

4. Optimisation sous contraintes

4.1 Position du problème

Soit à résoudre le problème:

$$(Pb) \begin{cases} \min f(x) & x \in \mathbb{R}^n \\ h_i(x) = 0 & i = 1, \dots, p \\ g_i(x) \leq 0 & i = p+1, \dots, m \end{cases}$$

où f , h et g sont des fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}$ continûment différentiables. Les fonctions $h_i(x)$ sont des contraintes d'égalité et $g_i(x)$ sont les contraintes d'inégalité. On note χ l'ensemble des solutions du (Pb) , i.e:

$$\chi = \{x \in \mathbb{R}^n / h_i(x) = 0 \text{ et } g_i(x) \leq 0\}$$

■ **Exemple 4.1** Soit les contraintes suivantes:

$$\begin{cases} g_1(x) \equiv -x_1 \leq 0 \\ g_2(x) \equiv -x_2 \leq 0 \\ g_3(x) \equiv -(1-x_1) + x_2 \leq 0 \end{cases}$$

■ **Exemple 4.2** Soit les contraintes suivantes:

$$\begin{cases} g_1(x) \equiv x_2^2 + (x_1 + 0.5)^2 - 1 \leq 0 \\ g_2(x) \equiv x_2^2 + (x_1 - 0.5)^2 - 1 \leq 0 \end{cases}$$

■ **Exemple 4.3** Dans cet exemple, on souhaite illustrer l'effet des contraintes sur la solution optimale. On définit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ par:

$$f(x) = (x_1 - 1)^2 + 2(x_2 - 2)^2 \text{ et } g(x) = \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ -x_1 - x_2 \end{bmatrix}$$

La zone admissible du problème est représentée par la figure 4.3. Un aspect important de cet exemple est que la deuxième contrainte $g_2(x) = -x_1 - x_2 \leq 0$ n'a pas d'effet sur la solution. Bien que si cette contrainte est changée en $g_2(x) = -x_1 - x_2 \leq a$ pour certaine valeur positive, elle peut affecter la solution. ■

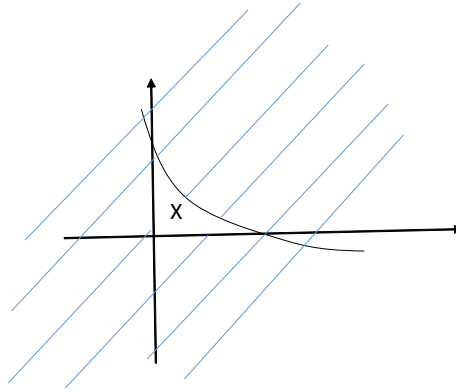


Figure 4.1: Zone admissible pour l'exemple 4.1

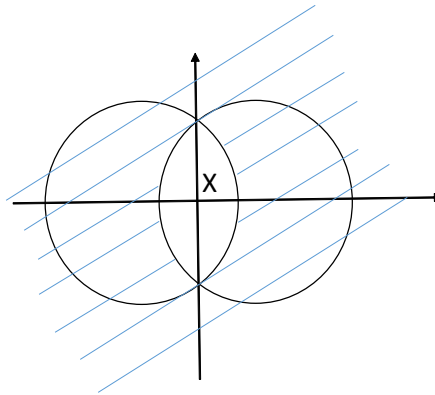


Figure 4.2: Zone admissible pour l'exemple 4.2

Définition 4.1.1 — Contrainte active et contrainte inactive. On dit qu'une contrainte d'inégalité $g_i(x) \leq 0$ est **active** en un point $x \in \chi$ si $g_i(x) = 0$. Elle est dite **inactive** si $g_i(x) < 0$. Par convention, toutes les contraintes h_i sont actives.

Définition 4.1.2 — Point régulier. Un point x^* qui satisfait $h_i(x) = 0$ est appelé un point régulier si les vecteurs gradients $\nabla h_1(x^*), \nabla h_2(x^*), \dots, \nabla h_p(x^*)$ sont linéairement indépendants, i.e, que la matrice définie par:

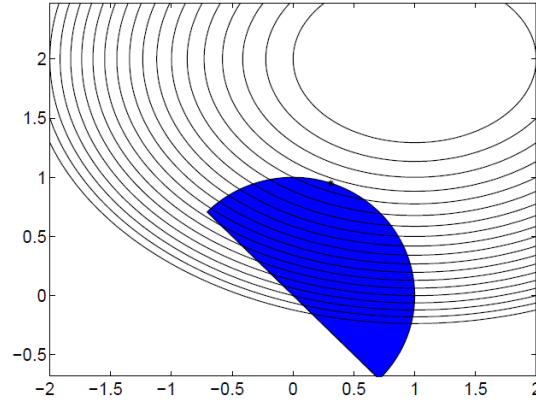
$$J(x^*) = \begin{bmatrix} \nabla h_1(x^*)^T \\ \nabla h_2(x^*)^T \\ \vdots \\ \nabla h_p(x^*)^T \end{bmatrix}$$

est de rang complet.

Dans le cas contraire, on dit que x^* est un point singulier ou encore point stationnaire.

■ **Exemple 4.4 — Question de cours de l'examen Janvier 2017 .** Soient les deux fonctions suivantes: $h_1(x) = -3x_1^2 - 4x_2 + x_3^2$ et $h_2(x) = 3x_1 + 2x_2 - x_3$ et soit $x^* = (1, 2, 1)^T$. Est-ce que x^* est un point régulier pour h_1 et h_2 . ■

En un point régulier, il est possible de caractériser le plan tangent en terme des gradients des contraintes.

Figure 4.3: Les contours de f et la zone admissible déterminée par $g(x) \leq 0$

Théorème 4.1.1 En un point régulier x^* de la surface S définie par $h(x^*) = 0$: $S = \{x \in \mathbb{R}^n \mid h(x) = 0\}$.

Le plan tangent $T = \text{span}(\nabla h_1, \nabla h_2, \dots, \nabla h_p)$ est défini par:

$$T = \{y : \nabla h(x)^T y = 0\}$$

■ **Lemme 4.1.2** Soit x^* un point régulier des contraintes $h(x^*) = 0$ et aussi un point extrême local de $f(x)$ sous ces contraintes, alors, pour tout $y \in T$ qui satisfait :

$$\nabla h(x^*)^T y = 0$$

doit aussi satisfaire:

$$\nabla f(x^*)^T y = 0$$

i.e, $\nabla f(x^*)$ est orthogonal au plan tangent T . Ceci implique que $\nabla f(x^*)$ est une combinaison linéaire des gradients de ∇h_i en x^* , ce qui conduit à l'introduction des multiplicateurs de Lagrange entre le gradient de ∇f et ∇h_i .

4.2 Optimisation avec contraintes d'égalité

Ici, le problème à minimiser est défini par:

$$(Pb) \begin{cases} \min f(x) & x \in \mathbb{R}^n \\ h_i(x) = 0 & i = 1, \dots, p \end{cases}$$

4.2.1 Condition nécessaire du premier ordre

Supposons que f et les fonctions h_i sont continûment différentiables. A chaque contrainte est associé un réel λ_i qui s'appelle le multiplicateur de Lagrange.

Théorème 4.2.1 Pour qu'un point régulier x^* soit un extrême local pour Pb , il est nécessaire que:

$$\exists \lambda^* / \nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) = 0$$

La combinaison linéaire de la fonction d'objective et les contraintes s'appelle le *Lagrangien*:

$$\ell(x, \lambda) = f(x) + \lambda^T h(x)$$

R Si le Lagrangien $\ell(x, \lambda)$ est convexe en x , alors la solution est un minimum global.

■ **Exemple 4.5** Considérons le problème suivant:

$$\begin{cases} \min & f(x) = x_1^2 + x_2^2 \\ & h(x) \triangleq x_1^2 + 2x_2^2 - 1 = 0 \end{cases}$$

Les tracés des contours de f et de la fonction $h(x)$ sont reportés dans Fig. 4.4.

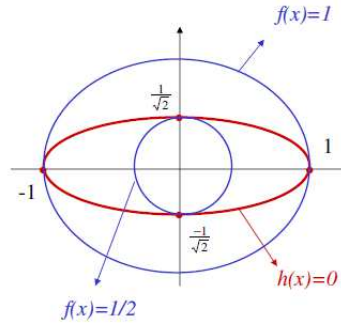


Figure 4.4: Contours de f intersectant avec $h(x) = 0$

Il clair que le problème admet quatre extrêmes; deux maximums correspondant à $(1,0)$ et $(-1,0)$ et deux minimums en $(0, \frac{1}{\sqrt{2}})$ et $(0, -\frac{1}{\sqrt{2}})$.

- Le Lagrangien de ce problème est le suivant:

$$\ell(x, \lambda) = f(x) + \lambda h(x) = x_1^2 + x_2^2 + \lambda(x_1^2 + 2x_2^2 - 1)$$

- La condition nécessaire du premier ordre:

$$\nabla \ell(x^*, \lambda^*) = 0 \Leftrightarrow \nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) = 0$$

Donc, trouver les solutions du problème revient à trouver les solutions x_i^* et le multiplicateur de Lagrange λ^* . Le système d'équation à résoudre sera alors composé par deux équations résultantes du gradient de Lagrangien en plus de l'équation $h(x^*) = 0$:

$$\begin{cases} x_1^* + \lambda^* x_1^* = 0 & \Leftrightarrow x_1^* = 0 \text{ ou } \lambda^* = -1 \\ x_2^* + 2\lambda^* x_2^* = 0 & \Leftrightarrow x_2^* = 0 \text{ ou } \lambda^* = -\frac{1}{2} \\ x_1^{*2} + 2x_2^{*2} - 1 = 0 \end{cases}$$

Alors, les solutions trouvées sont les suivantes:

$$x_1^* = 0 \text{ et } x_2^* = \pm \frac{1}{\sqrt{2}} \text{ et } \lambda^* = -\frac{1}{2}.$$

$$x_2^* = 0 \text{ et } x_1^* = \pm 1 \text{ et } \lambda^* = -1.$$

■

4.2.2 Conditions du second ordre

Théorème 4.2.2 — Conditions nécessaires de second ordre. Supposons que x^* est un minimum local de f sous les contraintes $h_i(x) = 0$ et que x^* est un point régulier de ces contraintes. Alors, il existe un $\lambda_i \in \mathbb{R}$ tel que:

1. $\nabla \ell(x^*, \lambda^*) = 0 \Leftrightarrow \nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) = 0$.

2. $y^T L(x^*, \lambda^*) y \geq 0$, $\forall y \in T$ avec $y \neq 0 | y^T \nabla h_i = 0$, où $L(x^*, \lambda^*) = F(x^*) + \sum \lambda_i^* H_i(x^*)$ est la matrice Hessienne de $\ell(x^*, \lambda^*)$, avec $F(x^*)$ et $H(x^*)$ sont les matrices Hessiennes de $f(x)$ et de $h(x)$, respectivement.

■ **Lemme 4.2.3 — Conditions suffisantes .**

Si $\exists x^*$ et $\exists \lambda^*$ tel que :

1. $\nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) = 0$ et $h(x^*) = 0$.
2. $y^T L(x^*, \lambda^*) y > 0$ pour $\forall y \neq 0 | y^T \nabla h_i(x^*) = 0 \Rightarrow x^*$ est un minimum local au sens strict de $f(x)$ avec $h(x) = 0$.
3. Si $y^T L(x^*, \lambda^*) y < 0$ pour les mêmes conditions, alors x^* est un maximum local au sens strict de $f(x)$ sous $h(x) = 0$.

■ **Exemple 4.6 — Suite de l'exemple précédent.** Les solutions trouvées dans l'exemple précédent

étaient: $x^* = \begin{bmatrix} 0 \\ \pm \frac{1}{\sqrt{2}} \end{bmatrix}$ et $\lambda^* = -\frac{1}{2}$ ou bien $x^* = \begin{bmatrix} \pm 1 \\ 0 \end{bmatrix}$ et $\lambda^* = -1$

La matrice Hessienne:

$$L(x^*, \lambda^*) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ pour } \lambda^* = -\frac{1}{2} \text{ ou bien } L(x^*, \lambda^*) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} \text{ pour } \lambda^* = -1$$

Maintenant, il faut choisir des y appartenant au plan tangent de $h(x)$ au point x^* :

$$\nabla h^T(x^*) = \begin{bmatrix} 0 \\ \pm 2\sqrt{2} \end{bmatrix} \text{ si } \lambda^* = -\frac{1}{2} \Rightarrow y^T \nabla h(x^*) = 0 \Leftrightarrow y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \text{ avec } y_1 \neq 0.$$

$$\nabla h^T(x^*) = \begin{bmatrix} 0 \\ y_2 \end{bmatrix} \text{ si } \lambda^* = -1 \Rightarrow y^T \nabla h(x^*) = 0 \Leftrightarrow y = \begin{bmatrix} 0 \\ y_2 \end{bmatrix} \text{ avec } y_2 \neq 0$$

$$\Rightarrow y^T L(x^*, \lambda^*) y = y_1^2 > 0 \text{ si } \lambda^* = -\frac{1}{2} \text{ ou bien } = -2y_2^2 < 0 \text{ si } \lambda^* = -1$$

$$\Rightarrow \text{deux minimums en } \begin{bmatrix} 0 \\ \pm \frac{1}{\sqrt{2}} \end{bmatrix} \text{ et deux maximums en } \begin{bmatrix} \pm 1 \\ 0 \end{bmatrix}$$

■

Exercice 4.1 — Examen semestriel Janvier 2017. Soient les deux fonctions $f(x, y) = xy$ et $g(x, y) = \frac{1}{x} + \frac{1}{y}$

1. Trouver les extrêma de la fonction f sous la contrainte $g(x, y) = \frac{2}{3}$.
2. Trouver les extrêma la fonction g sous la contrainte $f(x, y) = 9$.
3. Pour chaque cas, examiner la nature des extrêma trouvés.

■

4.3 Optimisation avec contraintes d'inégalité

Définition 4.3.1 Un point x^* satisfaisant les contraintes $h_i(x^*) = 0$ et $g_j(x^*) = 0 \forall j \in I(x^*)$ l'ensemble des indices des contraintes actives est appelé point régulier si les vecteurs gradients $\nabla h_i(x^*)$ et $\nabla g_j(x^*)$ sont linéairement indépendants.

4.3.1 Conditions nécessaires de Karush-Kuhn-Tucker (KKT) du premier ordre

Soit x^* un point qui minimise $f(x)$ sous $h_i(x) = 0$ et $g_i(x) \leq 0$. Si x^* est point régulier, alors $\exists \lambda_i, \mu_i \in \mathbb{R}$ tel que:

1. $\nabla f(x^*) + \sum \lambda_i \nabla h_i(x^*) + \sum \mu_i \nabla g_i(x^*) = 0$
2. $h(x^*) = 0$
3. $\mu_i g_i(x) = 0$ alors si $g_i(x)$ est active $\Rightarrow \mu_i \neq 0$ et si $g_i(x)$ est inactive $\Rightarrow \mu_i = 0$
4. $g_i(x) \leq 0$
5. $\mu_i \geq 0$. Si $g_i(x) \geq 0$ alors $\mu_i \leq 0$.

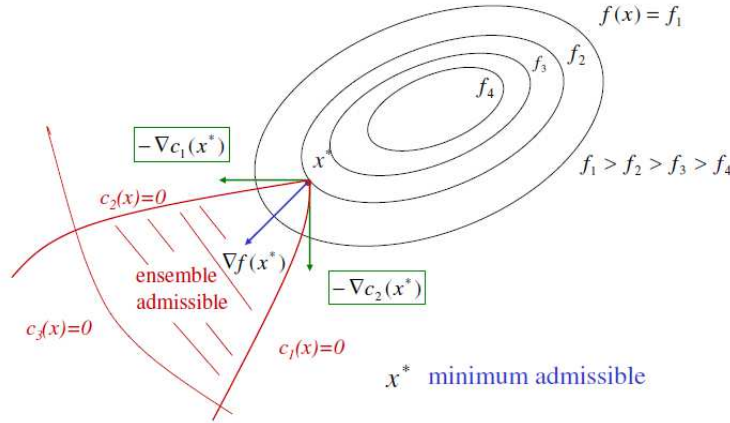


Figure 4.5: Les différentes directions des gradients

■ **Exemple 4.7** Soit à minimiser le problème décrit par le système suivant:

$$\begin{cases} \min & f(x) = x_1^2 - x_2 \\ & h(x) \triangleq x_1 + x_2 = 2 \\ & g_1(x) \triangleq x_1^2 + x_2^2 \leq 4 \\ & g_2(x) \triangleq 1 - x_1 \leq 0 \end{cases}$$

Le Lagrangien associé à ce problème s'écrit par:

$$\begin{aligned} \ell(x, \lambda, \mu_i) &= f(x) + \lambda h(x) + \mu_1 g_1(x) + \mu_2 g_2(x) \\ \Leftrightarrow \ell(x, \lambda, \mu_i) &= x_1^2 - x_2 + \lambda(x_1 + x_2 - 2) + \mu_1(x_1^2 + x_2^2 - 4) + \mu_2(1 - x_1) \end{aligned}$$

La condition nécessaire de KKT du premier ordre:

$$\begin{cases} \nabla \ell(x^*, \lambda^*, \mu_i^*) = 0 \\ h(x^*) = 0 \\ \mu_i g_i(x) = 0 \\ g_i(x) \leq 0 \\ \mu_i \geq 0 \end{cases} \Leftrightarrow \begin{cases} 2x_1 + \lambda + 2\mu_1 x_1 - \mu_2 = 0 \\ -1 + \lambda + 2\mu_1 x_2 = 0 \\ x_1 + x_2 = 2 \\ \mu_1(x_1^2 + x_2^2 - 4) = 0 \\ \mu_2(1 - x_1) = 0 \\ x_1^2 + x_2^2 - 4 \leq 0 \\ 1 - x_1 \leq 0 \\ \mu_1 \geq 0, \mu_2 \geq 0 \end{cases}$$

On doit examiner tous les cas possibles pour trouver une éventuelle solution. Il est clair que ce problème admet quatre cas possibles:

Premier cas: $\mu_1 = \mu_2 = 0$

$$\begin{cases} 2x_1 + \lambda = 0 \Rightarrow x_1 = -\frac{1}{2} \\ -1 + \lambda = 0 \Rightarrow \lambda = 1 \\ x_1 + x_2 = 2 \Rightarrow x_2 = 2 + \frac{1}{2} = \frac{5}{2} \\ x_1^2 + x_2^2 - 4 < 0 \\ 1 - x_1 < 0 \Rightarrow 1 - (-\frac{1}{2}) < 0 \text{ Impossible} \end{cases}$$

Donc, cette solution est écartée.

Deuxième cas: $\mu_1 \neq 0$ et $\mu_2 = 0$

$$\begin{cases} 2x_1 + \lambda + 2\mu_1 x_1 = 0 \\ -1 + \lambda + 2\mu_1 x_2 = 0 \\ x_1 + x_2 = 2 \\ x_1^2 + x_2^2 - 4 = 0 \\ 1 - x_1 < 0 \end{cases}$$

Dans ce cas là, le problème admet deux solutions: la première est $\mu_1 = \frac{1}{4}$, $\lambda = 0$, $x_1 =$

0, $x_2 = 2$. Cette solution est écartée car elle ne vérifie pas la contrainte $g_2(x) < 0$. La deuxième solution est $\mu_1 = -\frac{5}{4}$ qui est négative, donc, c'est une solution à écarter aussi.

Troisième cas: $\mu_1 = 0$ et $\mu_2 \neq 0$

$$\begin{cases} 2x_1 + \lambda - \mu_2 = 0 \Rightarrow \mu_2 = 3 \\ -1 + \lambda = 0 \Rightarrow \lambda = 1 \\ x_1 + x_2 = 2 \Rightarrow x_2 = 1 \\ x_1^2 + x_2^2 - 4 < 0 \\ 1 - x_1 = 0 \Rightarrow x_1 = 1 \end{cases}$$

La solution: $\mu_2 = 3$, $\lambda = 1$, $x_1 = 1$, $x_2 = 1$ est une solution acceptable qui vérifie toutes les conditions de KKT.

Troisième cas: $\mu_1 \neq 0$ et $\mu_2 \neq 0$

$$\begin{cases} 2x_1 + \lambda + 2\mu_1 x_1 - \mu_2 = 0 \\ -1 + \lambda + 2\mu_1 x_2 = 0 \\ x_1 + x_2 = 2 \\ x_1^2 + x_2^2 - 4 = 0 \\ 1 - x_1 = 0 \end{cases}$$

La résolution des trois dernières équations ne permet pas d'avoir une solution. Par conséquent, le problème admet une seule solution qui correspond au troisième cas: $x_1 = 1$, $x_2 = 1$, $\mu_1 = 0$, $\mu_2 = 3$, $\lambda = 1$.

Maintenant, il faut vérifier si cette solution est vraiment un point régulier pour qu'on puisse dire que c'est un optimum. Ici, on a deux contraintes actives; $h(x)$ et $g_2(x)$ puisque

$$\mu_2 \neq 0. \text{ Alors, } \nabla h(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ et } \nabla g_2(x^*) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Il est clair que si on forme une matrice dont les colonnes sont les gradients des contraintes actives;

$$J = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$$

On obtient une matrice qui a un déterminant différent de zéro, ce qui conduit à dire que les deux gradients sont indépendants donc l'optimum est un point régulier.

Graphiquement, on peut voir clairement que cet optimum est bien un minimum (voir Fig. 4.6).

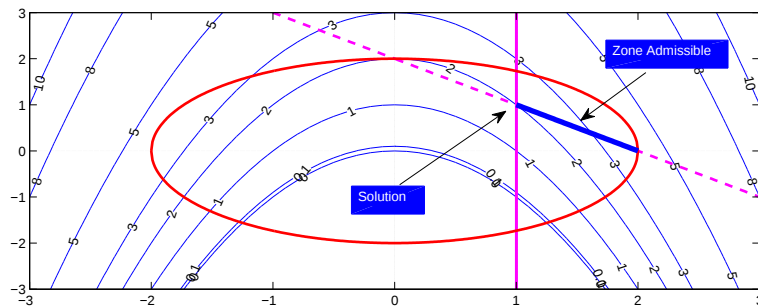


Figure 4.6: Le tracé des contours de f et les différentes contraintes

4.3.2 Conditions nécessaires et suffisantes de second ordre de KKT

Conditions nécessaires de second ordre

Supposons que les fonctions f , h et $g \in C^2$ et que x^* est un point régulier pour les contraintes actives. Si x^* est un point minimum relative, alors:

$\exists \lambda, \mu$, tel que:

1. $\nabla f(x^*) + \sum \lambda_i \nabla h_i(x^*) + \sum \mu_i \nabla g_i(x^*) = 0$
2. $h_i(x^*) = 0$
3. $\mu_i g_i(x) = 0$
4. $g_i(x) \leq 0$
5. $y^T L(x^*, \lambda^*, \mu^*) y \geq 0, \forall y \neq 0 \begin{cases} y^T \nabla h(x^*) = 0 \\ y^T \nabla g_j(x^*) = 0 \forall j \in I(x^*) \end{cases}$
6. $\mu_i \geq 0$.

où $L(x^*, \lambda^*, \mu^*) = F(x^*) + \sum \lambda_i^* H_i(x^*) + \sum \mu_i^* G_i(x^*)$ avec L, F, H , et G sont les matrices Hessiennes de la Lagrangienne, fonction de coût, les contraintes d'égalité et d'inégalité, respectivement.

Conditions suffisantes de second ordre

Soit, $f, h, g \in C^2$. Les conditions suffisantes pour qu'un point x^* , qui satisfait les contraintes, soit un minimum sont:

1. $\nabla f(x^*) + \sum \lambda_i \nabla h_i(x^*) + \sum \mu_i \nabla g_i(x^*) = 0$
2. $h_i(x^*) = 0$
3. $\mu_i g_i(x) = 0$
4. $g_i(x) \leq 0$
5. $y^T L(x^*, \lambda^*, \mu^*) y > 0, \forall y \neq 0 \begin{cases} y^T \nabla h(x^*) = 0 \\ y^T \nabla g_j(x^*) = 0 \forall j \in I(x^*) \end{cases}$
6. $\mu_i \geq 0$.

Alors, x^* est un minimum local au sens strict de $f(x)$ avec $h(x) = 0$ et $g(x) \leq 0$.

De plus, si $y^T L(x^*, \lambda^*, \mu^*) y < 0, \forall y \neq 0 \begin{cases} y^T \nabla h(x^*) = 0 \\ y^T \nabla g_j(x^*) = 0 \forall j \in I(x^*) \end{cases}$,

alors x^* est un maximum local au sens strict de $f(x)$ avec $h(x) = 0$ et $g(x) \leq 0$.

R Si les contraintes d'inégalité sont données sous forme $g_i(x) \geq 0$, alors les correspondants multiplicateurs de Lagrange sont définis comme suit: $\mu_i^* \leq 0$.

■ **Exemple 4.8** Reprenons l'exemple précédent en traitant les conditions de second ordre pour déterminer la nature du l'optimum par rapport aux variables x . Dans le cas où l'optimum est un minimum, les conditions nécessaires sont données par:

$$y^T L(x^*, \lambda^*, \mu^*) y > 0 \quad \forall y \neq 0 \mid \begin{bmatrix} \nabla h(x^*) \\ \nabla g_{i \mid \forall j \in I(x^*)}(x^*) \end{bmatrix} y = 0$$

de telle façon que le Hessien de $L(x^*, \lambda^*, \mu^*)$ par rapport à x soit DP pour n'importe quel vecteur y orthogonal aux gradients des contraintes actives en x^* . Pour une condition nécessaire, il suffit avec $L(x^*, \lambda^*, \mu^*)$ SDP.

$$L(x, \lambda, \mu) = x_1^2 - x_2 + \lambda(x_1 + x_2 - 2) + \mu_1(x_1^2 + x_2^2 - 4) + \mu_2(1 - x_1)$$

Les vecteurs appartenant au plan tangent au point x^* formé par les contraintes actives:

$$y^T \nabla h(x^*) = 0 \Rightarrow [y_1 \ y_2] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0 \Rightarrow y^T = [y_1 \ -y_1] .$$

$$y^T \nabla g_2(x^*) = 0 \Rightarrow [y_1 \ y_2] \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 0 \Rightarrow y^T = [0 \ y_2] .$$

$$L(x^*, \lambda^*, \mu^*) = \begin{bmatrix} \frac{\partial L^2}{\partial x_1^2} & \frac{\partial L^2}{\partial x_1 \partial x_2} \\ \frac{\partial L^2}{\partial x_1 \partial x_2} & \frac{\partial L^2}{\partial x_2^2} \end{bmatrix}$$

Pour ∇h :

$$y^T L(x^*, \lambda^*, \mu^*) y = [y_1 \ -y_1]^T \begin{bmatrix} 2+2\mu_1^* & 0 \\ 0 & 2\mu_1^* \end{bmatrix} = 2y_1^2 > 0$$

Pour ∇g_2 :

$$y^T L(x^*, \lambda^*, \mu^*) y = [0 \ y_1]^T \begin{bmatrix} 2+2\mu_1^* & 0 \\ 0 & 2\mu_1^* \end{bmatrix} \begin{bmatrix} 0 \\ y_1 \end{bmatrix} = 0$$

Selon le résultat obtenu, il est clair que la condition nécessaire est vérifiée mais non la suffisante.

■

■ **Exemple 4.9** Trouver la solution optimale du problème:

$$\begin{cases} \min_{x_1, x_2} & f(x) = (x_1^2 + 3)^2 + x_2^2 \\ & g_1(x) \triangleq -x_1 + x_2 - 2 \leq 0 \\ & g_2(x) \triangleq x_1^2 - x_2 - 4 \leq 0 \end{cases}$$

les contours de f et la zone admissible sont reportés par la Fig 4.7.

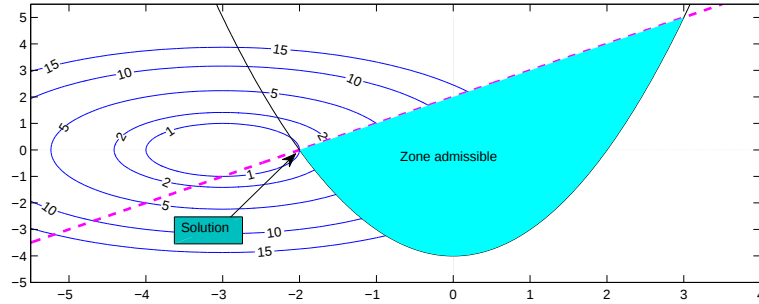


Figure 4.7: Le tracé des contours de f et la zone admissible

Conditions nécessaires du premier ordre

Le Lagrangien de ce problème est défini comme suit:

$$\ell(x, \mu) = (x_1 + 3)^2 + x_2^2 + \mu_1(-x_1 + x_2 - 2) + \mu_2(x_1^2 - x_2 - 4)$$

Les conditions nécessaires de KKT du premier ordre sont:

$$\begin{cases} \nabla \ell(x^*, \mu_i^*) = 0 \\ \mu_1 g_1(x) = 0 \\ \mu_2 g_2(x) = 0 \\ g_1(x) \leq 0 \\ g_2(x) \leq 0 \\ \mu_1 \geq 0, \mu_2 \geq 0 \end{cases} \Leftrightarrow \begin{cases} 2(x_1 + 3) + \mu_1(-1) + \mu_2 2x_1 = 0 \\ 2x_2 + \mu_1(1) + \mu_2(-1) = 0 \\ \mu_1(-x_1 + x_2 - 2) = 0 \\ \mu_2(x_1^2 - x_2 - 4) = 0 \\ -x_1 + x_2 - 2 \leq 0 \\ x_1^2 - x_2 - 4 \leq 0 \end{cases} \dots (1)$$

Premier cas $\mu_1 = 0, \mu_2 = 0$

$$(1) \Leftrightarrow \begin{cases} 2(x_1 + 3) = 0 \Rightarrow x_1 = -3 \\ 2x_2 = 0 \Rightarrow x_2 = 0 \\ -x_1 + x_2 - 2 < 0 \Rightarrow -(-3) + (0) - 2 = 1 > 0 \text{ Impossible, donc solution écartée} \\ x_1^2 - x_2 - 4 < 0 \end{cases}$$

Deuxième cas $\mu_1 \neq 0, \mu_2 = 0$

$$(1) \Leftrightarrow \begin{cases} 2(x_1 + 3) - \mu_1 = 0 \\ 2x_2 + \mu_1 = 0 \\ -x_1 + x_2 - 2 = 0 \\ x_1^2 - x_2 - 4 < 0 \end{cases} \Rightarrow \mu_1 = 1, x_1 = -5/2, x_2 = -1/2$$

Si on remplace ces valeurs dans la dernière contrainte inactive ($g_2(x) < 0$), on trouve que c'est une solution à écarter.

Troisième cas $\mu_1 = 0, \mu_2 \neq 0$

$$(1) \Leftrightarrow \begin{cases} 2(x_1 + 3) + 2\mu_2 x_1 = 0 \\ 2x_2 - \mu_2 = 0 \\ -x_1 + x_2 - 2 < 0 \\ x_1^2 - x_2 - 4 = 0 \end{cases}$$

La résolution de ce système d'équation peut être faite de plusieurs manières. Ici, on tire x_1 et x_2 des deux premières équations en fonction de μ_2 et on les remplace dans la contrainte active ($g_2(x) = 0$), on trouve:

$$\mu_2^3 + 10\mu_2^2 + 17\mu_2 = 10$$

Les solutions sont: $\mu_2 = -7.58, -2.87, 0.46$. La solution qui semble être acceptable c'est la troisième puisque les deux premières sont négatives. Alors pour $\mu_2 = 0.46 \Rightarrow x_1 = -2.06, x_2 = 0.23$. Ces deux solutions ne vérifient pas la contrainte inactive ($g_1(x) < 0$), ce qui écarte cette solution.

Quatrième cas $\mu_1 \neq 0, \mu_2 \neq 0$

$$(1) \Rightarrow \begin{cases} 2(x_1 + 3) - \mu_1 + 2\mu_2 x_1 = 0 \\ 2x_2 + \mu_1 - \mu_2 = 0 \\ -x_1 + x_2 - 2 = 0 \\ x_1^2 - x_2 - 4 = 0 \end{cases}$$

La résolution des deux contraintes actives donne: $x_1^* = (3, 5)$ et $x_2^* = (-2, 0)$. La première solution mène à $\mu_2 = -22/9 < 0$, donc, solution écartée. Pour $x_2^* = (-2, 0)$ on trouve que $\mu_1 = \mu_2 = 2/5$, alors, c'est une solution acceptable pour ce problème.

Maintenant, il reste à vérifier si cette solution est un point régulier pour qu'il soit un point optimal. Suivant les solutions obtenues, le problème admet deux contraintes actives. Les vecteurs gradients de ces contraintes sont:

$$\nabla g_1(x_2^*) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ et } \nabla g_2(x_2^*) = \begin{bmatrix} 2(-2) \\ -1 \end{bmatrix}, \text{ alors, il est clair que: } \text{rang} \begin{bmatrix} -1 & -4 \\ 1 & -1 \end{bmatrix} = 2$$

Donc, ils sont linéairement indépendants, ce qui vérifie la condition pour un point régulier.

Conditions suffisantes de second ordre

Pour étudier la nature du point optimal x^* , il faut analyser la nature du Hessien par rapport au plan tangent engendré par les gradients des contraintes actives (conditions de KKT).

$$-y^T \nabla g_1(x^*) = 0 \Leftrightarrow [y_1 \ y_2] \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0 \Rightarrow y = [y_1 \ y_1].$$

$$-y^T \nabla g_2(x^*) = 0 \Leftrightarrow [y_1 \ y_2] \begin{bmatrix} -4 \\ -1 \end{bmatrix} = 0 \Rightarrow y = [y_1 \ -4y_1]. \text{ Le Hessien:}$$

$$L(x^*, \mu^*) = \begin{bmatrix} \frac{\partial^2 \ell}{\partial x_1^2} & \frac{\partial^2 \ell}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \ell}{\partial x_2 \partial x_1} & \frac{\partial^2 \ell}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 + 2\mu_2^* & 0 \\ 0 & 2 \end{bmatrix}$$

Pour le vecteur $\nabla g_1(x^*)$

$$y^T L(x^*, \mu_i^*) y = [y_1 \ y_1] \begin{bmatrix} 14/5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1 \end{bmatrix} = 14/5 y_1^2 > 0$$

Pour le vecteur $\nabla g_1(x^*)$

$$y^T L(x^*, \mu_i^*) y = [y_1 \quad -4y_1] \begin{bmatrix} 14/5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ -4y_1 \end{bmatrix} = 14/5 y_1^2 > 0$$

Donc, la condition nécessaire et suffisante de second ordre est vérifiée pour que x^* soit un minimum strict de f sous $g_1(x)$ et $g_2(x)$. ■

4.4 Méthodes de programmation non linéaire avec contraintes

Les problèmes d'optimisation avec contraintes sont plus difficiles à résoudre dues aux restrictions additionnées que la solution optimale doit satisfaire.

La majorité des méthodes proposées pour surmonter ce type de problème sont basées sur l'un des concepts suivants :

- Extension de la méthodologie linéaire à travers des approximations linéaires répétées.
- Transformation du problème avec contraintes en problème sans contraintes par le biais des fonctions de pénalisation.
- Usage des tolérances flexibles pour accommoder les vecteurs $\vec{x}$ en régions possibles et non possibles.

Dans la présente section, nous étudierons les deux derniers concepts.

4.4.1 Méthode de fonction de pénalité

Le principe de cette méthode est de transformer un problème d'optimisation avec contraintes en problème sans contraintes ou bien en une séquence de problèmes sans contraintes.

Par exemple, on suppose qu'il est demandé de trouver le minimum du problème suivant:

$$\begin{cases} \min & f(x) = (x_1 - 3)^2 + (x_2 + 2)^2 \\ & h(x) \triangleq x_1 + x_2 - 4 = 0 \end{cases}$$

Il est possible de déduire une nouvelle fonction d'objective sans contraintes;

$$P(x) = (x_1 - 3)^2 + (x_2 + 2)^2 + (x_1 + x_2 - 4)^2$$

en rajoutant le carré de la contrainte à la fonction f comme une pénalité. Durant la minimisation de $P(x)$, le vecteur $\vec{x}$ est forcé par la pénalité à satisfaire la contrainte.

Les méthodes des pénalités peuvent être réparties en deux classes:

1. **Méthodes paramétriques:** Sont caractérisées par un ou plusieurs paramètres ajustables qui pondèrent la fonction de pénalité formée par des contraintes. Ces méthodes peuvent être réparties en:
 - *Méthodes des pénalités intérieures:* Maintenir loin les bordures de la région admissible à travers l'addition de la fonction de pénalité.
 - *Méthodes des pénalités extérieures:* Générer une séquence de points non admissibles alors que la limite atteint une solution admissible.
 - *Méthodes mixtes:* sont largement utilisées dans le cas des contraintes d'égalité. Quelques contraintes sont satisfaites et d'autres non, mais dans l'ensemble, toutes les contraintes sont satisfaites dans une tolérance quand la solution est atteinte.
2. **Méthodes non-paramétriques** Ils traitent la fonction d'objective comme une contrainte supplémentaire qui est fixée successivement à partir des informations développées dans la solution du problème.

Formellement, la transformation du problème est la suivante:

$$\begin{aligned}
 & \text{Minimiser } f(\vec{x}) \quad \vec{x} \in E^n \\
 & \text{sous } m \text{ contraintes d'égalité} \\
 & \quad h_j(\vec{x}) = 0 \quad j = 1, \dots, m \\
 & \text{et } m' \text{ contraintes d'inégalité} \\
 & \quad g_j(\vec{x}) \leq 0 \quad j = 1, \dots, m' \\
 & \quad \Downarrow \\
 & \text{Minimiser } P(x^{(k)}, \rho^{(k)}) = f(x^{(k)}) + \sum_{i=1}^m \rho_i^{(k)} H(h_i(x^{(k)})) + \sum_{i=1}^{m'} \rho_i^{(k)} G(g_i(x^{(k)}))
 \end{aligned}$$

avec

$P(x^{(k)}, \rho^{(k)})$: Fonction augmentée généralisée ou la fonction de pénalisation.

$\rho_i^{(k)}$: Facteurs de pondération.

$H(h_i(x^{(k)}))$: Fonctionnel de $h_i(x^{(k)})$.

$G(g_i(x^{(k)}))$: Fonctionnel de $g_i(x^{(k)})$.

k : est le pas de la recherche numérique.

Les fonctionnels des contraintes d'inégalité peuvent être choisies comme suit:

1. $G_1(g_i(x^{(k)})) \rightarrow \infty$ si $g_i(x) \rightarrow 0^+ \Rightarrow x$ est un point d'intérieur.
2. $G_2(g_i(x^{(k)})) \rightarrow \infty$ si $g_i(x) \rightarrow 0^- \Rightarrow x$ est un point d'extérieur.

Pour une contrainte d'égalité, la sélection fonctionnelle typique est:

$H(h_i(x) \rightarrow 0$ si $h_i(x) \rightarrow 0$.

Une forme typique pour les contraintes d'égalité est $H(h_i(x) = h_i^2(x)$.

De plus, pour toutes les sélections précédentes, il est exigé que:

$$\begin{aligned}
 \lim_{k \rightarrow \infty} \sum_{i=1}^{m'} \rho_i^{(k)} G(g_i(x^{(k)})) &= 0 \\
 \lim_{k \rightarrow \infty} \sum_{i=1}^m \rho_i^{(k)} H(h_i(x^{(k)})) &= 0 \\
 \lim_{k \rightarrow \infty} |P(x^{(k)}, \rho^{(k)}) - f(x^{(k)})| &= 0
 \end{aligned}$$

Ceci veut dire que, l'effet des contraintes dans la fonction augmentée est graduellement diminué dans la mesure que la recherche progresse et complètement éliminé dans les limites, de telle façon que la valeur de la fonction augmentée converge vers la même valeur de $f(x)$ et que l'extrême de $P(x)$ est celui de $f(x)$.

■ **Exemple 4.10** Soit $g(x) \leq 0$ un ensemble admissible qui représente en même temps une contrainte d'inégalité. Les pénalités qui vérifient des points intérieurs peuvent être choisies comme:

1. $G(x) = \frac{-1}{g(x)}$
2. $G(x) = \frac{1}{g(x)^2}$
3. $G(x) = \ln(1 - \frac{1}{g(x)})$

avec x appartenant à cette ensemble admissible.

A titre d'exemple, soit à minimiser $f(x) = 2x$ sous la contrainte $g(x) = 1 - x \leq 0$.

$G(x) = \frac{-1}{g(x)} \Rightarrow P(x, \rho_k) = 2x + \rho_k \frac{1}{x-1}$

Quand à la pénalité extérieure, les fonctions additives peuvent être choisies comme:

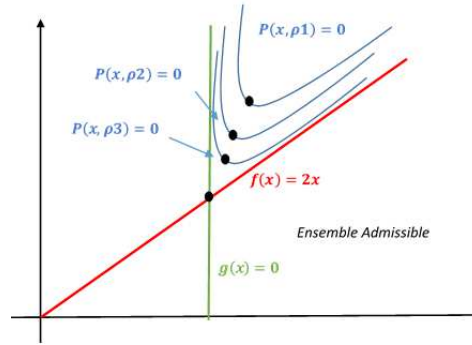


Figure 4.8: Effet de ρ_k sur la solution optimale dans la pénalité intérieure ($\rho_1 > \rho_2 > \rho_3$)

1. $G(g(x)) = \begin{cases} g(x)^q & \text{si } g(x) > 0 \\ 0 & \text{si } g(x) \leq 0 \end{cases} \Rightarrow G(g(x)) = \max(g(x), 0)^q$
2. $G(g(x)) = \begin{cases} a^{g(x)} - 1 & \text{si } g(x) > 0 \\ 0 & \text{si } g(x) \leq 0 \end{cases}$

Un autre exemple traitant le cas de pénalité extérieure. Soit à minimiser le même problème précédent avec la même contrainte.

Pour le cas de pénalité extérieure, $G(g(x)) = \max(g(x), 0)^2$ et le problème augmenté sera donné par:

$$P(x, \rho_k) = 2x + \rho_k \max(1 - x, 0)^2$$

■

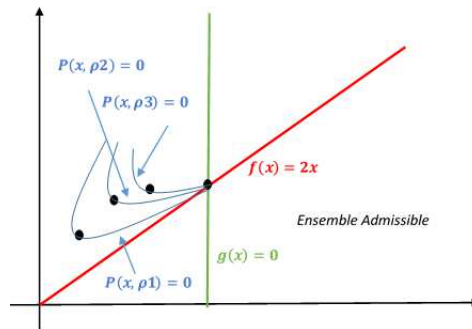


Figure 4.9: Effet de ρ_k sur la solution optimale dans la pénalité extérieure ($\rho_3 > \rho_2 > \rho_1$)

