

Si la fonction réelle f est différentiable sur Ω et strictement convexe (resp. strictement concave), alors elle possède au plus un point stationnaire et donc, au plus, un minimum (resp. maximum) en un point de Ω .

En effet, si $x_* \in \Omega$ est un point stationnaire de f et si f est strictement convexe, alors, selon l'inégalité (3.111) dans le cas d'une fonction strictement convexe,

$$f(x) > f(x_*)$$

pour tout $x \neq x_*$ appartenant à Ω . Dès lors, f ne peut être stationnaire et minimale qu'en un seul point x_* de Ω . □

EXEMPLE 3.40 La fonction

$$f(x, y) = x^2 + y^2$$

présentée à la figure 3.8 est strictement convexe sur \mathbb{R}^2 . En effet, sa matrice hessienne est donnée par

$$H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

et est définie positive en tout point de \mathbb{R}^2 .

Conformément au théorème énoncé ci-dessus, la fonction f présente un seul minimum absolu en son seul point stationnaire $(0, 0)$. ◇

3.10.4 Optimisation avec contraintes d'égalité.

Dans de nombreux problèmes, on désire identifier le point x maximisant ou minimisant une fonction f , non pas parmi tous les x appartenant au domaine de définition de f mais seulement parmi ceux qui vérifient une ou plusieurs contraintes du type $g_k(x) = 0$ pour tout $k \in \{1, 2, \dots, \ell\}$.

Très souvent, ce sont ces contraintes qui donnent du sens au problème ; sans contrainte il n'y aurait aucune solution optimale.

EXEMPLE 3.41 Déterminer les dimensions d'une boîte en carton parallélépipédique et dépourvue de couvercle dont la construction demande le moins de carton possible tout en ayant une contenance V déterminée.

Le problème peut être exprimé sous la forme de la recherche des dimensions x, y et z de la boîte minimisant $A = xy + 2xz + 2yz$ sous la contrainte $xyz = V$.

Manifestement, le problème n'a pas de sens sans la contrainte $xyz = V$. Celle-ci peut être utilisée pour éliminer la variable z de l'expression de A (si $V \neq 0$, les trois dimensions de la boîte diffèrent de zéro) :

$$z = \frac{V}{xy} \quad \Rightarrow \quad A(x, y) = xy + \frac{2V}{y} + \frac{2V}{x}$$

Les points stationnaires de A sont les solutions de

$$\begin{cases} \frac{\partial A}{\partial x} = y - \frac{2V}{x^2} = 0 \\ \frac{\partial A}{\partial y} = x - \frac{2V}{y^2} = 0 \end{cases}$$

soit, tenant compte de $x, y > 0$,

$$x = y = \sqrt[3]{2V} \quad \text{et} \quad z = \sqrt[3]{\frac{V}{4}} = \frac{x}{2}$$

Pour vérifier qu'il s'agit bien d'un minimum, on calcule aisément

$$\frac{\partial^2 A}{\partial x^2} = \frac{4V}{x^3}, \quad \frac{\partial^2 A}{\partial y^2} = \frac{4V}{y^3}, \quad \frac{\partial^2 A}{\partial x \partial y} = \frac{\partial^2 A}{\partial y \partial x} = 1$$

et

$$H(\sqrt[3]{2V}, \sqrt[3]{2V}) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

En appliquant le critère de Sylvester, on constate que cette matrice est définie positive puisque

$$2 > 0, \quad \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3 > 0$$

et le point stationnaire identifié est bien un minimum local.

Le problème admettant une solution et la fonction $A(x, y)$ ne possédant qu'un seul point stationnaire, celui correspond au minimum recherché. \diamond

Mathématiquement, un problème d'optimisation avec contraintes d'égalité peut être écrit sous la forme

$$\mathcal{P} \begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ \text{s.c. } g_k(x) = 0, \quad \forall k \in \{1, 2, \dots, \ell\} \end{cases} \quad (3.112)$$

ou

$$\mathcal{P} \begin{cases} \max_{x \in \mathbb{R}^n} f(x) \\ \text{s.c. } g_k(x) = 0, \quad \forall k \in \{1, 2, \dots, \ell\} \end{cases} \quad (3.113)$$

où la notation "s.c." est l'abréviation de l'expression "sous contrainte". La fonction f est appelée la *fonction cible* ou *fonction objectif*. Les points qui vérifient les contraintes $g_k(x) = 0$ pour tout $k \in \{1, 2, \dots, \ell\}$ sont dits *admissibles*. Pour qu'un problème d'optimisation ait un sens, que la solution existe et qu'elle ne soit pas complètement déterminée par les seules contraintes d'égalité, on aura généralement $\ell < n$.

Dans les cas où les contraintes d'égalités $g_k(x) = 0$ peuvent être résolues par rapport à ℓ variables x_1, x_2, \dots, x_ℓ , ces relations peuvent être substituées dans l'expression dont on cherche les extrema pour éliminer ces ℓ variables. Le problème se ramène alors à la recherche des extrema d'une fonction des $n - \ell$ variables $x_{\ell+1}, x_{\ell+2}, \dots, x_n$ (Cf. exemple

3.41). Cependant, cette méthode n'est pas toujours applicable et peut s'avérer très lourde, en particulier lorsqu'elle conduit à rompre la symétrie par rapport aux différentes variables du problème initial.

L'énoncé suivant fournit une méthode alternative particulièrement élégante.

Si les fonctions $f, g_1, g_2, \dots, g_\ell$ définissant le problème d'optimisation avec contraintes d'égalité \mathcal{P} sont différentiables dans un voisinage de la solution x_* de \mathcal{P} et si la matrice

$$G_* = (\nabla g_1(x_*) \quad \nabla g_2(x_*) \quad \cdots \quad \nabla g_\ell(x_*)) \quad (3.114)$$

est de rang maximal, *i.e.* si les gradients des contraintes sont linéairement indépendants en x_* , alors il existe $\Lambda_* = (\lambda_{1*}, \lambda_{2*}, \dots, \lambda_{\ell*}) \in \mathbb{R}^\ell$ tel que (x_*, Λ_*) est un point stationnaire du *Lagrangien*¹¹

$$L(x, \Lambda) = f(x) - \sum_{k=1}^{\ell} \lambda_k g_k(x)$$

Démontrons cet énoncé dans le cas où le problème d'optimisation ne possède qu'une seule contrainte d'égalité $g(x) = 0$ et où f présente un maximum local au point x_* en lequel $\nabla g(x_*) \neq 0$.

En l'absence de la contrainte $g(x) = 0$, la différentielle de f au point x_* est nulle, *i.e.*

$$df(x_*) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_*) dx_i = 0$$

quels que soient les accroissements dx_i . En effet, la fonction f étant différentiable, il vient, pour tout dx tel que $x_* + dx$ reste dans un certain voisinage de x_* ,

$$f(x_* + dx) - f(x_*) = df(x_*, dx) + o(|dx|) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_*) dx_i + o(|dx|) \leq 0$$

En considérant successivement les accroissements dx du type $dx e_k$ et $-dx e_k$ (pour $k = 1, 2, \dots, n$) on en déduit que

$$\begin{cases} \frac{\partial f}{\partial x_k}(x_*) dx + o(dx) \leq 0 \\ -\frac{\partial f}{\partial x_k}(x_*) dx + o(dx) \leq 0 \end{cases} \Rightarrow \frac{\partial f}{\partial x_k}(x_*) = 0$$

11. De façon alternative, le Lagrangien est parfois défini par $L(x, \Lambda) = f(x) + \sum_{k=1}^{\ell} \lambda_k g_k(x)$.

Ceci est sans influence sauf lorsqu'il s'agit d'examiner la signification pratique des multiplicateurs de Lagrange (Cf. page 325).

On retrouve donc la condition nécessaire d'annulation des dérivées partielles.

En présence de la contrainte $g(x) = 0$, on a toujours

$$f(x_* + dx) - f(x_*) = df(x_*, dx) + o(|dx|) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_*) dx_i + o(|dx|) \leq 0$$

mais uniquement pour des accroissements dx compatibles avec la contrainte $g(x) = 0$, c'est-à-dire tels que

$$dg(x_*, dx) = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(x_*) dx_i = 0 \quad (*)$$

Seuls $n - 1$ accroissements élémentaires dx_i peuvent donc être choisis librement.

Si dx vérifie (*), il en est de même de $-dx$. On a alors

$$df(x_*, dx) \leq 0 \quad \text{et} \quad df(x_*, -dx) = -df(x_*, dx) \leq 0$$

et on retrouve la condition d'annulation de la différentielle

$$df(x_*, dx) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_*) dx_i = 0$$

En introduisant un nombre λ pour l'instant quelconque, il vient

$$d(f - \lambda g)(x_*, dx) = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x_*) - \lambda \frac{\partial g}{\partial x_i}(x_*) \right) dx_i = 0$$

Supposons $\frac{\partial g}{\partial x_1}(x_*) \neq 0$, (on peut toujours trouver une dérivée partielle non nulle en vertu de l'hypothèse $\nabla g(x_*) \neq 0$). On peut alors choisir $\lambda = \lambda_*$ tel que

$$\frac{\partial f}{\partial x_1}(x_*) - \lambda_* \frac{\partial g}{\partial x_1}(x_*) = 0$$

de sorte que l'expression

$$d(f - \lambda_* g) = \sum_{i=2}^n \left(\frac{\partial f}{\partial x_i}(x_*) - \lambda_* \frac{\partial g}{\partial x_i}(x_*) \right) dx_i = 0$$

ne comporte plus que $n - 1$ accroissements dx_i que l'on peut supposer indépendants. Il vient dès lors

$$\frac{\partial f}{\partial x_i}(x_*) - \lambda_* \frac{\partial g}{\partial x_i}(x_*) = 0 \quad \forall i \in \{2, 3, \dots, n\}$$

En introduisant le Lagrangien $L(x, \lambda) = f(x) - \lambda g(x)$ et en regroupant les résultats précédents, on trouve donc

$$\frac{\partial L}{\partial x_i}(x_*, \lambda_*) = 0 \quad \forall i \in \{1, 2, \dots, n\}$$

De plus,

$$\frac{\partial L}{\partial \lambda}(x_*, \lambda_*) = -g(x_*) = 0$$

Dès lors, (x_*, λ_*) est un point stationnaire de $L(x, \lambda)$.

□

Les coefficients $\lambda_1, \lambda_2, \dots, \lambda_\ell$ introduits dans l'énoncé ci-dessus sont appelés les *multiplicateurs de Lagrange*.

La condition sur le rang de la matrice G_* est appelée *l'hypothèse de qualification des contraintes*. Elle demande que les gradients des contraintes g_1, g_2, \dots, g_ℓ évalués en x_* soient linéairement indépendants. Dans le cas où le problème ne comporte qu'une contrainte d'égalité $g(x) = 0$, cette hypothèse se traduit par

$$\nabla g(x_*) \neq 0 \tag{3.115}$$

Cette condition exprime alors la possibilité de définir la normale en x_* à la surface (dans \mathbb{R}^3), la courbe (dans \mathbb{R}^2) ou l'hyper-surface (dans \mathbb{R}^n avec $n > 3$)

$$\Sigma = \{x \in \mathbb{R}^n : g(x) = 0\} \tag{3.116}$$

composée par les points admissibles.

L'énoncé ci-dessus peut être aisément justifié géométriquement lorsque $\ell = 1$ (Fig. 3.13). Dans ce cas, on a

$$L(x, \lambda) = f(x) - \lambda g(x) \tag{3.117}$$

et la condition de stationnarité du Lagrangien, fonction de $n + 1$ variables, exprimée dans l'énoncé conduit à

$$\begin{cases} \nabla f(x_*) = \lambda_* \nabla g(x_*) \\ g(x_*) = 0 \end{cases} \tag{3.118}$$

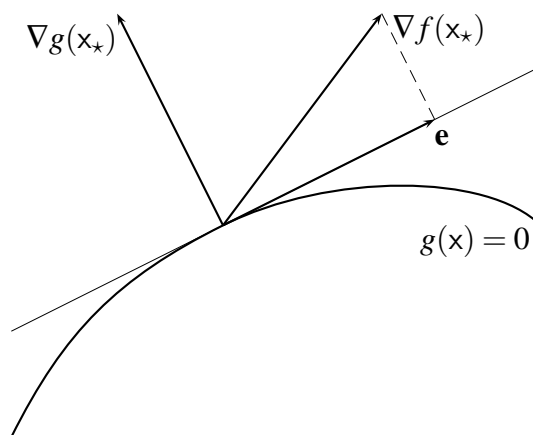


FIGURE 3.13

La deuxième condition exprime simplement que le point x_* est admissible. La première condition demande que les gradients de f et g , évalués à l'optimum x_* , soient parallèles. Comme $\nabla g(x_*)$ est normal à l'(hyper-)surface $g(x) = 0$, il doit en être de même de $\nabla f(x_*)$. Dans le cas contraire, la projection du vecteur $\nabla f(x_*)$ sur le plan tangent à l'(hyper-)surface $g(x) = 0$ serait un vecteur non nul \mathbf{e} tel que

$$D_{\mathbf{e}}f(x_*) = \mathbf{e} \cdot \nabla f(x_*) > 0 \quad \text{et} \quad D_{-\mathbf{e}}f(x_*) = -\mathbf{e} \cdot \nabla f(x_*) < 0$$

et f ne pourrait être extrémale en x_* . Ainsi donc, la stationnarité du Lagrangien apparaît comme une condition nécessaire d'extrémalité.

Le résultat de la page 318 permet d'aborder un problème d'optimisation \mathcal{P} avec contraintes d'égalité comme un problème d'optimisation non contraint dans un espace de dimension $n + \ell$. Il suffit de rechercher la solution parmi les points stationnaires du Lagrangien

$$L(x, \Lambda) = f(x) - \sum_{k=1}^{\ell} \lambda_k g_k(x)$$

Le Lagrangien étant une fonction de $n + \ell$ variables, les conditions de stationnarité conduisent à un système de $n + \ell$ équations pour les $n + \ell$ variables (x_*, Λ_*) .

EXEMPLE 3.42 Reprenons l'exemple 3.41 en utilisant l'approche développée dans cette section.

On constate que les fonctions

$$f(x, y, z) = xy + 2xz + 2yz \quad \text{et} \quad g(x, y, z) = xyz - V$$

vérifient les hypothèses du théorème. Dès lors, si l'optimum existe, il correspond à un point stationnaire du Lagrangien

$$L(x, y, z, \lambda) = xy + 2xz + 2yz - \lambda(xyz - V)$$

On forme donc le système de 4 équations à 4 inconnues

$$\begin{cases} \frac{\partial L}{\partial x} = y + 2z - \lambda yz = 0 \\ \frac{\partial L}{\partial y} = x + 2z - \lambda xz = 0 \\ \frac{\partial L}{\partial z} = 2x + 2y - \lambda xy = 0 \\ \frac{\partial L}{\partial \lambda} = -xyz + V = 0 \end{cases}$$

Soustrayant membre à membre les deux premières équations et remarquant qu'aucune des variables ne peut s'annuler à l'optimum, on obtient $x = y$, résultat qui pouvait aussi se déduire de la symétrie du problème.

Les deux dernières équations conduisent dès lors à

$$\lambda = \frac{4}{x}, \quad z = \frac{V}{x^2}$$

En substituant dans la première équation, on obtient, comme précédemment,

$$x_* = y_* = \sqrt[3]{2V} \quad \text{puis} \quad z_* = \sqrt[3]{\frac{V}{4}}$$

◇

EXEMPLE 3.43 Déterminons la distance de l'origine du plan à la courbe d'équation cartésienne

$$x^2y = 16$$

Le tracé (Fig. 3.14) de la courbe $x^2y = 16$ permet de se rendre compte de l'existence de deux points symétriques par rapport à l'axe OY et correspondant à la distance minimale recherchée.

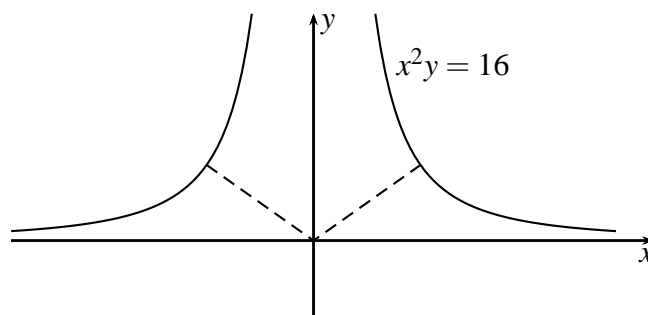


FIGURE 3.14

Le problème peut être exprimé comme suit, puisqu'il est équivalent de minimiser la distance considérée ou son carré,

$$\begin{aligned} \min \quad & f(x,y) = x^2 + y^2 \\ \text{s.c.} \quad & g(x,y) = x^2y - 16 = 0 \end{aligned}$$

Formons le Lagrangien

$$\begin{aligned} L(x,y,\lambda) &= f(x,y) - \lambda g(x,y) \\ &= x^2 + y^2 - \lambda(x^2y - 16) \end{aligned}$$

et recherchons les points stationnaires :

$$\begin{cases} \frac{\partial L}{\partial x} = 2x - 2\lambda xy = 0 \\ \frac{\partial L}{\partial y} = 2y - \lambda x^2 = 0 \\ \frac{\partial L}{\partial \lambda} = -x^2y + 16 = 0 \end{cases}$$

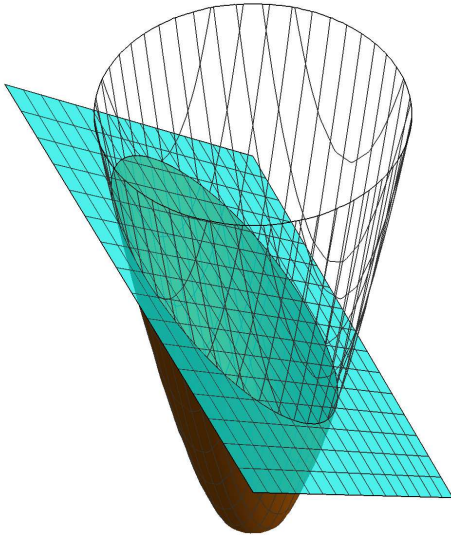
La solution $x = 0$ vérifiant la première équation ne pouvant être retenue, on tire de cette équation la relation $\lambda = 1/y$. De la troisième équation, on déduit $x^2 = 16/y$. Injectant ces expressions dans la deuxième équation, il vient

$$2y - \left(\frac{1}{y}\right) \frac{16}{y} = 0 \quad \Rightarrow \quad y = 2$$

On a donc $(x_*, y_*) = (\pm 2\sqrt{2}, 2)$ et la distance recherchée est égale à $\sqrt{f(x_*, y_*)} = 2\sqrt{3}$. ◇

EXEMPLE 3.44 Considérons le problème d'optimisation

$$\begin{aligned} \min \quad & f(x,y,z) = z \\ \text{s.c.} \quad & g_1(x,y,z) = x + y + z - 12 = 0 \\ & g_2(x,y,z) = x^2 + y^2 - z = 0 \end{aligned}$$



L'ensemble E des points admissibles est l'intersection du parabolôide $x^2 + y^2 = z$ et du plan $x + y + z = 12$. La fonction cible étant continue sur ce compact, elle réalise ses bornes supérieure et inférieure sur E . Le minimum recherché existe donc bien.

Les fonctions f , g_1 et g_2 sont (indéfiniment) continûment dérivables.

Pour examiner l'hypothèse de qualification des contraintes, on forme la matrice

$$G(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_2}{\partial x} \\ \frac{\partial g_1}{\partial y} & \frac{\partial g_2}{\partial y} \\ \frac{\partial g_1}{\partial z} & \frac{\partial g_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 1 & 2x \\ 1 & 2y \\ 1 & -1 \end{pmatrix}$$

Celle-ci est de rang maximal (=2) sauf si $x = y = -1/2$. Il n'existe cependant aucun point admissible avec de telles valeurs de x et y puisque les deux contraintes conduisent alors à des valeurs incompatibles de z . Dès lors, $\rho(G(x_*)) = 2$ et l'hypothèse de qualification des contraintes est vérifiée.

Le minimum recherché est donc associé à un point stationnaire du Lagrangien

$$L(x,y,z,\lambda_1,\lambda_2) = z - \lambda_1(x + y + z - 12) - \lambda_2(x^2 + y^2 - z)$$

Pour identifier ces points, formons le système

$$\begin{cases} \frac{\partial L}{\partial x} = -\lambda_1 - 2\lambda_2 x = 0 \\ \frac{\partial L}{\partial y} = -\lambda_1 - 2\lambda_2 y = 0 \\ \frac{\partial L}{\partial z} = 1 - \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial \lambda_1} = 12 - x - y - z = 0 \\ \frac{\partial L}{\partial \lambda_2} = z - x^2 - y^2 = 0 \end{cases}$$

Ce système possède les deux solutions

$$(x, y, z, \lambda_1, \lambda_2) = (2, 2, 8, 4/5, -1/5) \quad \text{et} \quad (x, y, z, \lambda_1, \lambda_2) = (-3, -3, 18, 6/5, 1/5)$$

L'évaluation de f montre que la première solution correspond au maximum de la fonction objectif, tandis que la seconde fournit le minimum recherché. \diamond

Dans l'application de la méthode des multiplicateurs de Lagrange, il convient d'être attentif à la signification précise de l'énoncé de la page 318.

D'une part, il faut se garder de penser que tous les points stationnaires du Lagrangien correspondent à des valeurs extrémales du problème d'optimisation \mathcal{P} .

D'autre part, les hypothèses du théorème ne doivent pas être oubliées. En toute généralité, la solution du problème d'optimisation \mathcal{P} , si elle existe, appartient à l'une des trois catégories suivantes :

- i. les points stationnaires du Lagrangien ;
- ii. les points où le Lagrangien n'est pas différentiable ;
- iii. les points où l'hypothèse de qualification des contraintes n'est pas vérifiée.

EXEMPLE 3.45 Soit le problème

$$\begin{aligned} \min \quad & f(x, y) = y \\ \text{s.c.} \quad & g(x, y) = x^2 - y^3 = 0 \end{aligned}$$

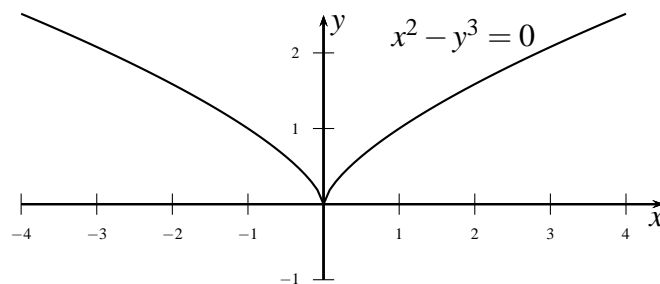


FIGURE 3.15

Recherchons les points stationnaires du Lagrangien

$$L(x, y, \lambda) = y - \lambda(x^2 - y^3)$$

en formant le système

$$\begin{cases} \frac{\partial L}{\partial x} = -2\lambda x = 0 \\ \frac{\partial L}{\partial y} = 1 + 3\lambda y^2 = 0 \\ \frac{\partial L}{\partial \lambda} = y^3 - x^2 = 0 \end{cases}$$

Ce système d'équations ne possède pas de solution. Pourtant le problème d'optimisation possède bel et bien une solution : on vérifie aisément que la contrainte $g(x,y) = 0$ définit la courbe représentée sur la Fig. 3.15 et que le minimum de f est donc atteint à l'origine et vaut 0.

Cette solution n'apparaît pas parmi les points stationnaires du Lagrangien mais parmi les points où l'hypothèse de qualification des contraintes n'est pas vérifiée. En effet, on a

$$\nabla g(0,0) = \begin{pmatrix} 2x \\ -3y^2 \end{pmatrix}_{(0,0)} = 0$$

◇

Interprétation des multiplicateurs de Lagrange.

En plus de fournir une méthode élégante pour prendre en compte des contraintes d'égalité, les multiplicateurs de Lagrange possèdent une interprétation intéressante. Pour établir celle-ci, considérons le problème d'optimisation

$$\begin{aligned} \min f(x) \\ \text{s.c. } g(x) = c \end{aligned} \tag{3.119}$$

où c désigne une constante et le Lagrangien correspondant écrit sous la forme

$$L(x, \lambda) = f(x) - \lambda [g(x) - c] \tag{3.120}$$

La solution du problème d'optimisation (3.119) (si elle existe) et le point stationnaire (x_*, λ_*) de L qui y correspond dépendent de la valeur de la constante c définissant la contrainte. Supposant la dérivabilité de x_* et λ_* par rapport à c , on calcule

$$\frac{d}{dc} L(x_*(c), \lambda_*(c)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial x_{i*}}{\partial c} - \frac{d\lambda_*}{dc} [g(x_*) - c] - \lambda_* \sum_{i=1}^n \frac{\partial g}{\partial x_i} \frac{\partial x_{i*}}{\partial c} + \lambda_* \tag{3.121}$$

où toutes les dérivées partielles sont évaluées en $x = x_*$ et $\lambda = \lambda_*$. Le Lagrangien étant stationnaire en (x_*, λ_*) , on a

$$\begin{cases} \frac{\partial f}{\partial x_i}(x_*) = \lambda_* \frac{\partial g}{\partial x_i}(x_*) \\ g(x_*) = c \end{cases} \tag{3.122}$$

de sorte que

$$\frac{dL_\star}{dc} = \lambda_\star \tag{3.123}$$

$$\text{où } L_\star \equiv L(x_\star, \lambda_\star) = f(x_\star) - \lambda [g(x_\star) - c] = f(x_\star)$$

Ceci montre que le multiplicateur de Lagrange mesure la sensibilité de la valeur optimale de la fonction cible f à la valeur de la contrainte.¹²

Ce résultat peut être aisément étendu aux problèmes d'optimisation comportant un nombre ℓ quelconque de contraintes d'égalité.

Si le Lagrangien

$$L(x, \Lambda) = f(x) - \sum_{k=1}^{\ell} \lambda_k [g_k(x) - c_k]$$

associé au problème d'optimisation

$$\mathcal{P} \begin{cases} \max (\text{resp. } \min) f(x) \\ \text{s.c. } g_k(x) = c_k, \quad \forall k \in \{1, 2, \dots, \ell\} \end{cases} \tag{3.124}$$

est stationnaire en (x_\star, Λ_\star) , alors

$$\lambda_{k\star} = \frac{d}{dc_k} f(x_\star)$$

EXEMPLE 3.46 Considérons à nouveau l'exemple 3.41 relatif à la minimisation de l'aire latérale de la boîte parallélépipédique sans couvercle de volume V donné.

En injectant les dimensions optimales

$$(x_\star, y_\star, z_\star) = (\sqrt[3]{2V}, \sqrt[3]{2V}, \sqrt[3]{V/4})$$

dans l'expression de l'aire latérale, on trouve

$$A(x_\star, y_\star, z_\star) = 3\sqrt[3]{4V^2}$$

On vérifie aisément que

$$\frac{d}{dV} A(x_\star, y_\star, z_\star) = 2\sqrt[3]{4/V} = \lambda_\star$$

Très logiquement, si on désire augmenter la capacité V de la boîte, il faut en augmenter l'aire latérale, ce que traduit la valeur positive de λ_\star .

◇

12. On remarquera que pour que cette interprétation soit correcte, il faut que le Lagrangien soit défini par (3.120) et non pas exprimé sous la forme $L(x, \lambda) = f(x) + \lambda [g(x) - C]$.

Condition suffisante d'extrémalité.

Des conditions suffisantes d'extrémalité semblables à celles obtenues dans le cadre de l'optimisation non contrainte peuvent être obtenues et utilisées pour caractériser les points stationnaires du Lagrangien. Les principales différences sont liées au fait que tous les accroissements Δx ne sont pas admissibles et qu'il est nécessaire de tenir compte de la courbure des contraintes.

Des conditions suffisantes semblables à celles de la page 306 utilisent la matrice hessienne $H_{Lx,\star}$ du Lagrangien L par rapport à ses seules variables x évaluées en (x_\star, Λ_\star) , *i.e.*

$$[H_{Lx,\star}]_{i,j} = \frac{\partial^2 L}{\partial x_i \partial x_j}(x_\star, \Lambda_\star) \quad i, j \in \{1, 2, \dots, n\} \quad (3.125)$$

et l'application linéaire définie par la matrice des gradients des contraintes en x_\star

$$G_\star^T = \begin{pmatrix} \nabla g_1(x_\star)^T \\ \nabla g_2(x_\star)^T \\ \vdots \\ \nabla g_\ell(x_\star)^T \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x_\star) & \frac{\partial g_1}{\partial x_2}(x_\star) & \cdots & \frac{\partial g_1}{\partial x_n}(x_\star) \\ \frac{\partial g_2}{\partial x_1}(x_\star) & \frac{\partial g_2}{\partial x_2}(x_\star) & \cdots & \frac{\partial g_2}{\partial x_n}(x_\star) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_\ell}{\partial x_1}(x_\star) & \frac{\partial g_\ell}{\partial x_2}(x_\star) & \cdots & \frac{\partial g_\ell}{\partial x_n}(x_\star) \end{pmatrix} \quad (3.126)$$

Soient $f, g_1, g_2, \dots, g_\ell$ deux fois continûment dérivables au voisinage de x_\star telles que le Lagrangien $L(x, \Lambda)$ est stationnaire en (x_\star, λ_\star) .

- Si $H_{Lx,\star}$ est *définie positive* sur le noyau de G_\star^T , *i.e.* si

$$\Delta x^T H_{Lx,\star} \Delta x > 0 \quad \forall \Delta x \neq 0 \text{ tel que } G_\star^T \Delta x = 0$$

alors f admet un minimum sous contrainte en x_\star .

- Si $H_{Lx,\star}$ est *définie négative* sur le noyau de G_\star^T , *i.e.* si

$$\Delta x^T H_{Lx,\star} \Delta x < 0 \quad \forall \Delta x \neq 0 \text{ tel que } G_\star^T \Delta x = 0$$

alors f admet un maximum sous contrainte en x_\star .

- Si $H_{Lx,\star}$ est *indéfinie* sur le noyau de G_\star^T , *i.e.* si

$$\exists \Delta x_1, \Delta x_2 \neq 0 \text{ tels que } G_\star^T \Delta x_1 = G_\star^T \Delta x_2 = 0 \text{ et}$$

$$\Delta x_1^T H_{Lx,\star} \Delta x_1 < 0, \quad \Delta x_2^T H_{Lx,\star} \Delta x_2 > 0$$

alors f n'admet ni minimum ni maximum sous contrainte en x_\star .

Afin de pouvoir caractériser la nature du point x_* correspondant à un point stationnaire (x_*, λ_*) du Lagrangien, il convient de pouvoir comparer les valeurs prises par la fonction cible f au point x_* et aux points admissibles $x_* + \Delta x$ dans un voisinage de x_* .

Par application de la formule de Taylor à la fonction L considérée comme une fonction de x uniquement, *i.e.* en traitant Λ_* comme un paramètre constant, et en tenant compte de la stationnarité du Lagrangien en (x_*, Λ_*) , on a

$$\begin{aligned} L(x_* + \Delta x, \Lambda_*) &= L(x_*, \Lambda_*) + \sum_{i=1}^n \frac{\partial L}{\partial x_i}(x_*, \Lambda_*) \Delta x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L}{\partial x_i \partial x_j}(\tilde{x}, \Lambda_*) \Delta x_i \Delta x_j \\ &= L(x_*, \Lambda_*) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L}{\partial x_i \partial x_j}(\tilde{x}, \Lambda_*) \Delta x_i \Delta x_j \end{aligned}$$

où $\tilde{x} = x_* + \theta \Delta x$ avec $\theta \in]0, 1[$.

Cette expression décrit également les variations de f puisque, si x_* et $x_* + \Delta x$ sont admissibles, alors

$$g_k(x_* + \Delta x) = g_k(x_*) = 0 \quad \forall k \in \{1, 2, \dots, \ell\} \quad (\ddagger)$$

de sorte que

$$L(x_* + \Delta x, \lambda_*) = f(x_* + \Delta x), \quad L(x_*, \lambda_*) = f(x_*)$$

et

$$f(x_* + \Delta x) - f(x_*) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L}{\partial x_i \partial x_j}(\tilde{x}, \lambda_*) \Delta x_i \Delta x_j \quad (\dagger)$$

Par continuité, si $\Delta x^T H_{Lx,*} \Delta x \neq 0$, le signe de (\dagger) est celui de $\Delta x^T H_{Lx,*} \Delta x$ pour Δx suffisamment petit, ce qui conduit au résultat énoncé. Par exemple, si $\Delta x^T H_{Lx,*} \Delta x > 0$ [resp. < 0] pour tout $\Delta x \neq 0$, alors

$$f(x_* + \Delta x) > f(x_*) \quad \left[\text{resp. } f(x_* + \Delta x) < f(x_*) \right]$$

et f présente un minimum [resp. maximum] local en x_* par rapport aux autres points admissibles au voisinage de ce point.

Si on peut trouver Δx_1 et Δx_2 non nuls tels que $\Delta x_1^T H_{Lx,*} \Delta x_1 > 0$ et $\Delta x_2^T H_{Lx,*} \Delta x_2 < 0$

$$f(x_* + \Delta x_1) > f(x_*) \quad \text{et} \quad f(x_* + \Delta x_2) < f(x_*)$$

de sorte que f ne présente ni un minimum ni un maximum local en x_* par rapport aux autres points admissibles au voisinage de ce point.

Dans l'examen des variations de f au voisinage de x_* , tous les accroissements Δx ne doivent pas être pris en compte. Il convient en effet de ne retenir que les accroissements Δx qui font de $x_* + \Delta x$ un point admissible. Injectant la condition (\ddagger) dans

$$g_k(x_* + \Delta x) = g_k(x_*) + \nabla g_k(x_*)^T \Delta x + O(\|\Delta x\|^2), \quad \forall k \in \{1, 2, \dots, \ell\}$$

on conclut que les accroissements admissibles sont caractérisés par

$$\nabla g_k(x_*)^T \Delta x = 0, \quad \forall k \in \{1, 2, \dots, \ell\}$$

soit, matriciellement,

$$G_*^T \Delta x = 0$$

□

Les conditions suffisantes ci-dessus demandent de former explicitement le noyau de G_*^T , ce qui n'est généralement pas aisé. On peut éviter cette difficulté en utilisant une variante du critère de Sylvester exprimant les conditions suffisantes établies plus haut à partir de la matrice hessienne complète du Lagrangien évaluée au point stationnaire, *i.e.*

$$H_{B_*} = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} & -\frac{\partial g_1}{\partial x_1} & \cdots & -\frac{\partial g_\ell}{\partial x_1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial^2 L}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 L}{\partial x_n^2} & -\frac{\partial g_1}{\partial x_n} & \cdots & -\frac{\partial g_\ell}{\partial x_n} \\ -\frac{\partial g_1}{\partial x_1} & \cdots & -\frac{\partial g_1}{\partial x_n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ -\frac{\partial g_\ell}{\partial x_1} & \cdots & -\frac{\partial g_\ell}{\partial x_n} & 0 & \cdots & 0 \end{pmatrix}_{(x_*, \Lambda_*)} = \begin{pmatrix} H_{Lx, *} & -G_* \\ -G_*^T & 0 \end{pmatrix} \quad (3.127)$$

La matrice H_{B_*} est formée à partir de la matrice $H_{Lx, *}$ en lui ajoutant des blocs associés à G_* et est dès lors appelée la *matrice hessienne bordée*. Elle contient toutes les informations nécessaires pour mettre en oeuvre les conditions suffisantes établies ci-dessus.

Soit (x_*, Λ_*) un point stationnaire du Lagrangien.

- Si les $n - \ell$ derniers mineurs principaux de H_{B_*} sont alternativement strictement positifs et strictement négatifs, le dernier d'entre-eux étant de même signe que $(-1)^n$, alors x_* est un maximum local.
- Si les $n - \ell$ derniers mineurs principaux de H_{B_*} sont tous de même signe que $(-1)^\ell$, alors x_* est un minimum local.
- Si les $n - \ell$ derniers mineurs principaux de H_{B_*} sont non nuls et ne vérifient aucune des deux règles de signes ci-dessus, alors x_* est un point de selle.

La démonstration peut être trouvée dans [17].

□

Les $n - \ell$ derniers mineurs principaux utilisés dans ce test sont les déterminants des matrices formées des éléments de H_{B^*} situés à l'intersection des r dernières lignes et r dernières colonnes de H_{B^*} avec $r \in \{2\ell + 1, \dots, n + \ell\}$. Les exemples qui suivent en fournissent des illustrations.

Remarquons que le nombre de mineurs principaux à prendre en compte est égal à la dimension $n - \ell$ du noyau de G_*^T , *i.e.* au nombre de directions indépendantes selon lesquelles on peut se déplacer sur la surface définie par les contraintes d'égalité (*directions admissibles*).

EXEMPLE 3.47 Poursuivons l'exemple 3.42 étudié précédemment.

Partant de l'expression de l'unique contrainte

$$g(x, y, z) = xyz - V = 0$$

et du Lagrangien

$$L(x, y, z, \lambda) = xy + 2xz + 2yz - \lambda(xyz - V)$$

on peut former la matrice hessienne bordée

$$H_B = \begin{pmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y} & \frac{\partial^2 L}{\partial x \partial z} & -\frac{\partial g}{\partial x} \\ \frac{\partial^2 L}{\partial y \partial x} & \frac{\partial^2 L}{\partial y^2} & \frac{\partial^2 L}{\partial y \partial z} & -\frac{\partial g}{\partial y} \\ \frac{\partial^2 L}{\partial z \partial x} & \frac{\partial^2 L}{\partial z \partial y} & \frac{\partial^2 L}{\partial z^2} & -\frac{\partial g}{\partial z} \\ -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} & -\frac{\partial g}{\partial z} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 - \lambda z & 2 - \lambda y & -yz \\ 1 - \lambda z & 0 & 2 - \lambda x & -xz \\ 2 - \lambda y & 2 - \lambda x & 0 & -xy \\ -yz & -xz & -xy & 0 \end{pmatrix}$$

Évaluant les différents éléments au point stationnaire

$$(x, y, z, \lambda) = (\sqrt[3]{2V}, \sqrt[3]{2V}, \sqrt[3]{V/4}, 2\sqrt[3]{4/V})$$

on obtient

$$H_{B^*} = \begin{pmatrix} 0 & -1 & -2 & -\sqrt[3]{V^2/2} \\ -1 & 0 & -2 & -\sqrt[3]{V^2/2} \\ -2 & -2 & 0 & -\sqrt[3]{4V^2} \\ -\sqrt[3]{V^2/2} & -\sqrt[3]{V^2/2} & -\sqrt[3]{4V^2} & 0 \end{pmatrix}$$

Puisque le problème est posé dans \mathbb{R}^3 ($n = 3$) et comporte une contrainte ($\ell = 1$), on doit examiner le signe des $3 - 1 = 2$ derniers mineurs principaux, *i.e.*

$$\begin{vmatrix} 0 & -2 & -\sqrt[3]{V^2/2} \\ -2 & 0 & -\sqrt[3]{4V^2} \\ -\sqrt[3]{V^2/2} & -\sqrt[3]{4V^2} & 0 \end{vmatrix} = -4V\sqrt[3]{2V} < 0$$

et

$$\begin{vmatrix} 0 & -1 & -2 & -\sqrt[3]{V^2/2} \\ -1 & 0 & -2 & -\sqrt[3]{V^2/2} \\ -2 & -2 & 0 & -\sqrt[3]{4V^2} \\ -\sqrt[3]{V^2/2} & -\sqrt[3]{V^2/2} & -\sqrt[3]{4V^2} & 0 \end{vmatrix} = -6V\sqrt[3]{2V} < 0$$

Ces deux mineurs principaux étant négatifs, soit du même signe que $(-1)^\ell = -1$, on en déduit que le point stationnaire considéré correspond à un minimum local du problème d'optimisation sous contrainte, *i.e.* l'aire latérale de la boîte est minimale pour les dimensions considérées. \diamond

EXEMPLE 3.48 Examinons la nature des points stationnaires du Lagrangien identifiés dans l'exemple 3.43. On a

$$g(x, y) = x^2y - 16$$

et

$$L(x, y, \lambda) = x^2 + y^2 - \lambda(x^2y - 16)$$

de sorte que

$$H_{B^*} = \begin{pmatrix} 2 - 2\lambda y & -2\lambda x & -2xy \\ -2\lambda x & 2 & -x^2 \\ -2xy & -x^2 & 0 \end{pmatrix}_{(\pm 2\sqrt{2}, 2, 1/2)} = \begin{pmatrix} 0 & \mp 2\sqrt{2} & \mp 8\sqrt{2} \\ \mp 2\sqrt{2} & 2 & -8 \\ \mp 8\sqrt{2} & -8 & 0 \end{pmatrix}$$

Le problème implique $n = 2$ variables et $\ell = 1$ contrainte. Il convient donc d'examiner seulement le signe du déterminant ($n - \ell = 1$) de la matrice hessienne bordée, soit

$$\begin{vmatrix} 0 & \mp 2\sqrt{2} & \mp 8\sqrt{2} \\ \mp 2\sqrt{2} & 2 & -8 \\ \mp 8\sqrt{2} & -8 & 0 \end{vmatrix} = -768 < 0$$

Le signe de ce mineur étant celui de $(-1)^\ell = -1$, on en déduit que les points stationnaires correspondent bien à un minimum et définissent la distance de l'origine à la courbe. \diamond

EXEMPLE 3.49 Caractérisons les points stationnaires du Lagrangien identifié dans l'exemple 3.44. Le problème d'optimisation comporte les deux contraintes

$$\begin{cases} g_1(x, y, z) = x + y + z - 12 = 0 \\ g_2(x, y, z) = x^2 + y^2 - z = 0 \end{cases}$$

Il est caractérisé par le Lagrangien

$$L(x, y, z, \lambda_1, \lambda_2) = z - \lambda_1(x + y + z - 12) - \lambda_2(x^2 + y^2 - z)$$

Dès lors, la matrice hessienne bordée s'écrit

$$H_B = \begin{pmatrix} -2\lambda_2 & 0 & 0 & -1 & -2x \\ 0 & -2\lambda_2 & 0 & -1 & -2y \\ 0 & 0 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 & 0 \\ -2x & -2y & 1 & 0 & 0 \end{pmatrix}$$

Le problème comportant $n = 3$ variables et $\ell = 2$ contraintes, la nature des points stationnaires peut être déterminée en examinant le signe du déterminant de la matrice.

Au point stationnaire $(-3, -3, 18, 6/5, 1/5)$, on a

$$\det H_{B^*} = \begin{vmatrix} -2/5 & 0 & 0 & -1 & 6 \\ 0 & -2/5 & 0 & -1 & 6 \\ 0 & 0 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 & 0 \\ 6 & 6 & 1 & 0 & 0 \end{vmatrix} = -20 < 0$$

qui est du même signe que $(-1)^n$, ce qui confirme que ce point stationnaire correspond à un maximum local de la fonction objectif.

Au point stationnaire $(2, 2, 8, 4/5, -1/5)$, on a

$$\det H_{B^*} = \begin{vmatrix} 2/5 & 0 & 0 & -1 & -4 \\ 0 & 2/5 & 0 & -1 & -4 \\ 0 & 0 & 0 & -1 & 1 \\ -1 & -1 & -1 & 0 & 0 \\ -4 & -4 & 1 & 0 & 0 \end{vmatrix} = 20 > 0$$

qui est du même signe que $(-1)^\ell$. Il s'agit donc d'un minimum local de la fonction objectif.

◇

Dans bien des cas, comme le montre l'exemple ci-dessous, on peut se passer des conditions suffisantes d'extrémalité pour caractériser les points stationnaires du Lagrangien. Si l'existence du maximum ou du minimum sous contrainte peut être établie par ailleurs, par exemple en utilisant des arguments physiques ou en exploitant le fait que l'ensemble des points admissibles forme un compact sur lequel la fonction cible est continue, on sait que le minimum ou le maximum recherché correspond soit à un point stationnaire du Lagrangien, soit à un point où le Lagrangien n'est pas dérivable, soit à un point où l'hypothèse de qualification des contraintes n'est pas vérifiée. Il suffit donc de déterminer ces points particuliers et de comparer les valeurs prises par la fonction cible en ceux-ci pour en trouver le minimum ou le maximum et ainsi obtenir la solution du problème d'optimisation.

Cette procédure ne permet d'identifier que les extrema absolus, pas de caractériser les minima et maxima locaux.

EXEMPLE 3.50 Montrons par un calcul d'optimisation que la moyenne géométrique d'un ensemble de valeurs positives est toujours inférieure ou égale à la moyenne arithmétique, *i.e.* que, pour ensemble de n nombres positifs x_1, x_2, \dots, x_n on a

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Le problème peut être abordé en déterminant la valeur maximale de la moyenne géométrique correspondant à une valeur fixée $\mu > 0$ de la moyenne arithmétique, ce qui est équivalent à la recherche du maximum de la fonction

$$f(x_1, x_2, \dots, x_n) = x_1 x_2 \cdots x_n$$

sur l'ensemble

$$E = \left\{ (x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0, g(x_1, x_2, \dots, x_n) = 0 \right\}$$

où

$$g(x_1, x_2, \dots, x_n) = x_1 + x_2 + \cdots + x_n - n\mu$$

Les contraintes de positivité des variables ne peuvent être vérifiées avec un signe d'égalité. En effet, si une variable x_i au moins est nulle, la moyenne géométrique est également nulle, ce qui correspond à la valeur minimale de f .

Ignorant les contraintes de positivité des variables, le Lagrangien correspondant à ce problème d'optimisation est donné par

$$L(x_1, x_2, \dots, x_n, \lambda) = x_1 x_2 \cdots x_n - \lambda(x_1 + x_2 + \cdots + x_n - n\mu)$$

Les points stationnaires sont les solutions de

$$\frac{\partial L}{\partial x_i} = \left(\prod_{\substack{j=1 \\ j \neq i}}^n x_j \right) - \lambda = 0, \quad i \in \{1, 2, \dots, n\}$$

$$\frac{\partial L}{\partial \lambda} = x_1 + x_2 + \cdots + x_n - n\mu = 0$$

La seule solution de ce système est

$$x_* = (x_1, x_2, \dots, x_n) = (\mu, \mu, \dots, \mu), \quad \lambda_* = \mu^{n-1}$$

Ce point stationnaire du Lagrangien correspond à des valeurs positives des variables et est donc admissible.

L'ensemble admissible E étant compact et la fonction f étant continue sur celui-ci, on sait que f réalise ses bornes supérieure et inférieure en des points de E . Le maximum recherché existe donc.

Puisque f et g sont partout différentiables et que $\nabla g \neq 0$, le maximum recherché correspond à un point stationnaire du Lagrangien. Comme (x_*, λ_*) est l'unique point stationnaire de L celui-ci correspond au maximum recherché.

En $x = x_*$, on a

$$\sqrt[n]{x_1 x_2 \cdots x_n} = \mu = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

ce qui démontre que la valeur maximale de la moyenne géométrique est égale à la moyenne arithmétique, comme annoncé initialement.

Notons que le calcul de la matrice hessienne bordée conduirait à

$$H_{B_*} = \begin{pmatrix} 0 & \alpha & \alpha & \cdots & \alpha & \alpha & -1 \\ \alpha & 0 & \alpha & \cdots & \alpha & \alpha & -1 \\ \alpha & \alpha & 0 & \cdots & \alpha & \alpha & -1 \\ \vdots & & & \ddots & & & \vdots \\ \alpha & \alpha & \alpha & & 0 & \alpha & -1 \\ \alpha & \alpha & \alpha & \cdots & \alpha & 0 & -1 \\ -1 & -1 & -1 & \cdots & -1 & -1 & 0 \end{pmatrix} \quad \text{où} \quad \alpha = \mu^{n-2} > 0$$

dont on peut montrer que les derniers mineurs principaux sont

$$(-1)^k k \alpha^{k-1} \quad k \in \{2, 3, \dots, n-1, n\}$$

Les signes de ces mineurs sont alternativement positifs et négatifs et le signe du déterminant de la matrice hessienne bordée complète est celui de $(-1)^n$. Ceci confirme que x_* est bien un maximum local de f sous la contrainte considérée. Ce calcul peut cependant être évité par le raisonnement mené ci-dessus. \diamond

3.10.5 Optimisation avec contraintes d'inégalité.

De nombreux problèmes d'optimisation impliquent la recherche du minimum ou du maximum d'une fonction f parmi les points appartenant à un ensemble admissible $E \subset \mathbb{R}^n$ d'intérieur non vide, ce qui peut s'écrire sous la forme

$$\min_{x \in E} f(x) \quad \text{ou} \quad \max_{x \in E} f(x) \quad (3.128)$$

Souvent, le domaine admissible E est défini par une ou plusieurs contraintes d'inégalité, par exemple,

$$E = \{x \in \mathbb{R}^n : g_1(x) \leq 0, g_2(x) \leq 0, \dots, g_\ell(x) \leq 0\} \quad (3.129)$$

où g_k ($i = 1, 2, \dots, \ell$) désignent des fonctions réelles. Nous supposons ici que les fonctions f et g_k ($k = 1, 2, \dots, \ell$) sont continûment dérivables.

Ces problèmes peuvent être abordés en deux temps en utilisant les outils développés dans les sections précédentes. Selon les conditions nécessaires établies à la section 3.10.1, l'extremum recherché x_* appartient à l'intérieur de l'ensemble E ou est situé sur la frontière de E .

D'une part, la recherche des points stationnaires de f et leur caractérisation en utilisant les conditions suffisantes de la section 3.10.2 permettent d'identifier les extrema locaux situés à l'intérieur de E .

D'autre part, les techniques développées à la section 3.10.4 permettent d'identifier les extrema locaux appartenant à la frontière de E . En effet, les points x appartenant à la frontière de E sont caractérisés par $g_k(x) = 0$ pour une ou plusieurs valeurs de i . Les contraintes correspondantes sont dites *actives*; les contraintes telles que $g_k(x_*) > 0$ sont *inactives*. Les contraintes actives peuvent donc être utilisées pour éliminer des variables explicitement ou en utilisant la technique des multiplicateurs de Lagrange.

Finalement, s'ils existent, le minimum et le maximum absolus recherchés peuvent être trouvés en rassemblant les résultats relatifs à l'intérieur de E et à sa frontière.¹³

EXEMPLE 3.51 La distribution de la température au sein d'une plaque métallique circulaire est décrite par l'expression

$$T(x, y) = T_0(x^2 + 2y^2 - x) \quad \text{pour} \quad x^2 + y^2 \leq 1$$

où T_0 est une constante strictement positive connue. Recherchons les températures extrêmes dans cette plaque.

Les extrema locaux situés à l'intérieur de la plaque sont parmi les points stationnaires de T et sont donc les solutions de

$$\begin{cases} \frac{\partial T}{\partial x} = T_0(2x - 1) = 0 \\ \frac{\partial T}{\partial y} = 4T_0y = 0 \end{cases}$$

Ce système possède la solution unique $(x_0, y_0) = (1/2, 0)$ appartenant à l'intérieur de la plaque. En ce point, comme en tout autre, la matrice hessienne est donnée par

$$H(x_0, y_0) = \begin{pmatrix} 2T_0 & 0 \\ 0 & 4T_0 \end{pmatrix}$$

qui est définie positive de sorte que le point $(1/2, 0)$ correspond à un minimum local de T avec

$$T(1/2, 0) = -\frac{T_0}{4}$$

Étudions maintenant les variations de la température sur le bord de la plaque circulaire. En ces points, on a $y^2 = 1 - x^2$ et la température est donc décrite par

$$T(x, \pm\sqrt{1-x^2}) = T_0(2-x-x^2), \quad x \in [-1, 1]$$

On sera attentif au fait que cette expression n'a de sens que pour $x \in [-1, 1]$ et doit donc être étudiée sur cet intervalle en tenant compte de la présence d'éventuels extrema aux extrémités de celui-ci.

13. On notera que si la fonction objectif est stationnaire en un point x_* appartenant à la frontière de E et si la matrice hessienne de f en ce point est définie positive (*resp.* définie négative), la fonction f présente un minimum (*resp.* un maximum) local en x_* indépendamment des restrictions imposées via E .

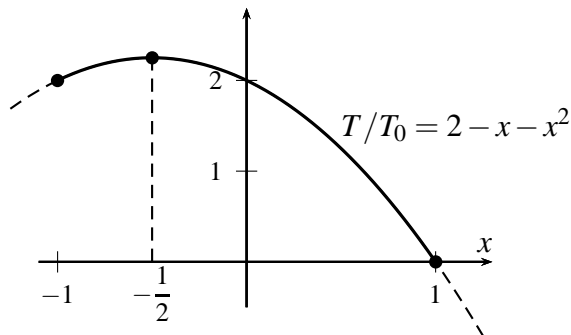


FIGURE 3.16

L'étude des variations de la température sur le bord de la plaque (Fig. 3.16), montre que la température passe par des minima locaux en $x = -1$ et en $x = +1$ et présente un maximum local en $x = -1/2$. En ces points caractéristiques, on a

$$T(-1, 0) = 2T_0, \quad T(-1/2, \pm\sqrt{3}/2) = \frac{9T_0}{4}, \quad T(1, 0) = 0$$

Physiquement, il est évident que la température doit atteindre une valeur minimale et une valeur maximale en des points particuliers de la plaque. Mathématiquement, l'existence de ces extrema globaux est assurée par le fait que la fonction $T(x, y)$ est continue sur le compact $x^2 + y^2 \leq 1$ et réalise donc ses bornes supérieure et inférieure en des points de celui-ci. Les extrema globaux devant également être des extrema locaux, il suffit de comparer les valeurs prises par la température aux points particuliers identifiés à l'intérieur de la plaque et sur son bord. On en déduit dès lors les résultats suivants :

- Température minimale $T_{min} = -T_0/4$ réalisée en $(1/2, 0)$;
- Température maximale $T_{max} = 9T_0/4$ réalisée en $(-1/2, \sqrt{3}/2)$ et en $(-1/2, -\sqrt{3}/2)$.

La figure Fig. 3.17 fournit une représentation des variations de la température au sein de la plaque qui permet de confirmer le résultat obtenu.

On remarquera que les points $(-1, 0)$ et $(1, 0)$ où la température présente une minimum local par rapport aux points du bord de la plaque ne correspondent pas à des minima locaux par rapport à l'ensemble des points de la plaque.

Si on se déplace exclusivement le long du bord de la plaque, la température passe effectivement par une valeur minimale au point $(1, 0)$. Cependant, si, à partir de ce point, on se dirige vers l'intérieur de la plaque, la température continue de diminuer. En effet, le calcul du gradient en ce point

$$\nabla T(1, 0) = T_0 \mathbf{e}_x$$

montre que la température décroît dans la direction de $-\mathbf{e}_x$ et que des températures inférieures à $T(1, 0) = 0$ sont donc rencontrées au voisinage de $(1, 0)$ en des points $(1 - \varepsilon, 0)$ pour $\varepsilon > 0$ suffisamment petit. Dès lors, la température ne présente pas d'extremum local en $(1, 0)$.

Les variations de la température au voisinage du point $(-1, 0)$ peuvent être menées de la même façon et conduisent à des conclusions identiques.

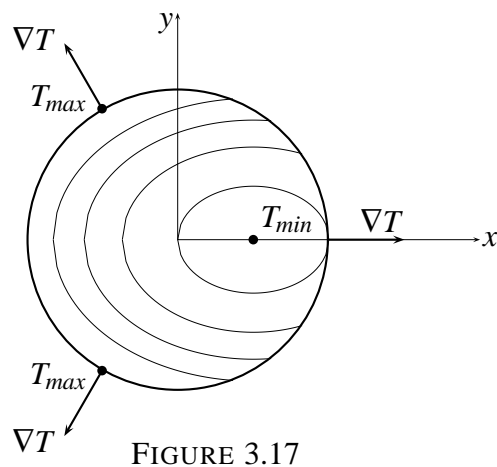


FIGURE 3.17

Aux deux points $(-1/2, \pm\sqrt{3}/2)$ où la température maximale $T_{max} = 9T_0/4$ est réalisée, le gradient de température est donné par

$$\nabla T(-1/2, \pm\sqrt{3}/2) = (-2\mathbf{e}_x \pm 2\sqrt{3}\mathbf{e}_y)T_0$$

qui est orienté selon la normale extérieure à la plaque, laquelle est parallèle au vecteur

$$\mathbf{s} = \frac{-\mathbf{e}_x \pm \sqrt{3}\mathbf{e}_y}{2}$$

Dès lors, tout déplacement effectué depuis les points $(-1/2, \pm\sqrt{3}/2)$ vers l'intérieur de la plaque se projette négativement sur le gradient et conduit à une diminution de la température, ce qui confirme l'existence de maxima aux points $(-1/2, \pm\sqrt{3}/2)$. \diamond

Dans l'exemple précédent, l'existence des extrema absolus constitue la clé de la justification du raisonnement. Si cette existence n'est pas avérée, la comparaison des valeurs de la fonction étudiée aux différents points remarquables identifiés à l'intérieur et sur la frontière de E ne suffit pas pour conclure.

EXEMPLE 3.52 Recherchons les extrema de la fonction

$$f(x, y) = x^2 + y^2$$

sur le domaine

$$E = \{(x, y) : (x - 1)^2 + (y - 1)^2 \geq 1\}$$

Le domaine E exclut les points du plan à l'intérieur du disque de centre (1,1) et de rayon unitaire. (Fig. 3.18).

Les points stationnaires de f sont les solutions de

$$\begin{cases} \frac{\partial f}{\partial x} = 2x = 0 \\ \frac{\partial f}{\partial y} = 2y = 0 \end{cases}$$

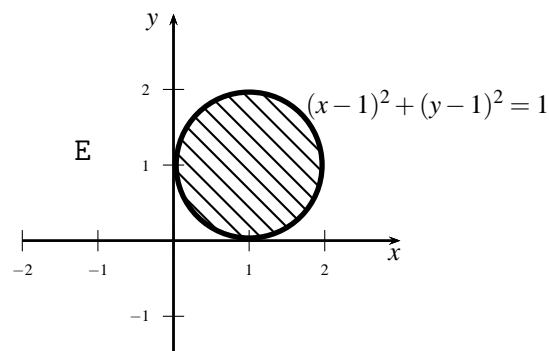


FIGURE 3.18

Ce système possède une solution unique $(0, 0) \in E$. On peut utiliser la matrice hessienne pour caractériser ce point stationnaire. Cependant, dans le cas présent, il est plus simple de constater que

$$f(0, 0) = 0 < f(x, y) \quad \forall (x, y) \neq (0, 0)$$

de sorte que la fonction f présente un minimum absolu sur E au point $(0, 0)$.

Afin d'identifier un éventuel maximum, étudions les variations de f sur la frontière de E . En ces points, la contrainte est active, *i.e.*

$$g(x, y) \equiv 1 - (x - 1)^2 - (y - 1)^2 = 0$$

Cette expression ne permettant pas d'éliminer aisément une variable, construisons le Lagrangien

$$L(x, y, \lambda) = x^2 + y^2 - \lambda [1 - (x - 1)^2 - (y - 1)^2]$$

Notant que f et g sont différentiables et que ∇g ne s'annule en aucun point de la frontière de E , les extrema de f appartenant à la frontière de E se trouvent parmi les points stationnaires du Lagrangien, *i.e.* parmi les solutions de

$$\begin{cases} \frac{\partial L}{\partial x} = 2x + 2\lambda(x - 1) = 0 \\ \frac{\partial L}{\partial y} = 2y + 2\lambda(y - 1) = 0 \\ \frac{\partial L}{\partial \lambda} = (x - 1)^2 + (y - 1)^2 - 1 = 0 \end{cases}$$

Des deux premières équations, on déduit $x = y$. Injectant ce résultat dans la dernière équation, on trouve les deux solutions

$$(x_1, y_1) = \left(\frac{\sqrt{2}-1}{\sqrt{2}}, \frac{\sqrt{2}-1}{\sqrt{2}} \right) \quad \text{et} \quad (x_2, y_2) = \left(\frac{\sqrt{2}+1}{\sqrt{2}}, \frac{\sqrt{2}+1}{\sqrt{2}} \right)$$

En ces points, on a

$$f(x_1, y_1) = (\sqrt{2}-1)^2 \quad \text{et} \quad f(x_2, y_2) = (\sqrt{2}+1)^2$$

qui sont respectivement les valeurs minimales et maximales prises par la fonction f parmi les points appartenant à la frontière de E . En effet, la fonction f est continue sur la frontière de E , qui est un compact, de sorte qu'elle y réalise ses bornes supérieure et inférieure.

En comparant les valeurs prises par la fonction f aux points particuliers identifiés à l'intérieur de E et sur sa frontière, on pourrait être tenté de conclure que le maximum de f vaut $(\sqrt{2}+1)^2$ et est réalisé au point (x_2, y_2) . Cependant, il n'en est rien : f n'est pas bornée supérieurement sur E et ne possède donc pas de maximum. Pour s'en convaincre, il suffit de considérer les points admissibles $(x, y) = (-\alpha^2, -\alpha^2) \in E$ pour lesquels

$$f(-\alpha^2, -\alpha^2) = 2\alpha^4$$

peut être rendu arbitrairement grand. L'ensemble admissible E n'étant pas compact (au contraire de sa frontière), on ne peut s'appuyer sur la continuité de f sur cet ensemble pour établir l'existence des extrema.

En conclusion, la fonction f présente son minimum absolu en $(0, 0)$, où elle s'annule, mais ne possède pas de maximum sur E . ◇

3.11 Fonctions implicites.

Une équation du type

$$0 = f(x, y) \tag{3.130}$$

définit implicitement une fonction $y(x)$ sur un sous-ensemble $E \subset \mathbb{R}$ si, $\forall x \in E$, l'équation (3.130) possède une solution y unique.

Dans ce cas, on est tenté d'écrire

$$\begin{aligned} 0 &= \frac{d}{dx} f(x, y(x)) \\ &= \frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x)) y'(x) \end{aligned} \tag{3.131}$$

et donc

$$y'(x) = -\frac{\frac{\partial f}{\partial x}(x, y(x))}{\frac{\partial f}{\partial y}(x, y(x))} \quad \text{si} \quad \frac{\partial f}{\partial y}(x, y(x)) \neq 0 \tag{3.132}$$