} & \cdots & \dfrac{\partial^2 F}{\partial x_1 \partial x_n} \\
\dfrac{\partial g_1}{\partial x_2} & \cdots & \dfrac{\partial g_k}{\partial x_2} & \dfrac{\partial^2 F}{\partial x_2 \partial x_1} & \cdots & \dfrac{\partial^2 F}{\partial x_2 \partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\dfrac{\partial g_1}{\partial x_n} & \cdots & \dfrac{\partial g_k}{\partial x_n} & \dfrac{\partial^2 F}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 F}{\partial x_n^2}
\end{array}\right)
$$

Notons le mineur principal qui contient $\partial^2 F/\partial x_1^2$ comme dernier élément de la diagonale principale par $|\ \boldsymbol{H}_1\ |$. $|\ \boldsymbol{H}_2\ |$ correspond au mineur principal qui contient $\partial^2 F/\partial x_2^2$ comme dernier élément de la diagonale principale et ainsi de suite. La condition suffisante pour l'existence d'un minimum ou d'un maximum dépend des signes des mineurs principaux $|\ \boldsymbol{H}_{k+1}\ |$, $|\ \boldsymbol{H}_{k+2}\ |$, $\ldots,|\ \boldsymbol{H}_n\ |=|\ \boldsymbol{H}\ |$.

Mentionnons qu'il existe au moins une contrainte (k ≥ 1) et que, par conséquent, $|\ \boldsymbol{H}_1\ |$ n'intervient jamais dans les calculs.

La condition suffisante pour l'existence d'un extremum

La fonction $f(x_1, \ldots, x_n)$ *soumise aux* k *contraintes :*

$$g_j(x_1, \ldots, x_n) = 0, \;\; j = 1, \ldots, k$$

admet :

- *un maximum au point candidat si les mineurs principaux* $\mid \boldsymbol{H}_{k+1} \mid$, $\mid \boldsymbol{H}_{k+2} \mid, \ldots, \mid \boldsymbol{H}_n \mid$ *sont de signe alterné, le signe de* $\mid \boldsymbol{H}_{k+1} \mid$ *étant celui de* $(-1)^{k+1}$,

- *un minimum si les mineurs principaux* $\mid \boldsymbol{H}_{k+1} \mid$, $\mid \boldsymbol{H}_{k+2} \mid, \ldots, \mid \boldsymbol{H}_n \mid$ *sont de même signe, celui de* $(-1)^k$.

Exemple