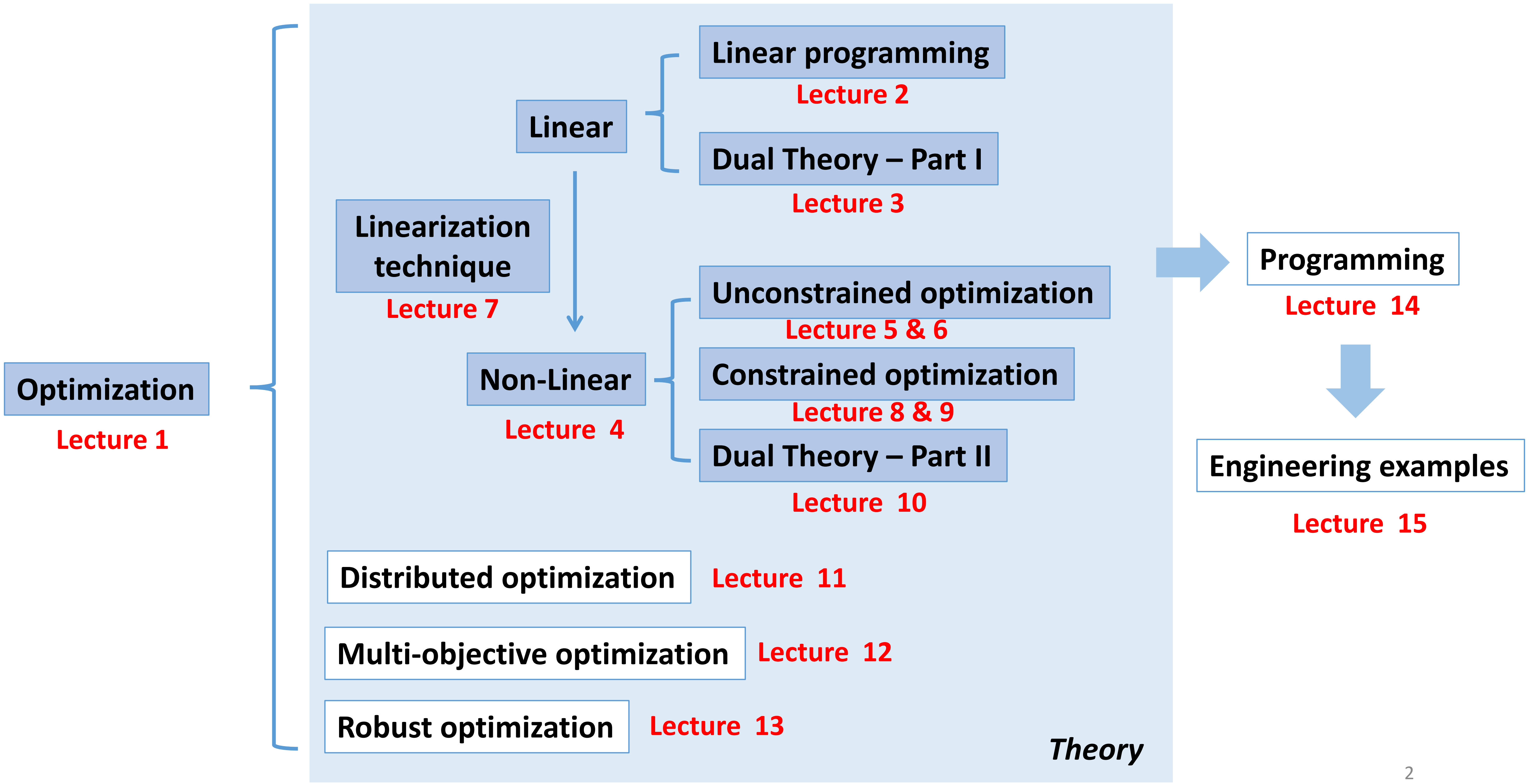
MAEG4070 Engineering Optimization

Lecture 10 Dual Theory – Part II

Yue Chen
MAE, CUHK
email: yuechen@mae.cuhk.edu.hk
Nov 2, 2022

Content of this course (tentative)



Dual Optimization

The primal problem:

$$\min_{x} \; f(x)$$
$$\text{s.t. } h_i(x) = 0, \forall i = 1, ..., m$$
$$g_j(x) \leq 0, \forall j = 1, ..., r$$

The Lagrangian function is

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{r} \mu_j g_j(x), \;\; \mu_j \geq 0, \forall j$$

We define the dual function as

$$Q(\lambda, \mu) = \min_{x} L(x, \lambda, \mu)$$

Then the dual optimization is

$$\max_{\lambda, \mu} \; Q(\lambda, \mu)$$
$$\text{s.t. } \mu \geq 0$$

Explanation – Finding the best lower bound

Still remember what we have learned in Lecture 3?
Dual problem can be interpreted as finding the best lower bound of a minimization problem.

Consider the optimization:

$$\min_x \ f(x)$$
$$\text{s.t. } g_i(x) \leq 0, \forall i = 1, ..., m$$

We want to find the best lower bound v.
Since v is a lower bound, we have

$$f(x) \geq v, \forall x \in \{x \mid g_i(x) \leq 0, \forall i = 1, ..., m\}$$

It means that the following equation has **no feasible solution**:

$$\begin{cases} f(x) < v \\ g_i(x) \leq 0, \forall i = 1, ..., m \end{cases}$$

Explanation – Finding the best lower bound

It means that the following equation has **no feasible solution**:

$$\begin{cases} f(x) < v \\ g_i(x) \leq 0, \forall i = 1, ..., m \end{cases} \tag{1}$$

It is further equivalent to the following condition has **no feasible solution**:

$$\exists \mu > 0,\ f(x) + \sum_{i=1}^{m} \mu_i g_i(x) < v \tag{2}$$

(Otherwise, if (2) has no feasible solution, but (1) has. Suppose the solution of (1) is x^*, it satisfies $f(x^*) < v, g_i(x^*) \leq 0, \forall i = 1, ..., m$; then for any $\mu > 0$, we have $f(x^*) + \sum_{i=1}^{m} \mu_i g_i(x^*) < v$; so x^* is also a feasible solution of (2). We have contradiction!!)

Explanation – Finding the best lower bound

Moreover, (2) has no feasible solution is equivalent to

$$\exists \mu > 0,\ f(x) + \sum_{i=1}^{m} \mu_i g_i(x) \geq v, \forall x$$

which is equivalent to

$$\exists \mu > 0,\ \underbrace{\min_x \{f(x) + \sum_{i=1}^{m} \mu_i g_i(x)\}}_{L(x,\mu)} \geq v$$

If a function is larger or equal to some value, the minimal of that function is still larger or equal to that value.

As we want to find the "best" lower bound, we want v to be as large as possible, so

$$\max_{\mu > 0} \min_x L(x, \mu)$$

Dual problem

Dual Optimization

Construct the dual optimization of this LP:

$$
\begin{aligned}
\min_x \ & c^T x \\
\text{s.t. } & A_1 x \geq b_1 \text{ or } \boxed{-A_1 x \leq -b_1} \\
& A_2 x = b_2 \text{ or } \boxed{-A_2 x = -b_2}
\end{aligned}
$$

The Lagrangian is

$$
\begin{aligned}
L(x, \mu, \lambda) = & \ c^T x - \mu^T (A_1 x - b_1) - \lambda^T (A_2 x - b_2) \\
= & \ (c^T - \mu^T A_1 - \lambda^T A_2) x + \mu^T b_1 + \lambda^T b_2
\end{aligned}
$$

If any element i of $c^T - \mu^T A_1 - \lambda^T A_2 > 0$, let $x_i \to -\infty$, $x_{j \neq i} = 0$, then $L = -\infty$.
If any element i of $c^T - \mu^T A_1 - \lambda^T A_2 < 0$, let $x_i \to \infty$, $x_{j \neq i} = 0$, then $L = -\infty$.
Therefore, $c^T - \mu^T A_1 - \lambda^T A_2 = 0$ and $\min_x L = \mu^T b_1 + \lambda^T b_2$.
The dual problem is

$$
\begin{aligned}
\max_{\mu, \lambda} \ & \mu^T b_1 + \lambda^T b_2 \\
\text{s.t. } & A_1^T \mu + A_2^T \lambda = c \\
& \mu \geq 0
\end{aligned}
$$

Principles for LP duality (Lecture 3)

Primal LP		**Dual LP**	
Objective: min **Objective coefficient: c^T** **Constraint coefficient: (A,b)**		**Objective: max** **Objective coefficient: b^T** **Constraint coefficient: (A^T,c)**	
Vars:	***n*-th variable** **≥0** **≤0** **free**	**Cons:**	***n*-th constraint** **≤** **≥** **=**
Cons:	***m*-th constraint** **≤** **≥** **=**	**Vars:**	***m*-th variable** **≤0** **≥0** **free**

Principles for LP duality (Lecture 3)

Primal problem

$$\min_{x} \; c^T x$$
$$\text{s.t.} \; A_1 x \geq b_1$$
$$A_2 x = b_2$$

Dual problem

$$\max_{\mu,\lambda} \; \mu^T b_1 + \lambda^T b_2$$
$$\text{s.t.} \; A_1^T \mu + A_2^T \lambda = c$$
$$\mu \geq 0$$

- The dual problem derived from Lagrangian provides a third interpretation of LP duality
- The dual theory for LP is a special case
- For LP, using the principles in the TABLE may be more convenient

Dual Optimization

The dual problem is always a convex optimization, i.e. $-Q(\lambda, \mu)$ is a convex function.

$$-Q(\lambda, \mu) = -\min_x L(x, \mu, \lambda) = \max_x -L(x, \mu, \lambda)$$

Notice that $-L(x, \mu, \lambda)$ is a linear function (also convex) of μ, λ.
Moreover,
1. $\max\{c_1(x), ..., c_m(x)\}$ is a convex function if all $c_i, \forall i = 1, ..., m$ are convex.
2. if a function $c(x, y)$ is convex about y, then $\max_x c(x, y)$ is also a convex function.
Therefore, $-Q(\lambda, \mu)$ is a convex function.

Example in Lecture 4 - Minimization

$\mathcal{C}$ is a convex set, $f(x,y)$ is convex in the (x,y) space, then the function

$$g(x) = \min_{y \in \mathcal{C}} f(x,y)$$

is convex in x

Proof: For any x_1 and x_2, the minimizers are y_1 and y_2

$$g(x_1) = f(x_1, y_1),\ g(x_2) = f(x_2, y_2)$$

$$\begin{aligned} g(tx_1 + (1-t)x_2) &= \min_{y \in \mathcal{C}} f(tx_1 + (1-t)x_2, y) \\ &\leq f(tx_1 + (1-t)x_2, ty_1 + (1-t)y_2) \\ &\leq tf(x_1, y_1) + (1-t)f(x_2, y_2) \\ &= tg(x_1) + (1-t)g(x_2) \end{aligned}$$

Example

Construct the dual optimization of:

$$\begin{aligned} \min_{x_1,x_2} \ & x_1^2 + x_2^2 \\ \text{s.t. } & x_1 + x_2 \leq 2 \\ & x_1 - x_2 = 0 \end{aligned}$$

The Lagrangian is

$$\begin{aligned} L(x,\mu,\lambda) = & \ x_1^2 + x_2^2 + \mu(x_1 + x_2 - 2) + \lambda(x_1 - x_2) \\ = & \ x_1^2 + (\mu+\lambda)x_1 + x_2^2 + (\mu-\lambda)x_2 - 2\mu \\ = & \ (x_1 + \frac{\mu+\lambda}{2})^2 + (x_2 + \frac{\mu-\lambda}{2})^2 - \frac{(\mu+\lambda)^2}{4} - \frac{(\mu-\lambda)^2}{4} - 2\mu \end{aligned}$$

Let $x_1 = -\frac{\mu+\lambda}{2}$ and $x_2 = -\frac{\mu-\lambda}{2}$, then

$$\min_x L(x,\mu,\lambda) = -\frac{1}{2}(\mu^2 + \lambda^2) - 2\mu$$

Example

Therefore, the dual optimization is

$$\max_{\mu,\lambda} \; -\frac{1}{2}(\mu^2 + \lambda^2) - 2\mu$$
$$\text{s.t. } \mu \geq 0$$

which is equivalent to

$$\min_{\mu,\lambda} \; \frac{1}{2}(\mu^2 + \lambda^2) + 2\mu$$
$$\text{s.t. } \mu \leq 0$$

$\frac{1}{2}(\mu^2 + \lambda^2) + 2\mu$ and $-\mu$ are all convex.

Example

Construct the dual optimization of:

$$\min_{x_1,x_2} \; x_1^2 + x_2^2$$
$$\text{s.t. } x_1 + x_2 \geq 4$$
$$x_1 \geq 0, x_2 \geq 0$$

The Lagrangian is

$$\begin{aligned} L(x,\mu) = &\; x_1^2 + x_2^2 - \mu_1(x_1 + x_2 - 4) - \mu_2 x_1 - \mu_3 x_2 \\ = &\; x_1^2 - (\mu_1 + \mu_2)x_1 + x_2^2 - (\mu_1 + \mu_3)x_2 + 4\mu_1 \\ = &\; (x_1 - \frac{\mu_1 + \mu_2}{2})^2 + (x_2 - \frac{\mu_1 + \mu_3}{2})^2 - \frac{(\mu_1 + \mu_2)^2}{4} - \frac{(\mu_1 + \mu_3)^2}{4} + 4\mu_1 \end{aligned}$$

Therefore,

$$\min_{x_1,x_2} L(x,\mu) = -\frac{(\mu_1 + \mu_2)^2}{4} - \frac{(\mu_1 + \mu_3)^2}{4} + 4\mu_1$$

The dual optimization problem is:

$$\max_{\mu_1,\mu_2,\mu_3} \; -\frac{(\mu_1 + \mu_2)^2}{4} - \frac{(\mu_1 + \mu_3)^2}{4} + 4\mu_1$$
$$\text{s.t. } \mu_1 \geq 0, \mu_2 \geq 0, \mu_3 \geq 0$$

Dual Theorems

Weak duality: suppose the optimal value of the primal problem is f^* and the optimal value of the dual problem is Q^*, then $f^* \geq Q^*$**.**

Proof: Suppose the optimal solution of the dual problem is (λ^*, μ^*)
The optimal solution of the primal problem is x^*.

$$
\begin{aligned}
Q^* = \ & \min_x L(x, \lambda^*, \mu^*) \\
= \ & \min_x f(x) + \sum_{i=1}^{m} \lambda_i^* h_i(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \\
\leq \ & \min_{x \in \mathcal{X}} f(x) + \sum_{i=1}^{m} \lambda_i^* h_i(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \\
\leq \ & \min_{x \in \mathcal{X}} f(x) = f^*
\end{aligned}
$$

where $\mathcal{X} = \{x | h_i(x) = 0, \forall i = 1, ..., m; g_j(x) \leq 0, j = 1, ..., r\}$.

Example

Consider the optimization problem:

$$\min_{x_1,x_2} \; x_1x_2$$
$$\text{s.t. } x_1^2 + x_2^2 \leq 2$$

We know that $(x_1 + x_2)^2 \geq 0$, so

$$x_1x_2 \geq -\frac{x_1^2 + x_2^2}{2} \geq -1$$

Equality holds when $x^* = (1, -1)$ or $x^* = (-1, 1)$.

$\mu \geq 0$

The Lagrangian is

$$L(x, \mu) = x_1x_2 + \mu(x_1^2 + x_2^2 - 2)$$

1. If $\mu = 0$, $\min_x L(x, \mu) \to -\infty$.

Example

2. If $\mu \neq 0$, take the derivative of $L(x, \mu)$:

$$\frac{\partial L}{\partial x_1} = x_2 + 2\mu x_1, \ \frac{\partial L}{\partial x_2} = x_1 + 2\mu x_2$$

$$H(x) = \begin{bmatrix} 2\mu & 1 \\ 1 & 2\mu \end{bmatrix}$$

$H(x)$ is positive semidefinite iff $4\mu^2 - 1 \geq 0$, i.e. $\mu \geq \frac{1}{2}$.
1) if $\mu = \frac{1}{2}$, let $\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = 0$, we have $x_1 = -x_2$.

$$L(x, \mu) = -x_1^2 + \frac{1}{2}(2x_1^2 - 2) = -1$$

2) If $\mu > \frac{1}{2}$, let $\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = 0$, we have $x_1 = x_2 = 0$. Therefore, $Q(\mu) = -2\mu$.

$$\max_{\mu} Q(\mu) = -1.$$

Dual Theorems

Recall that we have $f^* \geq Q^*$ (weak duality). In some problems, we have $f^* = Q^*$, which is called ***strong duality***.

Slater's condition: if the primal is a convex problem (i.e. f and $g_j, \forall j$ are convex, $h_i, \forall i$ are affine), and there exists at least one strictly feasible x:

$$g_j(x) < 0, \forall j = 1, \dots, r; \; h_i(x) = 0, \forall i = 1, \dots, m$$

Then strong duality holds.

An important refinement: strict inequalities only need to hold over functions $g_j(.)$ that are not affine.

Comparison

Weak duality: suppose the optimal value of the primal problem is f^* and the optimal value of the dual problem is Q^*, then $f^* \geq Q^*$.

Weak duality: Let x_0, λ_0 be a feasible solution of the primal problem and the dual problem, respectively. We have $c^T x_0 \geq b^T \lambda_0$.

Similar

Strong duality:

- Refined version of Slater's condition indicates that strong duality holds for an LP if it is feasible.
- Similarly, strong duality holds for the dual LP if it is feasible
- Moreover, the dual of dual LP is the primal LP
- We nearly always have strong duality for LPs.

Example

Complementary and slackness condition

$0 \leq (2 - x_1 - x_2) \perp \mu \geq 0$

In the previous example, we already have

Primal problem

$$\min_{x_1,x_2} \; x_1^2 + x_2^2$$
$$\text{s.t. } x_1 + x_2 \leq 2$$
$$x_1 - x_2 = 0$$

Dual problem

$$\max_{\mu,\lambda} \; -\frac{1}{2}(\mu^2 + \lambda^2) - 2\mu$$
$$\text{s.t. } \mu \geq 0$$

It is easy to check that the primal problem is a convex optimization, and there is at least one strictly feasible point, e.g. $x = (0.5,0.5)$. Therefore, we can apply Slater's condition and should have $f^* = Q^*$.

The optimal solution of the primal problem is $x^* = (0,0)$ with $f^* = 0$.
The optimal solution of the dual problem is $\mu^* = 0, \lambda^* = 0$ with $Q^* = 0$.

Thanks！