

DUALITÉ LAGRANGIENNE

Dans ce chapitre, nous intéressons au problème d'optimisation avec contraintes suivant :

$$\begin{cases} \min f(x), \\ h(x) = 0, \\ g(x) \leq 0, \end{cases} \quad (4.1)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ avec $m \leq n$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ sont des fonction de classe $\mathcal{C}^1$.

On introduit la fonction de Lagrange (ou le Lagrangien) associé au problème (4.1) et définie par :

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{i=1}^p \mu_i g_i(x), \quad (4.2)$$

où $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ et $\mu \in \mathbb{R}_+^p$.

Remarque 4.1. Trouver une solution du problème (4.1) consiste à déterminer un point optimal x^* et les multiplicateurs associés λ^* et μ^* . Les approches primales de résolution se concentrent sur la détermination du point x^* , les multiplicateurs λ et μ ne servant souvent qu'à vérifier l'optimalité de x^* . Les méthodes duales quant à elles, mettent l'accent sur la recherche des multiplicateurs optimaux.

4.1 Problème primal et problème dual

Définition 4.1. (Fonction primale et problème primal). On appelle fonction primale associé au problème (4.1), la fonction définie par :

$$L_P(x) := \sup_{(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p} \mathcal{L}(x, \lambda, \mu) = \begin{cases} f(x) & \text{si } x \in \Omega, \\ +\infty & \text{si } x \notin \Omega, \end{cases} \quad (4.3)$$

où Ω est le domaine admissible primal défini par :

$$\Omega := \{x \in \mathbb{R}^n : L_P(x) < +\infty\}.$$

Dans cas, on a :

$$\min_{x \in \Omega} f(x) = \inf_{x \in \Omega} L_P(x)$$

Par conséquent, le problème primal est défini par :

$$\min_{x \in \Omega} f(x) = \inf_{x \in \Omega} \left[\sup_{(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p} \mathcal{L}(x, \lambda, \mu) \right]. \quad (4.4)$$

Définition 4.2. (Fonction duale et problème dual). On appelle fonction duale associé au problème (4.1), la fonction définie par :

$$L_D(\lambda, \mu) := \begin{cases} \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) & \text{si } (\lambda, \mu) \in \Omega^*, \\ -\infty & \text{si } (\lambda, \mu) \notin \Omega^*, \end{cases} \quad (4.5)$$

où Ω^* est le domaine admissible dual défini par :

$$\Omega^* = \{(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p : L_D(\lambda, \mu) > -\infty\}.$$

Ainsi,

$$\max_{(\lambda, \mu) \in \Omega^*} L_D(\lambda, \mu) = \sup_{(\lambda, \mu) \in \Omega^*} L_D(\lambda, \mu)$$

On définit alors le problème dual associé au problème (4.1) de la façon suivante :

$$\max_{(\lambda, \mu) \in \Omega^*} L_D(\lambda, \mu) = \sup_{(\lambda, \mu) \in \Omega^*} \left[\inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) \right]. \quad (4.6)$$

La fonction duale (4.5) possède de nombreuses propriétés remarquables facilitant la solution du problème dual (4.6).

Lemme 4.1. Nous avons :

1. La fonction duale $L_D(\lambda, \mu)$ est concave.
2. Le domaine admissible dual Ω^* est un ensemble convexe.

Démonstration. 1. Soient $(\lambda_1, \mu_1), (\lambda_2, \mu_2) \in \mathbb{R}^m \times \mathbb{R}_+^p$ et $\alpha \in [0, 1]$. Pour tout $x \in \mathbb{R}^n$, nous avons :

$$\begin{aligned} & \mathcal{L}[x, (1-\alpha)\lambda_1 + \alpha\lambda_2, (1-\alpha)\mu_1 + \alpha\mu_2] \\ &= f(x) + [(1-\alpha)\lambda_1 + \alpha\lambda_2]^T h(x) + [(1-\alpha)\mu_1 + \alpha\mu_2]^T g(x) \\ &= (1-\alpha)[f(x) + \lambda_1^T h(x) + \mu_1^T g(x)] + \alpha[f(x) + \lambda_2^T h(x) + \mu_2^T g(x)] \\ &= (1-\alpha)\mathcal{L}(x, \lambda_1, \mu_1) + \alpha\mathcal{L}(x, \lambda_2, \mu_2). \end{aligned}$$

D'autre part, on a :

$$\begin{cases} L_D(\lambda_1, \mu_1) \leq \mathcal{L}(x, \lambda_1, \mu_1), \\ L_D(\lambda_2, \mu_2) \leq \mathcal{L}(x, \lambda_2, \mu_2) \end{cases}$$

donc, nous avons :

$$\begin{aligned} (1-\alpha)L_D(\lambda_1, \mu_1) + \alpha L_D(\lambda_2, \mu_2) &\leq (1-\alpha)\mathcal{L}(x, \lambda_1, \mu_1) + \alpha\mathcal{L}(x, \lambda_2, \mu_2) \\ &= \mathcal{L}[x, (1-\alpha)\lambda_1 + \alpha\lambda_2, (1-\alpha)\mu_1 + \alpha\mu_2], \quad \forall x \in \mathbb{R}^n, \\ &\leq L_D[(1-\alpha)\lambda_1 + \alpha\lambda_2, (1-\alpha)\mu_1 + \alpha\mu_2], \end{aligned}$$

d'où $L_D(\lambda, \mu)$ est concave.

2. Soient $(\lambda_1, \mu_1), (\lambda_2, \mu_2) \in \Omega^*$ et $\alpha \in [0, 1]$. Nous avons :

$$\begin{cases} (\lambda_1, \mu_1) \in \Omega^* \Rightarrow L_D(\lambda_1, \mu_1) > -\infty, \\ (\lambda_2, \mu_2) \in \Omega^* \Rightarrow L_D(\lambda_2, \mu_2) > -\infty. \end{cases}$$

donc, on a :

$$L_D[(1-\alpha)\lambda_1 + \alpha\lambda_2, (1-\alpha)\mu_1 + \alpha\mu_2] \geq (1-\alpha)L_D(\lambda_1, \mu_1) + \alpha L_D(\lambda_2, \mu_2) > -\infty.$$

Par conséquent, nous avons :

$$(1-\alpha)(\lambda_1, \mu_1) + \alpha(\lambda_2, \mu_2) \in \Omega^* \Rightarrow \Omega^* \text{ est convexe.}$$

□

4.2 Exemples : cas linéaire

Exemple 4.1. On considère le problème d'optimisation linéaire suivant :

$$\begin{cases} \min c^T x, \\ Ax = b, \\ x \geq 0 \end{cases}$$

A est une matrice d'ordre $m \times n$ et $x, b, c \in \mathbb{R}^n$. Déterminer le problème dual.

La fonction de Lagrange est donnée par :

$$\mathcal{L}(x, \lambda, \mu) = c^T x + \lambda^T (b - Ax) - \mu^T x = \lambda^T b + (c - A^T \lambda - \mu)^T x.$$

la fonction dual est donnée par :

$$L_D(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) = \lambda^T b + \inf_{x \in \mathbb{R}^n} (c - A^T \lambda - \mu)^T x$$

Nous avons :

$$\inf_{x \in \mathbb{R}^n} (c - A^T \lambda - \mu)^T x \Leftrightarrow \mu = c - A^T \lambda.$$

Donc, le problème dual est donnée par :

$$\begin{cases} \max \lambda^T b, \\ A^T \lambda \leq c. \end{cases}$$

Exemple 4.2. Soit le problème d'optimisation linéaire suivant :

$$\begin{cases} \min c^T x \\ Ax \geq b, \\ x \geq 0. \end{cases}$$

La fonction de Lagrange est donnée par :

$$\mathcal{L}(x, \nu, \mu) = c^T x + \nu^T (b - Ax) - \mu^T x = \nu^T b + (c - A^T \nu - \mu)^T x.$$

Donc, la fonction duale est donnée par :

$$L_D(\nu, \mu) = \nu^T b + \inf_{x \in \mathbb{R}^n} (c - A^T \nu - \mu)^T x.$$

Nous avons :

$$\inf_{x \in \mathbb{R}^n} (c - A^T \nu - \mu)^T x = 0 \Leftrightarrow \mu = c - A^T \nu \text{ et } \nu \geq 0.$$

Par conséquent, le problème dual est donné par :

$$\begin{cases} \max \nu^T b \\ A^T \nu \leq c, \\ \nu \geq 0. \end{cases}$$

4.3 Exemples : cas quadratique

Exemple 4.3. On considère le problème d'optimisation quadratique suivant :

$$\begin{cases} \min \frac{1}{2}x^T Qx + c^T x \\ Ax \leq b, \end{cases}$$

où Q est une matrice d'ordre n , symétrique et définie positive et $c \in \mathbb{R}^n$, A est une matrice d'ordre $m \times n$ et $b \in \mathbb{R}^m$. Déterminer le problème dual.

La fonction de Lagrange est donnée par :

$$\mathcal{L}(x, \mu) = \frac{1}{2}x^T Qx + c^T x + \mu^T (Ax - b) = -\mu^T b + \frac{1}{2}x^T Qx + c^T x + \mu^T Ax.$$

La fonction duale est donnée par :

$$L_D(\mu) = -\mu^T b + \inf_{x \in \mathbb{R}^n} \left[\frac{1}{2}x^T Qx + c^T x + \mu^T Ax \right].$$

Nous avons :

$$\inf_{x \in \mathbb{R}^n} \left[\frac{1}{2}x^T Qx + c^T x + \mu^T Ax \right] \text{ admet une solution unique } x^* = -Q^{-1}(A^T \mu + c).$$

D'autre part, nous avons :

$$\langle Qx^*, x^* \rangle + \langle c, x^* \rangle + \langle Ax^*, \mu \rangle = 0$$

Alors, la fonction dual est donnée par :

$$\begin{aligned} L_D(\mu) &= -\mu^T b - \frac{1}{2} \langle Qx^*, x^* \rangle \\ &= -\mu^T b - \frac{1}{2} (\mu^T A + c^T) Q^{-1} (A^T \mu + c) \\ &= -\mu^T b - \frac{1}{2} (\mu^T A^T Q^{-1} A \mu) - \frac{1}{2} (c^T Q^{-1} c) - c^T Q^{-1} A^T \mu. \end{aligned}$$

Ainsi, le problème duale est donné par :

$$\begin{cases} \max -\mu^T b - \frac{1}{2} (\mu^T A^T Q^{-1} A \mu) - \frac{1}{2} (c^T Q^{-1} c) - c^T Q^{-1} A^T \mu \\ \mu \geq 0. \end{cases}$$

4.4 Liens entre problème primal et problème dual

Proposition 4.1. Pour tout $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p$, nous avons :

$$L_D(\lambda, \mu) \leq \inf_{x \in \Omega} L_P(x). \quad (4.7)$$

Dans ce cas, la différence suivante :

$$\sup_{(\lambda, \mu) \in \Omega^*} L_D(\lambda, \mu) - \inf_{x \in \Omega} L_P(x)$$

est appelé le saut de dualité.

Démonstration. Soit la fonction de Lagrange suivante :

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x).$$

Pour x un point réalisable de problème (4.1) et pour chaque $\lambda \in \mathbb{R}^m$ et $\mu \in \mathbb{R}_+^p$, nous avons :

$$\lambda^T h(x) + \mu^T g(x) = \mu^T g(x) \leq 0,$$

alors, $\mu \geq 0$ et $g(x) \leq 0$, on a :

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \mu^T g(x) \leq f(x), \quad \text{pour tout } x \text{ réalisable.}$$

Par conséquent, nous avons :

$$L_D(\lambda, \mu) = \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \leq \inf_{x \in \Omega} f(x) = \inf_{x \in \Omega} L_P(x).$$

□

Définition 4.3. (Point-selle). Soient $\bar{x} \in \mathbb{R}^n$ et $(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^m \times \mathbb{R}_+^p$. Le point $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point de selle de la fonction de Lagrange si :

$$\forall x \in \mathbb{R}^n, \forall (\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p : \mathcal{L}(\bar{x}, \lambda, \mu) \leq \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) \leq \mathcal{L}(x, \bar{\lambda}, \bar{\mu}).$$

Autrement dit, si :

$$\max_{(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p} \mathcal{L}(\bar{x}, \lambda, \mu) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \bar{\lambda}, \bar{\mu})$$

Une condition suffisante d'existence d'une solution du problème (4.1) est donnée par le théorème suivant :

Théorème 4.1. (Condition suffisante d'existence d'une solution de (4.1)). Si $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point de selle de la fonction de Lagrange associé au problème (4.1), alors $\bar{x}$ est une solution du problème (4.1) et

$$\bar{\mu}_i g_i(\bar{x}) = 0, \quad i = 1, 2, \dots, p.$$

Démonstration. Supposons que $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point de selle de la fonction Lagrangienne. Alors, pour tout $(\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}_+^p$, on a :

$$\mathcal{L}(\bar{x}, \lambda, \mu) \leq \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}),$$

donc,

$$(\lambda - \bar{\lambda})^T h(\bar{x}) + (\mu - \bar{\mu})^T g(\bar{x}) \leq 0,$$

ce qui implique :

$$\begin{cases} \forall \lambda \in \mathbb{R}^m : (\lambda - \bar{\lambda})^T h(\bar{x}) \leq 0 \\ \forall \mu \in \mathbb{R}_+^p : (\mu - \bar{\mu})^T g(\bar{x}) \leq 0 \end{cases} \Rightarrow h(\bar{x}) = 0 \text{ et } g(\bar{x}) \leq 0. \quad (4.8)$$

Donc le point $\bar{x}$ est un point admissible du problème (4.1) et

$$\mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}) + \bar{\mu}^T g(\bar{x})$$

De plus, on pose dans (4.8) $\mu = 0$, on obtient :

$$\bar{\mu}^T g(\bar{x}) \geq 0 \text{ avec } \bar{\mu} \geq 0.$$

et on pose dans (4.8) $\mu = 2\bar{\mu}$, on obtient :

$$\bar{\mu}^T g(\bar{x}) \leq 0 \text{ avec } \bar{\mu} \geq 0.$$

Par conséquent, $\bar{\mu}^T g(\bar{x}) = 0$. Alors, nous avons :

$$\mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}).$$

D'autre part, on a :

$$\forall x \in \Omega : f(\bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) \leq \mathcal{L}(x, \bar{\lambda}, \bar{\mu}) = f(x) + \underbrace{\bar{\mu}^T g(x)}_{\leq 0} \leq f(x)$$

D'où le point $\bar{x}$ est solution du problème (4.1). □

Remarque 4.2. 1. Il n'existe pas toujours de point selle, et s'il en existe un, il n'est pas forcément unique.
2. La réciproque de Théorème 4.1 n'est pas forcément vraie, sauf dans le cas particulier de fonctions convexes.

Théorème 4.2. (Théorème de dualité).

1. Le point $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point-selle de la fonction Lagrangienne si et seulement si :
 - (a) $\bar{x}$ est solution du problème primal.
 - (b) $(\bar{\lambda}, \bar{\mu})$ est solution du problème dual.
 - (c) Le saut de dualité est nul c'est-à-dire :

$$\sup_{(\lambda, \mu) \in \Omega^*} L_D(\lambda, \mu) = \inf_{x \in \Omega} L_P(x)$$

2. Si le problème (4.1) admet une solution $\bar{x} \in \Omega$, et s'il existe $\bar{\lambda} \in \mathbb{R}^m$ et $\bar{\mu} \in \mathbb{R}_+^p$ tels que :

$$f(\bar{x}) = \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \bar{\lambda}, \bar{\mu}),$$

alors $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point-selle de la fonction Lagrangienne.

Démonstration. 1. On suppose que $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point-selle de la fonction de Lagrange. Alors, nous avons :

$$\max_{(\lambda, \mu) \in \Omega^*} \mathcal{L}(\bar{x}, \lambda, \mu) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \min_{x \in \Omega} \mathcal{L}(x, \bar{\lambda}, \bar{\mu}), \quad (4.9)$$

et d'après Théorème 4.1, on a :

$$\mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}) = \min_{x \in \Omega} \mathcal{L}(x, \bar{\lambda}, \bar{\mu}) \quad \text{et} \quad \bar{\mu}_i^T g_i(\bar{x}) = 0, \quad i = 1, 2, \dots, p. \quad (4.10)$$

- (a) D'après les relations (4.9) et (4.10), on a :

$$\forall x \in \Omega : \max_{(\lambda, \mu) \in \Omega^*} \mathcal{L}(x, \lambda, \mu) \geq \mathcal{L}(x, \bar{\lambda}, \bar{\mu}) \geq \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}),$$

soit

$$\min_{x \in \Omega} \left[\max_{(\lambda, \mu) \in \Omega^*} \mathcal{L}(x, \lambda, \mu) \right] \geq \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}).$$

En conséquence :

$$\min_{x \in \Omega} \left[\max_{(\lambda, \mu) \in \Omega^*} \mathcal{L}(x, \lambda, \mu) \right] = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}),$$

et $\bar{x}$ est un point solution du problème primal.

(b) De même, sachant que

$$f(\bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \max_{(\lambda, \mu) \in \Omega^*} \mathcal{L}(x, \lambda, \mu),$$

on a :

$$\forall (\lambda, \mu) \in \Omega^* : \min_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \leq \mathcal{L}(\bar{x}, \lambda, \mu) \leq f(\bar{x}),$$

Ce qui implique,

$$\max_{(\lambda, \mu) \in \Omega^*} \left[\min_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \right] \leq f(\bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}).$$

En conséquence :

$$f(\bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = \max_{(\lambda, \mu) \in \Omega^*} \left[\min_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \right].$$

Autrement dit, $(\bar{\lambda}, \bar{\mu})$ est solution du problème dual.

(c) On a démontré que :

$$f(\bar{x}) = \max_{(\lambda, \mu) \in \Omega^*} \left[\min_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \right] = \min_{x \in \Omega} \left[\max_{(\lambda, \mu) \in \Omega^*} \mathcal{L}(x, \lambda, \mu) \right].$$

Alors, Le saut de dualité est nul.

2. Démonstration laissée en exercice.

□

Remarque 4.3. 1. L'intérêt du problème dual est double : il est concave et les contraintes sont simples à prendre en compte (contraintes de positivité dans le cas général, pas de contrainte dans le cas de contraintes d'égalité).

2. Pour le résoudre, on peut par exemple utiliser :

- (a) une méthode de gradient dans le cas de contraintes d'égalité : c'est l'algorithme d'Uzawa.
- (b) une méthode de gradient projeté dans le cas de contraintes d'inégalité.
- (c) des méthodes de type Newton.