# MA 5037: Optimization Methods and Applications
# Duality

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: August 03, 2023/Last updated: June 22, 2025

# The primal problem and its Lagrangian function

We explore the dual problem by considering the general model:

$$
\begin{aligned}
f^* = \min\ & f(\boldsymbol{x}) \\
\text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0, \quad i = 1, 2, \cdots, m, \\
& h_j(\boldsymbol{x}) = 0, \quad j = 1, 2, \cdots, p, \\
& \boldsymbol{x} \in \boldsymbol{X},
\end{aligned} \tag{1}
$$

*where $f, g_i, h_j (i = 1, 2, \cdots, m, j = 1, 2, \cdots, p)$ are functions defined on the set $\boldsymbol{X} \subseteq \mathbb{R}^n$. Problem (1) will be referred to as the primal problem.*

The Lagrangian function of the problem is defined as

$$
L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{p} \mu_j h_j(\boldsymbol{x}), \quad (\boldsymbol{x} \in \boldsymbol{X}, \boldsymbol{\lambda} \in \mathbb{R}^m_+, \boldsymbol{\mu} \in \mathbb{R}^p)
$$

*where $\lambda_1, \lambda_2, \cdots, \lambda_m$ are nonnegative Lagrange multipliers associated with the inequality constraints, and $\mu_1, \mu_2, \cdots, \mu_p$ are the Lagrange multipliers associated with the equality constraints.*

## Definition of the dual problem

The dual objective function $q : \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R} \bigcup \{-\infty\}$ is defined to be

$$q(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}). \tag{2}$$

There may be values of $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ for which $q(\boldsymbol{\lambda}, \boldsymbol{\mu}) = -\infty$, we define the domain of the dual objective function as

$$dom(q) = \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^m_+ \times \mathbb{R}^p : q(\boldsymbol{\lambda}, \boldsymbol{\mu}) > -\infty\}.$$

*Then the dual problem is given by*

$$q^* = \max\ q(\boldsymbol{\lambda}, \boldsymbol{\mu}) \quad \text{subject to } (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in dom(q). \tag{3}$$

*That is,*

$$q^* = \max_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in dom(q)} \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}).$$

## Convexity of the dual problem

**Theorem:** *Consider the problem (1) with $f, g_i, h_j$ $(i = 1, 2, \cdots, m,$ $j = 1, 2, \cdots, p)$ being functions defined on the set $\boldsymbol{X} \subseteq \mathbb{R}^n$, and let $q$ be the function defined in (2). Then we have*

(a) $dom(q) = \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^m_+ \times \mathbb{R}^p : q(\boldsymbol{\lambda}, \boldsymbol{\mu}) > -\infty\}$ *is a convex set,*

(b) $q(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ *is a concave function over $dom(q)$.*

*Proof:*
(a) Taking $(\boldsymbol{\lambda}_1, \boldsymbol{\mu}_1), (\boldsymbol{\lambda}_2, \boldsymbol{\mu}_2) \in dom(q)$ and $\alpha \in [0, 1]$, we have

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) > -\infty \quad \text{and} \quad \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}_2, \boldsymbol{\mu}_2) > -\infty.$$

Since $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ is affine with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, we obtain

$$\begin{aligned} & q(\alpha\boldsymbol{\lambda}_1 + (1-\alpha)\boldsymbol{\lambda}_2, \alpha\boldsymbol{\mu}_1 + (1-\alpha)\boldsymbol{\mu}_2) \\ &= \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \alpha\boldsymbol{\lambda}_1 + (1-\alpha)\boldsymbol{\lambda}_2, \alpha\boldsymbol{\mu}_1 + (1-\alpha)\boldsymbol{\mu}_2) \\ &= \min_{\boldsymbol{x} \in \boldsymbol{X}} \big(\alpha L(\boldsymbol{x}, \boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) + (1-\alpha) L(\boldsymbol{x}, \boldsymbol{\lambda}_2, \boldsymbol{\mu}_2)\big). \end{aligned}$$

# Proof of the convexity of the dual problem

Therefore, we have

$$\begin{aligned} q(\alpha\boldsymbol{\lambda}_1 + (1-\alpha)\boldsymbol{\lambda}_2, \alpha\boldsymbol{\mu}_1 + (1-\alpha)\boldsymbol{\mu}_2) \\ \geq \alpha \min_{\boldsymbol{x}\in X} L(\boldsymbol{x}, \boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) + (1-\alpha)\min_{\boldsymbol{x}\in X} L(\boldsymbol{x}, \boldsymbol{\lambda}_2, \boldsymbol{\mu}_2) \\ = \alpha q(\boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) + (1-\alpha) q(\boldsymbol{\lambda}_2, \boldsymbol{\mu}_2) > -\infty. \end{aligned}$$

Hence, $\alpha(\boldsymbol{\lambda}_1, \boldsymbol{\mu}_1) + (1-\alpha)(\boldsymbol{\lambda}_2, \boldsymbol{\mu}_2) \in dom(q)$, and the convexity of $dom(q)$ is established.

(b) As noted in the proof of part (a), $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ is an affine function with respect to $(\boldsymbol{\lambda}, \boldsymbol{\mu})$. In particular, it is a concave function with respect to $(\boldsymbol{\lambda}, \boldsymbol{\mu})$. *Since $q(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is the minimum of concave functions, it must be concave.* □

# Weak duality theorem

**Theorem:** *Consider the primal problem (1) and its dual problem (3). Then $q^* \le f^*$, where $q^*$ and $f^*$ are optimal dual and primal values, respectively.*

*Proof:* Let us denote the feasible set of the primal problem by

$$S = \{\boldsymbol{x} \in \boldsymbol{X} : g_i(\boldsymbol{x}) \le 0, h_j(\boldsymbol{x}) = 0, 1 \le i \le m, 1 \le j \le p\}.$$

Then for any $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^m \times \mathbb{R}^p$, we have

$$\begin{aligned} q(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \le \min_{\boldsymbol{x} \in \boldsymbol{S}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \\ &= \min_{\boldsymbol{x} \in \boldsymbol{S}} \Big( f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{p} \mu_j h_j(\boldsymbol{x}) \Big) \\ &\le \min_{\boldsymbol{x} \in \boldsymbol{S}} f(\boldsymbol{x}). \end{aligned}$$

We thus obtain that

$$q(\boldsymbol{\lambda}, \boldsymbol{\mu}) \le \min_{\boldsymbol{x} \in \boldsymbol{S}} f(\boldsymbol{x}) = f^*$$

for any $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^m \times \mathbb{R}^p$. By taking the maximum over $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, we have $q^* \le f^*$. $\square$

# Supporting hyperplane theorem

**Theorem:** *Let $\boldsymbol{C} \subseteq \mathbb{R}^n$ be a convex set with nonempty interior and let $\boldsymbol{y} \notin \boldsymbol{C}$. Then there exists $\mathbf{0} \neq \boldsymbol{p} \in \mathbb{R}^n$ such that*

$$\boldsymbol{p}^\top \boldsymbol{x} \leq \boldsymbol{p}^\top \boldsymbol{y} \text{ for any } \boldsymbol{x} \in \boldsymbol{C}.$$

*Proof*: Since $\boldsymbol{y} \notin \boldsymbol{C} \supseteq int(\boldsymbol{C}) = int(cl(\boldsymbol{C}))$, it follows that $\boldsymbol{y} \notin int(cl(\boldsymbol{C}))$. Therefore, there exists a sequence $\{\boldsymbol{y}_k\}_{k \geq 1}$ satisfying $\boldsymbol{y}_k \notin cl(\boldsymbol{C})$ such that $\boldsymbol{y}_k \to \boldsymbol{y}$ as $k \to \infty$.

Since $cl(\boldsymbol{C})$ is convex and closed, it follows by the strict separation theorem (Theorem 10.1) that there exists $\mathbf{0} \neq \boldsymbol{p}_k \in \mathbb{R}^n$ such that

$$\boldsymbol{p}_k^\top \boldsymbol{x} < \boldsymbol{p}_k^\top \boldsymbol{y}_k, \quad \forall\, \boldsymbol{x} \in cl(\boldsymbol{C}).$$

Dividing the latter inequality by $\|\boldsymbol{p}_k\| \neq 0$, we obtain

$$\frac{\boldsymbol{p}_k^\top}{\|\boldsymbol{p}_k\|}(\boldsymbol{x} - \boldsymbol{y}_k) < 0, \quad \forall\, \boldsymbol{x} \in cl(\boldsymbol{C}). \tag{4}$$

# Proof of the supporting hyperplane theorem (cont'd)

Since the sequence $\{\frac{p_k}{\|p_k\|}\}_{k\geq 1}$ is bounded, it follows that there exists a subsequence $\{\frac{p_k}{\|p_k\|}\}_{k\in T}$ such that $\frac{p_k}{\|p_k\|} \to p$ as $k \xrightarrow{T} \infty$ for some $p \in \mathbb{R}^n$. Hence, $\|p\| = 1$ and in particular $p \neq \mathbf{0}$. Taking the limit as $k \xrightarrow{T} \infty$ in inequality (4), we obtain that

$$p^\top(x - y) \leq 0, \quad \forall\, x \in cl(C),$$

which implies the result since $C \subseteq cl(C)$. $\square$

## Separation of two convex sets

**Theorem:** *Let $\boldsymbol{C}_1, \boldsymbol{C}_2 \subseteq \mathbb{R}^n$ be two convex sets with nonempty interior such that $\boldsymbol{C}_1 \cap \boldsymbol{C}_2 = \emptyset$. Then there exists $\mathbf{0} \neq \boldsymbol{p} \in \mathbb{R}^n$ for which*

$$\boldsymbol{p}^\top \boldsymbol{x} \leq \boldsymbol{p}^\top \boldsymbol{y}, \quad \forall\, \boldsymbol{x} \in \boldsymbol{C}_1, \boldsymbol{y} \in \boldsymbol{C}_2.$$

*Proof:* The set $\boldsymbol{C}_1 - \boldsymbol{C}_2$ is a convex set with nonempty interior, and since $\boldsymbol{C}_1 \cap \boldsymbol{C}_2 = \emptyset$, it follows that $\mathbf{0} \notin \boldsymbol{C}_1 - \boldsymbol{C}_2$. By the supporting hyperplane theorem, it follows that there exists $\mathbf{0} \neq \boldsymbol{p} \in \mathbb{R}^n$ such that

$$\boldsymbol{p}^\top (\boldsymbol{x} - \boldsymbol{y}) \leq \boldsymbol{p}^\top \mathbf{0} = 0, \quad \forall\, \boldsymbol{x} \in \boldsymbol{C}_1, \boldsymbol{y} \in \boldsymbol{C}_2,$$

which is the same as the desired result. □

# A nonlinear version of Farkas lemma

**Theorem** (nonlinear Farkas lemma): *Let $\boldsymbol{X} \subseteq \mathbb{R}^n$ be a convex set and let $f, g_1, g_2, \cdots, g_m$ be convex functions over $\boldsymbol{X}$. Assume that there exists $\hat{\boldsymbol{x}} \in \boldsymbol{X}$ such that*

$$g_i(\hat{\boldsymbol{x}}) < 0, \quad i = 1, 2, \cdots, m.$$

*Let $c \in \mathbb{R}$. Then the following two claims are equivalent.*

(a) *The following implication holds :*

$$\boldsymbol{x} \in \boldsymbol{X}, \;\; g_i(\boldsymbol{x}) \leq 0, \;\; i = 1, 2, \cdots, m \implies f(\boldsymbol{x}) \geq c.$$

(b) *There exist $\lambda_1, \lambda_2, \cdots, \lambda_m \geq 0$ such that*

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} \left( f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) \right) \geq c. \tag{5}$$

# Proof of the nonlinear Farkas lemma

(b) $\Rightarrow$ (a): Suppose that there exist $\lambda_1, \lambda_2, \cdots, \lambda_m \geq 0$ such that (5) holds. Let $\boldsymbol{x} \in \boldsymbol{X}$ satisfy $g_i(\boldsymbol{x}) \leq 0, i = 1, 2, \cdots, m$. Then by (5) we have

$$f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) \geq c.$$

Since $g_i(\boldsymbol{x}) \leq 0$ and $\lambda_i \geq 0$ for $i = 1, 2, \cdots, m$,

$$f(\boldsymbol{x}) \geq c - \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) \geq c.$$

(a) $\Rightarrow$ (b): Assume that (a) holds. Consider the following two sets:

$$\begin{aligned}
\boldsymbol{S} &:= \{\boldsymbol{u} = (u_0, u_1, \cdots, u_m) : \exists\, \boldsymbol{x} \in \boldsymbol{X} \text{ s.t. } f(\boldsymbol{x}) \leq u_0, g_i(\boldsymbol{x}) \leq u_i, 1 \leq i \leq m\}, \\
\boldsymbol{T} &:= \{(u_0, u_1, \cdots, u_m) : u_0 < c, u_1 \leq 0, u_2 \leq 0, \cdots, u_m \leq 0\}.
\end{aligned}$$

# Proof of the nonlinear Farkas lemma (cont'd)

Note that $\boldsymbol{S}$ and $\boldsymbol{T}$ are convex with nonempty interiors and by (a), $\boldsymbol{S} \cap \boldsymbol{T} = \emptyset$. Therefore, by the separation theorem of two convex sets, it follows that there exists a vector $\boldsymbol{a} = (a_0, a_1, \cdots, a_m)^\top \neq \mathbf{0}$ such that

$$\min_{(u_0,u_1,\cdots,u_m)\in \boldsymbol{S}} \sum_{j=0}^{m} a_j u_j \geq \max_{(u_0,u_1,\cdots,u_m)\in \boldsymbol{T}} \sum_{j=0}^{m} a_j u_j. \qquad (6)$$

**Claim:** $\boldsymbol{a} \geq \mathbf{0}$. Consider the RHS of (6). Suppose that there exists an $a_i < 0$. By taking $u_i$ to be a negative number tending to $-\infty$ while fixing all the other components as zeros, we obtain that the RHS of (6) is $\infty$, which is a contradiction.

Since $\boldsymbol{a} \geq \mathbf{0}$, it follows that RHS of (6) $= a_0 c$, and we thus obtain

$$\min_{(u_0,u_1,\cdots,u_m)\in \boldsymbol{S}} \sum_{j=0}^{m} a_j u_j \geq a_0 c. \qquad (7)$$

# Proof of the nonlinear Farkas lemma (cont'd)

We will show that $a_0 > 0$. Suppose in contradiction that $a_0 = 0$. Then

$$\min_{(u_0, u_1, \cdots, u_m) \in S} \sum_{j=1}^{m} a_j u_j \geq 0.$$

Since we can take $u_i = g_i(\hat{\boldsymbol{x}})$, $i = 1, 2, \cdots, m$, it leads to

$$\sum_{j=1}^{m} a_j g_j(\hat{\boldsymbol{x}}) \geq 0.$$

which is impossible since $g_j(\hat{\boldsymbol{x}}) < 0 \; \forall \, j$ and $\boldsymbol{a} = (a_0, a_1, \cdots, a_m)^\top \neq \mathbf{0}$.

Now we can divide (7) by $a_0 > 0$ to obtain

$$\min_{(u_0, u_1, \cdots, u_m) \in S} \left( u_0 + \sum_{j=1}^{m} \tilde{a}_j u_j \right) \geq c, \quad \tilde{a}_j := \frac{a_j}{a_0}. \tag{8}$$

# Proof of the nonlinear Farkas lemma (cont'd)

Define a subset $\tilde{\boldsymbol{S}} \subseteq \boldsymbol{S}$ by

$$
\begin{aligned}
\tilde{\boldsymbol{S}} &:= \{\boldsymbol{u} = (u_0, u_1, \cdots, u_m) : \exists\, \boldsymbol{x} \in \boldsymbol{X} \text{ s.t. } f(\boldsymbol{x}) = u_0, g_i(\boldsymbol{x}) = u_i, 1 \le i \le m\} \\
&= \{(f(\boldsymbol{x}), g_1(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x})) : \boldsymbol{x} \in \boldsymbol{X}\}. \quad (\star)
\end{aligned}
$$

Then we have

$$
\begin{aligned}
\min_{(u_0, u_1, \cdots, u_m) \in \boldsymbol{S}} \left( u_0 + \sum_{j=1}^{m} \tilde{a}_j u_j \right) &\le \min_{(u_0, u_1, \cdots, u_m) \in \tilde{\boldsymbol{S}}} \left( u_0 + \sum_{j=1}^{m} \tilde{a}_j u_j \right) \\
\text{by } (\star) \quad &= \min_{\boldsymbol{x} \in \boldsymbol{X}} \left( f(\boldsymbol{x}) + \sum_{j=1}^{m} \tilde{a}_j g_j(\boldsymbol{x}) \right),
\end{aligned}
$$

which combined with (8) yields the desired result

$$
\min_{\boldsymbol{x} \in \boldsymbol{X}} \left( f(\boldsymbol{x}) + \sum_{j=1}^{m} \tilde{a}_j g_j(\boldsymbol{x}) \right) \ge c. \quad \square
$$

## Strong duality of convex problems (inequality constraints)

**Theorem:** *Consider the optimization problem*

$$\begin{aligned} f^* &= \min f(\boldsymbol{x}) \\ &\quad \text{subject to} \quad g_i(\boldsymbol{x}) \le 0,\ i = 1, 2, \cdots, m, \\ &\qquad\qquad\qquad \boldsymbol{x} \in \boldsymbol{X}, \end{aligned} \tag{9}$$

*where* $\boldsymbol{X}$ *is a convex set and* $f, g_i, i = 1, 2, \cdots, m$, *are convex functions over* $\boldsymbol{X}$. *Suppose that there exists* $\hat{\boldsymbol{x}} \in \boldsymbol{X}$ *for which* $g_i(\hat{\boldsymbol{x}}) < 0, i = 1, 2, \cdots, m$. *Suppose that problem (9) has a finite optimal value. Then the optimal value of the dual problem*

$$q^* = \max\{q(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in dom(q)\}, \tag{10}$$

*where*

$$q(\boldsymbol{\lambda}) = \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}),$$

*is attained, and the optimal values of the primal and dual problems are the same:*

$$f^* = q^*.$$

# Proof of the strong duality of convex problems

Since problem (9) has a finite optimal value, we have $f^* > -\infty$. It follows that the following implication holds:

$$\boldsymbol{x} \in \boldsymbol{X},\ \ g_i(\boldsymbol{x}) \leq 0,\ \ i = 1, 2, \cdots, m \Longrightarrow f(\boldsymbol{x}) \geq f^*.$$

By the nonlinear Farkas's lamma, there exist $\tilde{\lambda}_1, \tilde{\lambda}_2, \cdots, \tilde{\lambda}_m \geq 0$ such that

$$q(\tilde{\boldsymbol{\lambda}}) = \min_{\boldsymbol{x} \in \boldsymbol{X}} \left( f(\boldsymbol{x}) + \sum_{j=1}^{m} \tilde{\lambda}_j g_j(\boldsymbol{x}) \right) \geq f^*,$$

which combined with the weak duality theorem yields

$$q^* \geq q(\tilde{\boldsymbol{\lambda}}) \geq f^* \geq q^*.$$

Hence, $f^* = q^*$ and $\tilde{\boldsymbol{\lambda}}$ is an optimal solution of the dual problem. $\square$

# Complementary slackness conditions

**Theorem:** *Consider problem (9) and assume that $f^* = q^*$, where $q^*$ is the optimal value of the dual problem given by (10). If $\boldsymbol{x}^*$ and $\boldsymbol{\lambda}^*$ are optimal solutions of the primal and dual problems, respectively, then*

$$\boldsymbol{x}^* \in \arg\min \; L(\boldsymbol{x}, \boldsymbol{\lambda}^*),$$
$$\lambda_i^* g_i(\boldsymbol{x}^*) = 0, \quad i = 1, 2. \cdots, m.$$

*Proof:* First, we have

$$\begin{aligned} q^* &= q(\boldsymbol{\lambda}^*) = \min_{\boldsymbol{x} \in \boldsymbol{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}^*) \leq L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \\ &= f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \lambda_i^* g_i(\boldsymbol{x}^*) \leq f(\boldsymbol{x}^*) = f^*. \end{aligned}$$

Since $f^* = q^*$, all the inequalities in the above chain are satisfied as equalities. It follows that $\boldsymbol{x}^* \in \arg\min \; L(\boldsymbol{x}, \boldsymbol{\lambda}^*)$, $\sum_{i=1}^{m} \lambda_i^* g_i(\boldsymbol{x}^*) = 0$. Because of $\lambda_i^* \geq 0$ and $g_i(\boldsymbol{x}^*) \leq 0 \; \forall \, i = 1, 2. \cdots, m$, we obtain $\lambda_i^* g_i(\boldsymbol{x}^*) = 0 \; \forall \, i = 1, 2. \cdots, m$. □