西安交通大學

XI'AN JIAOTONG UNIVERSITY

The theory of dual problems and it's application

Major: Applied Mathematics

Name: Zhe Wang

Student ID: for fun

1 Theory

For an optimization problem:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, p, \end{aligned} \tag{1}$$

No assumption about the convexity of the problem, a common strategy is Lagrange Method, which come up with an approximation to the primal problem.

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \tag{2}$$

Why we can do this? In this article, I hope I can explain it clearly for you.

Remark 1.1. *Generally speaking, we are trying to figure out this problem in* $\mathbb{R}^N$. *What's more, according to Riesz representation theorem, for a Hilbert space, any bounded linear functional in it's dual space can be represented by inner product of the primal space. Thus, the Lagrangian multiplier here is actually the bounded linear functional in the dual space. To explore the mathematical background thoroughly, we set our optimization problem in Banach space.*

1.1 the disturbance for the primal problem

If you are trying to figure out an optimization problem by it's dual problem, it is better to give a disturbance to the primal one. Thus, different kind of disturbance will lead to different dual problems.

Definition 1.1. *We say* $\Phi(x, y)$ *is a disturbance to the primal problem, if it is convex and satisfies:*

$$\begin{aligned} \Phi : X \times Y \to R \\ \Phi(x, 0) = f(x) \end{aligned} \tag{3}$$

Then, our objective function comes clearly

$$\min_{x} \quad \Phi(x, 0) \tag{4}$$

We hope $\Phi(x, y)$ is convex, then the generalization version of the new problem can be formed as

$$\min_{x} \quad \Phi(x, y) \tag{5}$$

for $y \in Y$

Definition 1.2. *Let* $h(y) = \inf_{x} \Phi(x, y)$*, that is to say* $h(y)$ *is the point-wise lower bound for* $\Phi(x, y)$

Untill now, our original problem turned into the figuring out of $h(0)$

Lemma 1.1. *For any* $x \in X, \;\; h(x) \geq h^{**}(x)$

The proof is trivial, no details here.

Remark 1.2. *The condition of the equality is easy to be satisfied, as long as the subgradient of* h *exist at 0.*

Remark 1.3. *As a matter of fact, the existence of subgradient for a convex function is not a harsh demanding, as long as the function is closed and convex, with the existence of subgradient on the boundary.*

Based on Remark 1.2, we have $h(0) \geq h^{**}(0)$, the lower bound for the original problem is given out. It is a natural idea to estimate $h(0)$ with $h^{**}(0)$, and it is what we do in reality.

Lemma 1.2. *According to the definition of* $h^{**}(0)$

$$h^{**}(0) = \max_{y^* \in Y^*} \; -h^*(y^*) \tag{6}$$

This problem can be rewritten in Φ, which leads to our second lemma

Lemma 1.3. *For any* $y^* \in Y^*, \; h^*(y^*) = \Phi^*(0, y^*)$

Proof.

$$\begin{aligned} h^*(y^*) &= \sup_{y \in Y}(< y^*, y > -h(y)) \\ &= \sup_{y \in Y}(< y^*, y > - \inf_{x \in X} \Phi(x, y)) \\ &= \sup_{y \in Y} \sup_{x \in X} \{< 0, x > + < y^*, y > -\Phi(x, y)\} \\ &= \sup_{z \in X \times Y} \{< z^*, z > -\Phi(z)\} \\ &= \Phi^*(z^*) \end{aligned} \tag{7}$$

in which $z = (x, y) \in X \times Y, \; z^* = (0, y^*) \in (X \times Y)$ □

Let's make the whole process much more clear

- The original problem turned into the estimation of $h(0)$
- The estimation of $h(0)$ can be achieved via the calculation of $h^{**}(0)$
- The calculation of $h^{**}(0)$ can be turned into the figuring out of $\max\limits_{y^* \in Y^*} -h^*(y^*)$
- The last step is based on the equivalence of $h^*(y^*)$ and $\Phi^*(0, y^*)$

Now, we can give out the primal and dual problem

primal problem

$$\min_{x \in X} \Phi(x, 0) \tag{8}$$

dual problem

$$\min_{y^* \in Y^*} \Phi^*(0, y^*) \tag{9}$$

The primal problem is about x, Φ and domain , the dual problem is about y^*, Φ^* and range, it is such a wonderful and elegant transformation.

2 One example

Based on the foregoing structure, consider the following problem

Example 2.1.

$$\begin{aligned} &\min_x f(x) \\ &\textit{subject to} \quad g(x) \leq 0 \end{aligned} \tag{10}$$

Here comes the disturbance

$$\begin{aligned} &\min_x f(x) \\ &\text{subject to} \quad g(x) + y \leq 0 \end{aligned} \tag{11}$$

$\Phi(x, y)$ can be formed as :

$$\Phi(x, y) = \begin{cases} f(x) & \text{if } g(x) + y \leq 0 \\ \infty & \text{otherwise} \end{cases} \tag{12}$$

To get the dual problem, it is necessary to calculate

$$-\Phi^*(0, y^*) \tag{13}$$

The calculation is explicit

$$
\begin{aligned}
-\Phi^*(0,y^*) &= -\sup_{g(x)+y\leq 0}\{< y^*, y > -f(x)\} \\
&= -\sup_{x\in X, q\geq 0}\{< y^*, -g(x)-q > -f(x)\} \\
&= \inf_{x\in X, q\geq 0} f(x)+ < y^*, g(x) > + < y^*, q >
\end{aligned} \tag{14}
$$

Minimize it w.r.t q, we will get $-\infty$ unless for all $q \geq 0$, $< y^*, q >\geq 0$, that is to say $y^* \geq 0$

Our primal problem is transformed into

$$
-\Phi^*(0,y^*) = \begin{cases} \inf_x f(x)+ < y^*, g(x) > & \text{if } y^* \geq 0 \\ -\infty & \text{otherwise} \end{cases} \tag{15}
$$

In Hilbert space, we can change the inner product into a linear funcional, since $< y^*, g(x) >= \lambda g(x)$, then maximize that w.r.t λ:

$$
\max_{\lambda\geq 0}\inf_{x\in X} f(x)+ < \lambda, g(x) > \tag{16}
$$

It is not hard to verify that the original problem is equivalent to

$$
\min_{x\in X}\sup_{\lambda\geq 0} f(x)+ < \lambda, g(x) > \tag{17}
$$

(x, λ) is actually the saddle point for the primal-dual problem.