Chapitre

Optimisation

3.1 Introduction

Euclide formulait déjà des problèmes d'optimisation au III^{e} siècle avant J.-C. Ce prolifique mathématicien de l'Antiquité a rédigé l'un des plus célèbres textes de l'histoire des mathématiques : *les Éléments.* Cette œuvre comprend treize livres et couvre des domaines tels que la géométrie plane et solide ou la théorie des nombres rationnels et irrationnels. Durant vingt siècle, ce traité fut l'ouvrage de référence dans l'enseignement mathématique. Plus de mille éditions en furent tirées, depuis la première en 1482. Sir Isaac Newton (1642-1727), auteur du célèbre *Mathematical Principles of Natural Philosophy,* ainsi que Gottfried Wilhelm Leibniz (1646-1716) offrirent, à la fin du XVII^{e} siècle, les premiers outils de résolution de certains problèmes d'optimisation relatifs à la géométrie et à la physique. Ils avaient en effet inventé, simultanément mais néanmoins indépendamment, le calcul différentiel.

Comme indiqué au chapitre 1, 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). Les problèmes d'optimisation en économie sont souvent caractérisés par un nombre très élevé de variables et par la nécessité de trouver des solutions non négatives. C'est donc un type d'optimisation avec contraintes.

Nous allons reprendre ces deux types d'optimisation et les développer dans ce chapitre.

3.2 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). Le premier critère est fourni par le résultat 3.1.

Résultat 3.1 *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.

La figure 3.1 illustre ce résultat.

Le second critère fait appel à la dérivée seconde de la fonction. La dérivée seconde $f^{''}(x)$ d'une fonction est la dérivée de la dérivée première $f^{'}(x)$. Celle-ci mesure le taux de croissance ou de décroissance de $f^{'}(x)$, autrement dit le taux de croissance ou de décroissance de la pente de la tangente à la courbe. Le signe de $f^{''}(x)$ au point candidat $x = x_0$ fournit les renseignements nécessaires. Si $f^{''}(x_0)$ est positive, la pente de la tangente croît quand x croît en passant par x_0. Inversement, si $f^{''}(x_0)$ est négative, la pente de la tangente décroît quant x croît en passant par x_0. D'où le résultat 3.2 basé sur la dérivée seconde.

Résultat 3.2 *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.*

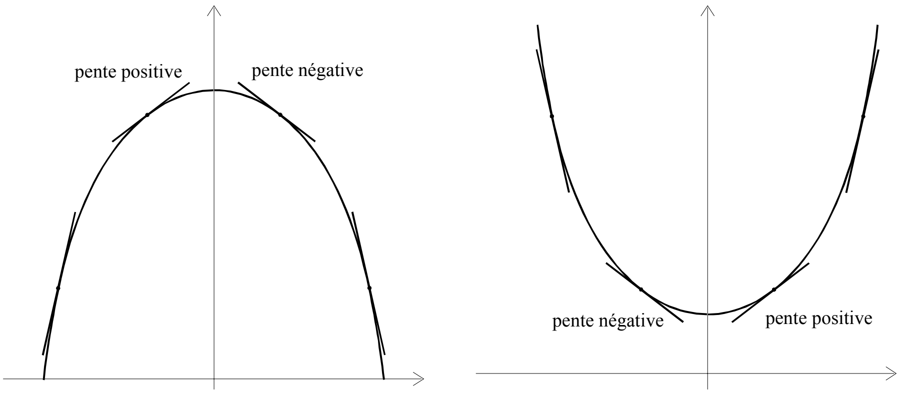

Figure 3.1 : Maximum et minimum d'une fonction à une variable

Exemple 3.1 *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 $y_2 = 1$.

Déterminons la nature des points candidats en utilisant le résultat 3.2 $f''(0) = -6 < 0$ 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).

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 3.2 (a) :

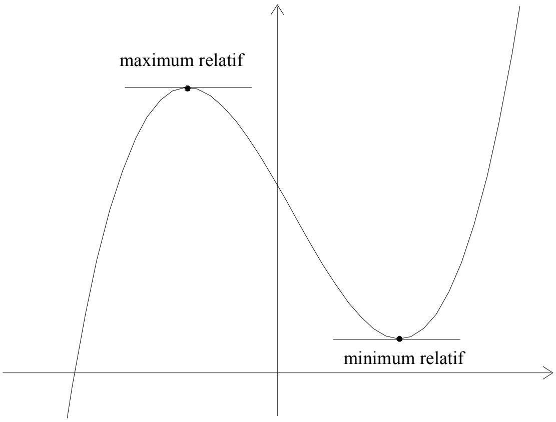

Figure 3.2 (a) : Extrema locaux

Cependant, 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 3.2 (b) donne un exemple d'extrema relatifs et absolus.

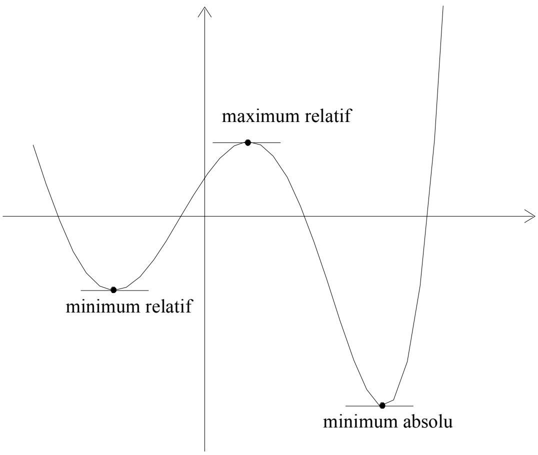

Figure 3.2 (b) : Extrema relatifs et absolus

Exemple 3.2 *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 :

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

Son bénéfice s'exprime par :

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

La dérivée de cette fonction est :

$$B^{'} = -20q + 1000$$

La solution de l'équation $B^{'} = 0$ est :

$$\begin{aligned} -20q + 1000 = 0 \\ q = 50 \end{aligned}$$

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 3.3. 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 :

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

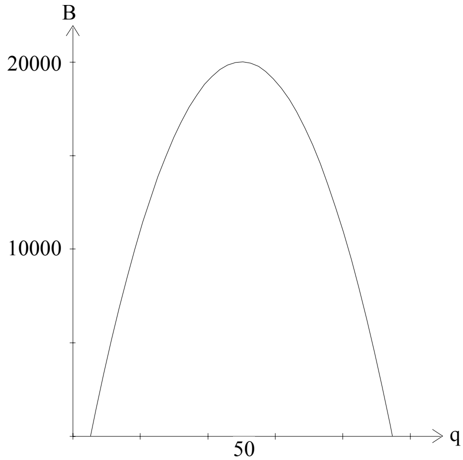

Figure 3.3 : Fonction bénéfice ($B = -10q^2 + 1000q - 5000$)

Considérons à présent le cas des fonctions à plusieurs variables indépendantes. Des exemples simples en sont fournis par des formules de mathématiques élémentaires. Ainsi, dans la formule de calcul du volume V d'un cylindre droit à base circulaire, $V = \pi r^2 h$, V est une fonction à deux variables indépendantes : r (rayon du cercle de base) et h (hauteur). De la même manière, dans la formule qui donne l'aire A d'un triangle quelconque,
$A = \frac{1}{2} xy \sin\alpha$, A est une fonction à trois variables indépendantes, x, y et α, qui traduisent respectivement la longeur de deux côtés du triangle et l'angle formé par ces deux côtés.

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 $y = y_0$, comme indiqué sur la figure 3.4 (a). 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 $y = y_0$; ce cas est illustré par la figure 3.4 (b). 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
:

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

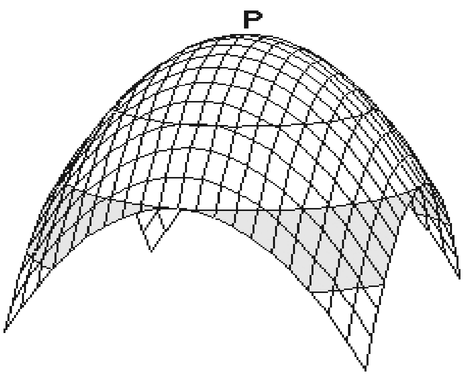

Figure 3.4 (a) : Maximum au point P

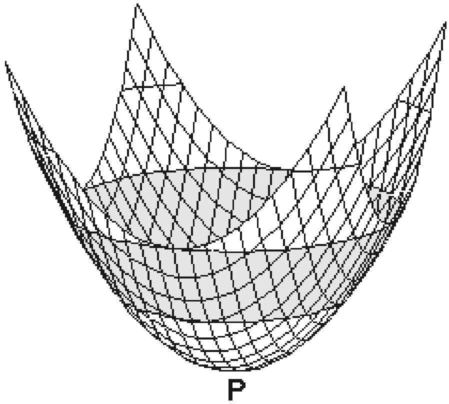

Figure 3.4 (b) : Minimum au point P

Cette condition est nécessaire mais pas suffisante. En effet, il existe des fonctions pour lesquelles $\partial f/\partial x = \partial f/\partial y = 0$ sans qu'il existe un extremum en ce point. Dans ce cas, on parle de **point-selle**. Bien que les deux tangentes soient horizontales, il est toujours possible de trouver un point situé au-dessus du point-selle et un autre au-dessous, ceci quelque soit le voisinage du point-selle considéré. Notons encore, qu'en un point-selle la fonction présente un minimum pour l'une des variables et un maximum pour l'autre variable. La figure 3.5 illustre cette situation.

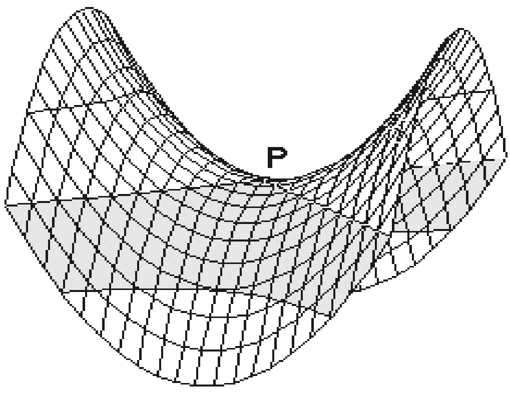

Figure 3.5 : Point-selle en P

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.

Le résultat 3.3 résume la situation pour les fonctions à deux variables.

Résultat 3.3 *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 $\boldsymbol{H}$, notés $\triangle_i$, les déterminants des sous-matrices de $\boldsymbol{H}$ obtenues en lui retirant ses $n-i$ dernières lignes et colonnes $(i = 1, \dots, n)$.

Dans le cas général, la recherche des extrema d'une fonction à plusieurs variables est basée sur le résultat 3.4.

Résultat 3.4 *Soit P le point en lequel :*

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

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.*

Dans le cas de fonctions à deux variables, on retrouve le résultat 3.3. 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 3.3 *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 3.6 en témoigne. En effet, en appliquant le résultat 3.4, 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 $(0; 0; 0)$.

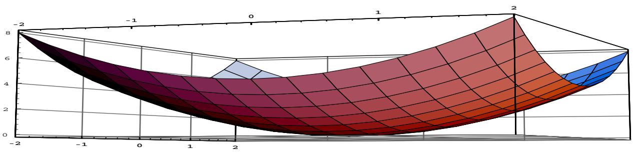

Figure 3.6 : Graphe de $z = f(x, y) = x^2 + y^2$

Exemple 3.4 *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 :

$$8x - y - 3x^2 = 0 \tag{3.1}$$

$$-x + 2y = 0 \tag{3.2}$$

Après substitution de $x = 2y$ (tiré de (3.2)) dans (3.1), 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 $x = \frac{5}{2}$ et $y = \frac{5}{4}$, 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 3.5 *Soit* $f(x, y, z) = x^4 - 17x^2 + 2y^2 + z^2 - 2xy - 2yz + 81$.

Annulons les premières dérivées partielles :

$$\frac{\partial f}{\partial x} = 4x^3 - 34x - 2y = 0 \tag{3.3}$$

$$\frac{\partial f}{\partial y} = 4y - 2x - 2z = 0 \tag{3.4}$$

$$\frac{\partial f}{\partial z} = 2z - 2y = 0 \tag{3.5}$$

Par simplification, de (3.5) on tire :

$$y = z \tag{3.6}$$

Substituons (3.6) dans (3.4) :

$$\begin{aligned} 4y - 2x - 2y &= 0 \\ 2y - 2x &= 0 \\ x &= y \end{aligned} \tag{3.7}$$

En introduisant (3.6) et (3.7) dans (3.3), 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 résultat 3.4, 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 :

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

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

Exemple 3.6 *Une firme aéronautique fabrique des avions qu'elle vend sur deux marchés étrangers. Soit q_1 le nombre d'avions vendus sur le premier marché et q_2 le nombre d'avions vendus sur le deuxième marché. Les fonctions de demande dans les deux marchés respectifs sont :*

$$p_1 = 60 - 2q_1$$
$$p_2 = 80 - 4q_2$$

P_1 et p_2 sont les deux prix de vente. La fonction de coût total de la firme est :

$$C = 50 + 40q$$

où q est le nombre total d'avions produits. Il faut trouver le nombre d'avions que la firme doit vendre sur chaque marché pour maximiser son bénéfice.

Comme $q = q_1 + q_2$, la fonction de coût devient :

$$\begin{aligned}C &= 50 + 40q\\ &= 50 + 40(q_1 + q_2)\\ &= 50 + 40q_1 + 40q_2\end{aligned}$$

Le revenu total R s'obtient en multipliant le prix par la quantité sur chaque marché :

$$\begin{aligned}R &= p_1q_1 + p_2q_2\\ &= (60 - 2q_1)q_1 + (80 - 4q_2)q_2\\ &= 60q_1 - 2q_1^2 + 80q_2 - 4q_2^2\end{aligned}$$

On obtient le bénéfice B en calculant la différence entre le revenu et le coût :

$$\begin{aligned}B &= R - C\\ &= 60q_1 - 2q_1^2 + 80q_2 - 4q_2^2 - (50 + 40q_1 + 40q_2)\\ &= 20q_1 - 2q_1^2 + 40q_2 - 4q_2^2 - 50\end{aligned}$$

Annulons les premières dérivées partielles :

$$\begin{aligned} \partial B/\partial q_1 &= -4q_1 + 20 = 0 \Rightarrow q_1 = 5 \\ \partial B/\partial q_2 &= -8q_2 + 40 = 0 \Rightarrow q_2 = 5 \end{aligned}$$

Il reste à vérifier que le point candidat $(q_1; q_2) = (5; 5)$ est un maximum ; pour cela, calculons les deuxièmes dérivées partielles :

$$\partial^2 B/\partial q_1^2 = -4 \quad \partial^2 B/\partial q_2^2 = -8 \quad \partial^2 B/\partial q_1 \partial q_2 = 0$$

La matrice hessienne est donc :

$$\boldsymbol{H} = \begin{pmatrix} -4 & 0 \\ 0 & -8 \end{pmatrix}$$

Par conséquent :

$$\begin{aligned} \triangle_1 &= -4 < 0 \\ \triangle_2 &= 32 > 0 \end{aligned}$$

Comme $\triangle_1 < 0$ et $\triangle_2 > 0$, il s'agit d'un maximum.
Le bénéfice maximum réalisé est égal à :

$$20(5) - 2(5)^2 + 40(5) - 4(5)^2 - 50 = 100$$

Quant aux prix, ils s'élèvent respectivement à :

$$\begin{aligned} p_1 &= 60 - 2(5) = 50 \\ p_2 &= 80 - 4(5) = 60 \end{aligned}$$

3.3 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. De même, une compagnie peut désirer maximiser ses ventes résultant de deux publicités alors qu'elle doit observer la contrainte du budget de publicité. Enfin, 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** qui est la plus largement répandue pour trouver les extrema d'une fonction soumise à des contraintes d'égalité. Cette méthode tient son nom du mathématicien franco-italien (1736-1813), à qui l'on doit le fameux *Mécanique céleste* (1788), ouvrage résumant sous une forme rigoureuse toutes les connaissances acquises en matière de mécanique depuis Newton. C'est dans le cadre de deux mémoires importants sur la théorie des équations, publiés en 1770 et 1771, qu'apparaît pour la première fois le résultat connu sous le terme de théorème de Lagrange.

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 $F(x, y, \lambda)$, 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 :**

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

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

La condition suffisante pour l'existence d'un extremum est fournie par le résultat 3.5.

Résultat 3.5 *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

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

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

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 $\mid \boldsymbol{H}_1 \mid$. $\mid \boldsymbol{H}_2 \mid$ 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 $\mid \boldsymbol{H}_{k+1} \mid$, $\mid \boldsymbol{H}_{k+2} \mid$, $\ldots, \mid \boldsymbol{H}_n \mid = \mid \boldsymbol{H} \mid$.

Mentionnons qu'il existe au moins une contrainte $(k \geq 1)$ et que, par conséquent, $\mid \boldsymbol{H}_1 \mid$ n'intervient jamais dans les calculs. La condition suffisante pour l'existence d'un extremum est donnée dans le résultat 3.6.

Résultat 3.6 *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* $|\boldsymbol{H}_{k+1}|, |\boldsymbol{H}_{k+2}|, \ldots, |\boldsymbol{H}_n|$ *sont de même signe, celui de* $(-1)^k$.

Exemple 3.9 *Trouver les extrema de la fonction objectif :*

$$f(x, y) = 5x^2 + 6y^2 - xy$$

sous la contrainte :

$$x + 2y = 24$$

La contrainte s'écrit $g(x, y) = x + 2y - 24 = 0$.
Le Lagrangien est donné par :

$$F(x, y, \lambda) = 5x^2 + 6y^2 - xy + \lambda(x + 2y - 24)$$

L'annulation des premières dérivées partielles fournit un système de trois équations à trois inconnues qu'il s'agit de résoudre.

$$\frac{\partial F}{\partial x} = 10x - y + \lambda = 0 \tag{3.12}$$

$$\frac{\partial F}{\partial y} = 12y - x + 2\lambda = 0 \tag{3.13}$$

$$\frac{\partial F}{\partial \lambda} = x + 2y - 24 = 0 \tag{3.14}$$

En éliminant λ des équations (3.12) et (3.13), on obtient $2y = 3x$ que l'on substitue dans (3.14).

$$\begin{gathered} x + 3x - 24 = 0 \\ 4x = 24 \end{gathered}$$

On obtient $x_0 = 6$.
Comme $x + 2y = 24$, on trouve $y_0 = 9$.
Pour déterminer si le point candidat $x_0 = 6$ et $y_0 = 9$ est un extremum, il faut calculer les dérivées partielles du deuxième ordre.

$$\frac{\partial^2 F}{\partial x^2} = 10 \qquad \frac{\partial^2 F}{\partial y^2} = 12 \qquad \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = -1$$

Les dérivées partielles de la contrainte $g(x, y) = x + 2y - 24$ sont :

$$\frac{\partial g}{\partial x} = 1 \text{ et } \frac{\partial g}{\partial y} = 2$$

La matrice hessienne bordée est donc

$$\boldsymbol{H} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 10 & -1 \\ 2 & -1 & 12 \end{pmatrix}$$

Puisqu'il s'agit d'une fonction à deux variables soumise à une contrainte, on utilise le résultat 3.5.

Comme $\mid \boldsymbol{H} \mid = -56 < 0$, la fonction objectif sous la contrainte $x+2y = 24$ possède un minimum en $x_0 = 6$ et $y_0 = 9$.

Nous avons donc trouvé la solution qui minimise la fonction objectif tout en respectant la contrainte. Remarquons que cette même fonction, si elle n'est pas soumise à la contrainte $x + 2y = 24$, ne possède pas un minimum au même point ! Le lecteur peut vérifier que sans la contrainte, cette fonction possède un minimum en $x_0 = y_0 = 0$.

Exemple 3.10 *Une entreprise fabrique trois types de machines : x_1, x_2 et x_3. La fonction de coût conjointe $C(x_1, x_2, x_3)$ est :*

$$C(x_1, x_2, x_3) = 4x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + x_2x_3 - 30x_2 - 30x_3$$

Combien de machines de chaque type l'entreprise doit-elle fabriquer pour minimiser son coût s'il lui faut un total de 100 machines ?

Il s'agit ici de minimiser la fonction de coût conjointe sous la contrainte :

$$x_1 + x_2 + x_3 = 100$$

Ainsi $g(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 100$.

La fonction de Lagrange $F(x_1, x_2, x_3, \lambda)$ est donnée par :

$$\begin{aligned} F(x_1, x_2, x_3, \lambda) &= C(x_1, x_2, x_3) + \lambda g(x_1, x_2, x_3) \\ &= 4x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + x_2x_3 - 30x_2 - 30x_3 \\ &+\lambda(x_1 + x_2 + x_3 - 100) \end{aligned}$$

Annulons les premières dérivées partielles :

$$\frac{\partial F}{\partial x_1} = 8x_1 - 2x_2 + \lambda = 0 \tag{3.15}$$

$$\frac{\partial F}{\partial x_2} = 4x_2 - 2x_1 + x_3 - 30 + \lambda = 0 \tag{3.16}$$

$$\frac{\partial F}{\partial x_3} = 2x_3 + x_2 - 30 + \lambda = 0 \tag{3.17}$$

$$\frac{\partial F}{\partial \lambda} = x_1 + x_2 + x_3 - 100 = 0 \tag{3.18}$$

Tirons λ de la première équation : $\lambda = 2x_2 - 8x_1$. En substituant λ dans les équations (3.16) et (3.17), on obtient avec l'équation (3.18) un système de trois équations à trois inconnues :

$$-10x_1 + 6x_2 + x_3 = 30 \tag{3.19}$$

$$-8x_1 + 3x_2 + 2x_3 = 30 \tag{3.20}$$

$$x_1 + x_2 + x_3 = 100 \tag{3.21}$$

De (3.21), on tire $x_1 = 100 - x_2 - x_3$ que l'on remplace dans (3.19) et (3.20) pour obtenir :

$$16x_2 + 11x_3 = 1030 \tag{3.22}$$

$$11x_2 + 10x_3 = 830 \tag{3.23}$$

De (3.23), on tire $x_3 = \dfrac{830 - 11x_2}{10}$ que l'on substitue dans (3.22).

$$16x_2 + \frac{11}{10}(830 - 11x_2) = 1030$$

$$\begin{aligned} 160x_2 - 121x_2 &= 1170 \\ 39x_2 &= 1170 \\ x_2 &= 30 \end{aligned}$$

En utilisant ce résultat dans (3.23), on obtient :

$$\begin{aligned} 330 + 10x_3 &= 830 \\ 10x_3 &= 500 \\ x_3 &= 50 \end{aligned}$$

Finalement, en se basant sur (3.21), on en conclut que $x_1 = 100 - 30 - 50 = 20$.

Ainsi, le point candidat est $x_1 = 20, x_2 = 30$ et $x_3 = 50$. Vérifions à présent qu'il s'agit bien d'un minimum. Pour cela, calculons :

$$\begin{array}{lllllllll} \dfrac{\partial g}{\partial x_1} & = & 1 & \dfrac{\partial g}{\partial x_2} & = & 1 & \dfrac{\partial g}{\partial x_3} & = & 1 \\[2ex] \dfrac{\partial^2 F}{\partial x_1^2} & = & 8 & \dfrac{\partial^2 F}{\partial x_2^2} & = & 4 & \dfrac{\partial^2 F}{\partial x_3^2} & = & 2 \\[2ex] \dfrac{\partial^2 F}{\partial x_1 \partial x_2} & = & -2 & \dfrac{\partial^2 F}{\partial x_1 \partial x_3} & = & 0 & \dfrac{\partial^2 F}{\partial x_2 \partial x_3} & = & 1 \end{array}$$

La matrice hessienne bordée est par conséquent donnée par :

$$\boldsymbol{H} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 8 & -2 & 0 \\ 1 & -2 & 4 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix}$$

On a alors :

$$\mid \boldsymbol{H}_2 \mid = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 8 & -2 \\ 1 & -2 & 4 \end{vmatrix} = -16 \text{ et } \mid \boldsymbol{H}_3 \mid = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 8 & -2 & 0 \\ 1 & -2 & 4 & 1 \\ 1 & 0 & 1 & 2 \end{vmatrix} = -39$$

Comme dans cet exemple il n'y a qu'une contrainte, ($k = 1$), on a donc :

$$\begin{aligned} \mid \boldsymbol{H}_{k+1} \mid &= \mid \boldsymbol{H}_2 \mid = -16 \\ \mid \boldsymbol{H}_{k+2} \mid &= \mid \boldsymbol{H}_n \mid = \mid \boldsymbol{H}_3 \mid = -39 \\ (-1)^k &= (-1)^1 = -1 \end{aligned}$$

Selon le résultat 3.6, les mineurs $| \boldsymbol{H}_2 |$ et $| \boldsymbol{H}_3 |$ sont du même signe que $(-1)^k$; le point $x_1 = 20$, $x_2 = 30$ et $x_3 = 50$ est donc un minimum.