

# PROBLÈMES D'OPTIMISATION AVEC CONTRAINTES D'INÉGALITÉS

**Définition 2.1.** Un problème de minimisation avec contraintes d'inégalité peut être exprimé de la façon suivante :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \leq 0. \end{cases}$$

Où  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  avec  $m \leq n$  et  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ .

Pour le problème général ci-dessus, nous adoptons les définitions suivantes.

**Définition 2.2.** Une contrainte d'inégalité  $g_j(x) \leq 0$  est dite active à  $x^*$  si  $g_j(x^*) = 0$ . Elle est non active à  $x^*$  si  $g_j(x^*) < 0$ .

Par convention, on considère une contrainte d'égalité  $h_i(x) = 0$  pour être toujours actif.

**Définition 2.3.** Soit  $x^*$  satisfait  $h(x^*) = 0$ ,  $g(x^*) \leq 0$  et soit  $J(x^*)$  l'ensemble d'index des contraintes d'inégalité active :

$$J(x^*) := \{j : g_j(x^*) = 0\}.$$

Alors, on dit que  $x^*$  est un point régulier si les vecteurs

$$\nabla h_i(x^*), \nabla g_j(x^*), 1 \leq i \leq m, j \in J(x^*)$$

sont linéairement indépendants.

## 2.1 Condition nécessaire du premier ordre (conditions de Karush-Kuhn-Tucker)

**Théorème 2.1.** (Conditions de KKT). Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $m \leq n$ , et  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  des fonctions de classe  $\mathcal{C}^1$ . Soit  $x^*$  est un point régulier du problème de minimisation suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \leq 0. \end{cases} \quad (2.1)$$

Si  $x^*$  est un point de minimum local de (2.1), alors il existe  $\lambda^* \in \mathbb{R}^m$  et  $\mu^* \in \mathbb{R}^p$  tels que :

1.  $\mu^* \geq 0$ .
2.  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ .
3.  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$ .

*Démonstration.* Soit  $x^*$  est un point de minimum local de  $f$  sur l'ensemble  $\{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$ . Alors,  $x^*$  est un point de minimum local de  $f$  sur l'ensemble suivant :

$$\{x \in \mathbb{R}^n : h(x) = 0, g_j(x) = 0, j \in J(x^*)\}.$$

D'après la condition de Lagrange, il existe  $\lambda^* \in \mathbb{R}^m$  et  $\mu^* \in \mathbb{R}^p$  tel que

$$Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T,$$

où pour tout  $j \notin J(x^*)$ , on a  $\mu_j^* = 0$ .

Pour compléter la preuve, il reste à montrer que pour tout  $j \in J(x^*)$ , on a  $\mu_j^* \geq 0$ . Nous utilisons une preuve par contradiction. Alors, on suppose qu'il existe  $j \in J(x^*)$  tel que  $\mu_j^* < 0$ . Soient  $\hat{S}$  une surface définie par :

$$\hat{S} = \{x \in \mathbb{R}^n : h(x) = 0, g_i(x) = 0, i \in J(x^*), i \neq j\}$$

et  $\hat{T}(x^*)$  un espace tangent défini par :

$$\hat{T}(x^*) = \{y \in \mathbb{R}^n : Dh(x^*)y = 0, Dg_i(x^*)y = 0, i \in J(x^*), i \neq j\}.$$

**Montrons l'affirmation suivante :** "il existe  $y \in \hat{T}(x^*)$  tel que  $Dg_j(x^*)y \neq 0$ ". En raisonnant par absurde. On suppose pour tout  $y \in \hat{T}(x^*)$  tel que

$$Dg_j(x^*)y = \nabla g_j(x^*)^T y = 0.$$

Donc, on a :

$$\nabla g_j(x^*) \in \hat{T}(x^*)^\perp.$$

D'après Lemme 1.1, on obtient :

$$\nabla g_j(x^*) \in \text{span}[\nabla h_k(x^*), k = 1, \dots, m, \nabla g_i(x^*), i \in J(x^*), i \neq j].$$

Mais cela contredit le fait que  $x^*$  est un point régulier, ce qui prouve notre affirmation. Sans perte de généralité, on suppose que nous avons  $y$  tel que  $Dg_j(x^*)y < 0$ .

Considérons la condition de Lagrange, réécrite comme

$$Df(x^*) + (\lambda^*)^T Dh(x^*) + \mu_j^* Dg_j(x^*) + \sum_{i \neq j} \mu_i^* Dg_i(x^*) = 0^T.$$

En multipliant la condition de Lagrange par  $y \in \hat{T}(x^*)$ , on obtient :

$$Df(x^*)y = -\mu_j^* Dg_j(x^*)y.$$

Par conséquent, on a :

$$Df(x^*)y < 0.$$

D'après Théorème 1.1, on peut trouver une courbe différentiable  $\{x(t) : t \in ]a, b[ \}$  sur  $\hat{S}$  tel qu'il existe  $t^* \in ]a, b[$  avec  $x(t^*) = x^*$  et  $\dot{x}(t^*) = y$ . Donc, on a :

$$\frac{d}{dt} f(x(t^*)) = Df(x^*)y < 0.$$

Alors, il existe  $\delta > 0$  tel que pour tout  $t \in ]t^*, t^* + \delta[$ , nous avons :

$$f(x(t)) < f(x(t^*)) = f(x^*).$$

D'autre part,

$$\frac{d}{dt} g_j(x(t^*)) = Dg_j(x^*) y < 0,$$

donc il existe  $\epsilon > 0$  tel que pour tout  $t \in [t^*, t^* + \epsilon]$ , nous avons :

$$g_j(x(t)) \leq g_j(x^*) = 0.$$

C'est à dire pour tout  $t \in ]t^*, t^* + \min\{\delta, \epsilon\}]$ , on a :

$$f(x(t)) < f(x^*).$$

Donc, contradiction avec que  $x^*$  est un point de minimum local. □

**Remarque 2.1.** 1. Les conditions  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$  sont signifiées que

(a) Si  $\mu_i^* = 0$ , alors  $g_i(x^*) < 0$  ou

(b) Si  $\mu_i^* > 0$ , alors  $g_i(x^*) = 0$ .

2. Pour un problème de maximisation avec contraintes d'inégalité suivant :

$$\begin{cases} \max f(x), \\ h(x) = 0, \\ g(x) \leq 0. \end{cases}$$

Les conditions de KKT sont données par :

(a)  $\mu^* \geq 0$ ,

(b)  $-Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ ,

(c)  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$

(d)  $h(x^*) = 0$ ,

(e)  $g(x^*) \leq 0$ .

Où par une forme équivalente suivante :

(a)  $\mu^* \leq 0$ ,

(b)  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ ,

(c)  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$

(d)  $h(x^*) = 0$ ,

(e)  $g(x^*) \leq 0$ .

3. Pour un problème de minimisation avec contraintes d'inégalité suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \geq 0. \end{cases}$$

Nous multiplions la fonction de contrainte d'inégalité par  $-1$  pour obtenir  $-g(x) \leq 0$ . Donc, les conditions de KKT sont données par :

(a)  $\mu^* \leq 0$ ,

(b)  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ ,

(c)  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$

(d)  $h(x^*) = 0,$

(e)  $g(x^*) \geq 0.$

4. Pour un problème de maximisation avec contraintes d'inégalité suivant :

$$\begin{cases} \max f(x), \\ h(x) = 0, \\ g(x) \geq 0. \end{cases}$$

Les conditions de KKT sont donnés par :

(a)  $\mu^* \geq 0,$

(b)  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T,$

(c)  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$

(d)  $h(x^*) = 0,$

(e)  $g(x^*) \geq 0.$

**Exemple 2.1.** On considère le problème suivant :

$$\begin{cases} \min f(x_1, x_2) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2, \\ x_1^2 + x_2^2 \leq 5, \\ 3x_1 + x_2 \leq 6. \end{cases}$$

On pose  $g_1(x_1, x_2) = x_1^2 + x_2^2 - 5$  et  $g_2(x_1, x_2) = 3x_1 + x_2 - 6$ . Donc, les conditions de KKT sont données par :

$$\begin{cases} \mu_1 \geq 0, \mu_2 \geq 0 \\ Df(x_1, x_2) + (\mu_1, \mu_2) Dg(x_1, x_2) = (0, 0), \\ \mu_1 g_1(x_1, x_2) = 0, \\ \mu_2 g_2(x_1, x_2) = 0. \end{cases} \Leftrightarrow \begin{cases} \mu_1 \geq 0, \mu_2 \geq 0 \\ 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 = 0, \\ 2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 = 0, \\ \mu_1 (x_1^2 + x_2^2 - 5) = 0, \\ \mu_2 (3x_1 + x_2 - 6) = 0. \end{cases}$$

1. Si  $\mu_1 = \mu_2 = 0$ , donc  $x_1^2 + x_2^2 < 5$  et  $3x_1 + x_2 < 6$  et

$$\begin{cases} 4x_1 + 2x_2 = 10, \\ 2x_1 + 2x_2 = 10. \end{cases} \Leftrightarrow x_1 = 0 \text{ et } x_2 = 5.$$

D'autre part, on a :

$$x_1^2 + x_2^2 - 5 = 20.$$

D'ou, contradiction.

2. Si  $\mu_1 > 0$  et  $\mu_2 = 0$ , donc  $x_1^2 + x_2^2 = 5$ . Nous avons :

$$\begin{cases} 4x_1 + 2x_2 + 2\mu_1 x_1 = 10, \\ 2x_1 + 2x_2 + 2\mu_1 x_2 = 10, \\ x_1^2 + x_2^2 = 5. \end{cases} \Leftrightarrow x_1 = 1, x_2 = 2, \mu_1 = 1.$$

Nous avons  $3x_1 + x_2 = 5$ , donc le point  $(1, 2)^T$  est satisfait les conditions de KKT.

## 2.2 Conditions d'optimalité du seconde ordre

On définit la matrice suivante :

$$\mathbf{L}(x, \lambda, \mu) = F(x) + [\lambda \mathbf{H}(x)] + [\mu \mathbf{G}(x)].$$

Où  $F(x)$  est la matrice hessienne de  $f$  au point  $x$ ,  $[\lambda \mathbf{H}(x)]$  une matrice définie par :

$$[\lambda \mathbf{H}(x)] = \lambda_1 \mathbf{H}_1(x) + \dots + \lambda_m \mathbf{H}_m(x)$$

et  $[\mu \mathbf{G}(x)]$  une matrice définie par :

$$[\mu \mathbf{G}(x)] = \mu_1 G_1(x) + \dots + \mu_p G_p(x).$$

Où  $G_k(x)$  est la matrice hessienne de  $g_k$  donnée par :

$$G_k(x) = \begin{bmatrix} \frac{\partial^2 g_k}{\partial x_1^2}(x) & \dots & \frac{\partial^2 g_k}{\partial x_n \partial x_1}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 g_k}{\partial x_1 \partial x_n}(x) & \dots & \frac{\partial^2 g_k}{\partial x_n^2}(x) \end{bmatrix}$$

Dans le théorème suivant, nous utilisons l'espace tangent suivant :

$$T(x^*) = \{y \in \mathbb{R}^n : Dh(x^*)y = 0, Dg_j(x^*)y = 0, j \in J(x^*)\}.$$

**Théorème 2.2.** (Conditions nécessaires du seconde ordre). Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $m \leq n$ , et  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  des fonctions de classe  $\mathcal{C}^2$ . Soit  $x^*$  est un point régulier du problème de minimisation suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \leq 0. \end{cases} \quad (2.2)$$

Si  $x^*$  est un point de minimum local du problème (2.2), alors il existe  $\lambda^* \in \mathbb{R}^m$  et  $\mu^* \in \mathbb{R}^p$  tels que

1.  $\mu^* \geq 0$ .
2.  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ .
3.  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$ .
4. Pour tout  $y \in T(x^*)$ , on a  $y^T \mathbf{L}(x^*, \lambda^*, \mu^*) y \geq 0$ .

**Théorème 2.3.** (Conditions suffisantes du seconde ordre). Soient  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $m \leq n$ , et  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  des fonctions de classe  $\mathcal{C}^2$ .

1. On suppose qu'il existe  $x^* \in \mathbb{R}^n$ ,  $\lambda^* \in \mathbb{R}^m$  et  $\mu^* \in \mathbb{R}^p$  tels que :

- (a)  $\mu^* \geq 0$ .
- (b)  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ .
- (c)  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$ .
- (d) Pour tout  $y \in \tilde{T}(x^*, \mu^*)$ ,  $y \neq 0$ , nous avons  $y^T \mathbf{L}(x^*, \lambda^*, \mu^*) y > 0$  avec

$$\begin{aligned} \tilde{T}(x^*, \mu^*) &= \left\{ y \in \mathbb{R}^n : Dh(x^*)y = 0, Dg_j(x^*)y = 0, i \in \tilde{J}(x^*, \mu^*) \right\}, \\ \tilde{J}(x^*, \mu^*) &= \{i : g_i(x^*) = 0, \mu_i^* > 0\}. \end{aligned}$$

Alors,  $x^*$  est un point de **minimum local strict** du problème de minimisation avec contraintes d'inégalité suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \leq 0. \end{cases}$$

2. On suppose qu'il existe  $x^* \in \mathbb{R}^n$ ,  $\lambda^* \in \mathbb{R}^m$  et  $\mu^* \in \mathbb{R}^p$  tels que :

(a)  $\mu^* \leq 0$ .

(b)  $Df(x^*) + (\lambda^*)^T Dh(x^*) + (\mu^*)^T Dg(x^*) = 0^T$ .

(c)  $\mu_i^* g_i(x^*) = 0$  pour tout  $i = 1, \dots, p$ .

(d) Pour tout  $y \in \tilde{T}(x^*, \mu^*)$ ,  $y \neq 0$ , nous avons  $y^T L(x^*, \lambda^*, \mu^*) y < 0$

Alors,  $x^*$  est un point de **maximum local strict** du problème de minimisation avec contraintes d'inégalité suivant :

$$\begin{cases} \max f(x), \\ h(x) = 0, \\ g(x) \leq 0. \end{cases}$$

**Exemple 2.2.** Résoudre le problème suivant :

$$\begin{cases} \min (x-1)^2 + y - 2 \\ y - x = 1, \\ x + y \leq 2. \end{cases}$$

1. on pose  $f(x, y) = (x-1)^2 + y - 2$ ,  $h(x, y) = y - x - 1$  et  $g(x, y) = x + y - 2$ . Les conditions de KKT sont données par :

$$\begin{cases} \mu \geq 0, \\ \nabla f(x, y) + \lambda \nabla h(x, y) + \mu \nabla g(x, y) = 0, \\ \mu g(x, y) = 0, \\ h(x, y) = 0, \\ g(x, y) \leq 0. \end{cases} \Leftrightarrow \begin{cases} \mu \geq 0, \\ 2(x-1) - \lambda + \mu = 0, \\ 1 + \lambda + \mu = 0, \\ \mu(x+y-2) = 0, \\ y - x = 1, \\ x + y \leq 2. \end{cases}$$

(a) Si  $\mu > 0$ , nous avons :

$$\begin{cases} y - x = 1, \\ x + y = 2. \end{cases} \Leftrightarrow x = 1/2 \text{ et } y = 3/2.$$

D'autre part, on a :

$$\begin{cases} \lambda + \mu = -1, \\ \lambda - \mu = -1 \end{cases} \Leftrightarrow \lambda = -1 \text{ et } \mu = 0, \text{ contradiction}$$

(b) Si  $\mu = 0$ , alors  $\lambda = -1$ . Donc,  $x = 1/2$  et  $y = 3/2$ .

2. Nous avons :

$$L(x^*, y^*, \lambda^*, \mu^*) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$\tilde{T}(x^*, y^*, \mu^*) = \{y : Dh(x^*, y^*)y = 0\} = \{y : \begin{bmatrix} -1 & 1 \end{bmatrix} y = 0\} = \{(a, a)^T : a \in \mathbb{R}\}.$$

Nous avons :

$$y^T L(x^*, y^*, \lambda^*, \mu^*) y = \begin{bmatrix} a & a \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a \\ a \end{bmatrix} = 2a^2 > 0, \forall a \neq 0.$$

D'où  $(1/2, 3/2)$  est un point de minimum local strict.