Convex Optimization

Chapter 5: Duality

Notes by Renxiong Liu

In this notes, we mainly discuss Lagrangian duality that are widely used for constrained optimization problem.

5.1 Lagrangian Duality

Firstly, we introduce some notions that will be frequently used in this note. For optimization problem (note: not necessary convex),

$$\begin{array}{lll}(\Pi_P)\ p^\star = & \text{minimize} & f_0(x) \\ & \text{subject to} & f_i(x) \leq 0,\ i = 1, 2 \ldots, m \leftarrow \lambda_i \\ & & h_j(x) = 0,\ j = 1, 2 \ldots, n \leftarrow v_i\end{array}$$

We call f_0 the primal function, $x \in \mathbb{R}^p$ the primal variables. Note the domain for this problem is

$$D = (\bigcap_{i=0}^{m} \mathbf{dom}(f_i)) \cap (\bigcap_{j=1}^{n} \mathbf{dom}(h_j))$$

For problems above, one very useful method is to consider the Lagrangian function,

$$L : \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$$

with domain $\mathbf{dim}(L) = D \times \mathbb{R}^m \times \mathbb{R}^n$:

$$L(x, \lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} v_j h_j(x)$$

where we call $\lambda = (\lambda_1, \ldots, \lambda_m)$ and $v = (v_1 \ldots, v_n)$ the dual variables.

To proceed, we define the dual function as

$$g(\lambda_1, \lambda_m, v_1, \ldots, v_n) = \inf_{x \in D} L(x, \lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n)$$

Note for any fixed $x \in D$, $L(x, \lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n)$ is an affine function with respect to $(\lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n)$, we have that $g(\lambda_1, \lambda_m, v_1, \ldots, v_n)$ must be convace with respect to $(\lambda_1, \lambda_m, v_1, \ldots, v_n)$.

5.2 Weak Duality

Given the dual function $g(\lambda_1, \lambda_m, v_1, \ldots, v_n)$ and corresponding dual variables $(\lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n)$, we define the following problem as the dual problem

$$\begin{array}{lll}(\Pi_D)\ d^\star = & \text{maximize} & g(\lambda_1, \lambda_m, v_1, \ldots, v_n) \\ & \text{subject to} & \lambda_i \geq 0,\ i = 1, 2 \ldots, m\end{array}$$

Now, we may wonder what's the relationship between primal problem and dual problem. Indeed, we can see that for any $x \in D$, $\lambda_i \geq 0$ and $v_j \in \mathbb{R}$, we have that

$$\begin{aligned} g(\lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n) &\leq L(x, \lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n) \\ &= f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{n} v_j h_j(x) \\ &\leq f_0(x) \end{aligned}$$

which means

$$g(\lambda_1, \ldots, \lambda_m, v_1, \ldots, v_n) \leq d^\star \leq p^\star \leq f_0(x)$$

Hence, once we find $\lambda^\star, v^\star$ and $x^\star$ such that $g(\lambda^\star, v^\star) = f_0(x^\star)$, then we will have that

$$\begin{aligned} x^\star &\in \arg\min_{x \in \mathcal{F}} f_0(x) \\ (\lambda^\star, v^\star) &\in \arg\max_{\lambda \geq 0} g(\lambda, v) \end{aligned}$$

To proceed, we first look at one example. For any l_p norm $\|\cdot\|_p$, we define the associated dual norm as

$$\|y\|_\star = \text{maximize}\, \langle y, x \rangle, \text{ subject to } \|x\|_p \leq 1 \tag{5.1}$$

Recall the Holder inequality:

$$\langle x, y \rangle \leq \|x\|_p \cdot \|y\|_q$$

where $p, q > 0$ and $\frac{1}{p} + \frac{1}{q} = 1$, we have that the dual norm of $\|\cdot\|_p$ is $\|\cdot\|_q$. Now, given any (convex) norm $\|\cdot\|$, we claim the conjugate function of $f(x) = \|x\|$ is

$$f^\star(y) = \begin{cases} 0, & \text{if } \|y\|_\star \leq 1 \\ \infty, & \text{otherwise} \end{cases}$$

This is because

$$f^\star(y) = \sup_x \langle y, x \rangle - \|x\| = \sup_x \|x\|(\langle y, \frac{x}{\|x\|} \rangle - 1) = \sup_x \|x\|(\|y\|_\star - 1) = \begin{cases} 0, & \text{if } \|y\|_\star \leq 1 \\ \infty, & \text{otherwise} \end{cases}$$