Duality

Duality Introductory Tutorial. Done with specific purpose thus many contents are omitted.

1 Lagrangian Dual Problem

We consider an optimization problem in a standard form

$$\begin{aligned} &\text{minimize } f_0(x) \\ &\text{subject to } f_i(x) \leq 0 \quad i = 1, 2, ..., m \\ &\qquad\qquad h_i(x) = 0 \quad i = 1, 2, ..., p \end{aligned} \tag{1}$$

where $x \in \mathbf{R}^n$ and the domain $\mathcal{D}$ is nonempty. We don't assume the problem is (1) is convex. We denote the optimal solution to (1) as p^*.

The ideal is augmenting the objective function with a weighted sum of constraints. We define the Lagrangian as $\mathcal{L}(x, \lambda, \nu) : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$.

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

where λ_i and ν_i is the Lagrangian multipliers (dual variables). We define the Lagrangian dual function $g(\lambda, \nu) : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$ as the minimum value of the Lagrangian function over x[1].

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \nu) \tag{3}$$

Since the dual function is the point-wise infimum of a family of affine functions[2] of (λ, ν), it is concave, even when (1) is not convex. The dual function yields the lower bound of the optimal value of (1). For $\lambda \succeq 0$ and ν, we have

$$g(\lambda, \nu) \leq p^* \tag{4}$$

The proof is simply and straightforward. For each feasible $\tilde{x}$, $\lambda \succeq 0$, and ν we have

$$\begin{aligned} \mathcal{L}(\tilde{x}, \lambda, \nu) &= f_0(\tilde{x}) + \sum_{i=1}^{m} \lambda_i f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x}) \\ &\leq f_0(\tilde{x}) \end{aligned} \tag{5}$$

[1]When the Lagrangian is unbounded, the minimum value can take $-\infty$. (inf and sup are similar to min and max, but are more useful in analysis because they better characterize special sets which may have no minimum or maximum. For instance, the positive real numbers $\mathbf{R}^+$ does not have a minimum. There is, however, exactly one infimum of the positive real numbers 0, which is smaller than all the positive real numbers and greater than any other real number which could be used as a lower bound.)

[2]An affine function is just a linear function plus a translation

When we minimize over all the feasible x, we have $g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \nu) \leq \min_{x} f_0(x) = p^*$.

Since we have a lower bound of the optimal value p^*, a natural question is: What is the best lower bound we can get? This leads to the following question.

$$\begin{aligned} &\text{maximize } g(\lambda, \nu) \\ &\text{subject to } \lambda \succeq 0 \end{aligned} \tag{6}$$

(6) is called the dual problem of (1). It is a convex optimization problem since the maximization problem and the constraints are all convex. We denote the optimal value as d^*. We often make the constraint $(\lambda, \nu) \in \textbf{dom } g$ where

$$\textbf{dom } g = \{(\lambda, \nu) \mid g(\lambda, \nu) > -\infty)\} \tag{7}$$

2 Weak and Strong Duality

By definition, d^* is the best lower bound of p^*. Naturally, we have

$$d^* \geq p^* \tag{8}$$

This property is called *weak duality*. The weak duality inequality holds when d^* and p^* are infinite. For example, if the primal problem is unbounded below (i.e. $p^* = -\infty$), we must have $d^* = -\infty$, that means the Lagrange dual problem is infeasible. Conversely, if the dual problem is unbounded above (i.e. $d^* = +\infty$), we must have $p^* = +\infty$, that means the primal problem is infeasible.

We refer the difference $d^* - p^*$ as the duality gap. Since the dual problem is convex, we can use the dual function to obtain the lower bound of the optimal value of some difficult problem.

If the duality gap equals to zero, we call the *strong duality* holds.

$$d^* = p^* \tag{9}$$

Strong duality usually does not hold. But if the primal problem in (1) is convex[3], we usually have the strong duality. There are many results that establish conditions on the problem, beyond convexity, under which strong duality holds. These conditions are called $constraint\ qualifications$[4]. One simple constraint qualification is $Slater's\ condition$: if $\exists x \in \textbf{relint } \mathcal{D}$ such that

$$f_i(x) > 0 \ for\ i = 0, 1, 2, ..., m \tag{10}$$

Linear inequalities do not need to hold with strict inequality.

[3]This mean $f_i(x)$ for $i = 0, 1, 2, ..., m$ are all convex

[4]There exists many different constraint qualifications

2.1 Max-min Characterization of Weak and Strong Duality

For simplification, we discuss the problems with no equality constraints in this section. To start with, we first note

$$\begin{aligned}\sup_{\lambda \succeq 0} \mathcal{L}(x, \lambda) &= \sup_{\lambda \succeq 0} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \right) \\ &= \begin{cases} f_0(x) & f_i(x) \leq 0 \ , \ i = 1, 2, ..., m \\ \infty & otherwise \end{cases}\end{aligned} \tag{11}$$

Thus, we can express p^* as

$$p^* = \inf_x \sup_{\lambda \succeq 0} \mathcal{L}(x, \lambda) \tag{12}$$

By definition, d^* can be expressed as

$$d^* = \sup_{\lambda \succeq 0} \inf_x \mathcal{L}(x, \lambda) \tag{13}$$

For the weak duality, we have

$$\sup_{\lambda \succeq 0} \inf_x \mathcal{L}(x, \lambda) \geq \inf_x \sup_{\lambda \succeq 0} \mathcal{L}(x, \lambda) \tag{14}$$

For the strong duality, we have

$$\sup_{\lambda \succeq 0} \inf_x \mathcal{L}(x, \lambda) = \inf_x \sup_{\lambda \succeq 0} \mathcal{L}(x, \lambda) \tag{15}$$

The inequality in (14) holds in general. We called it $Max - Min\ Inequality$. When the equality holds, we say the function satisfies the *saddle point* property.

3 Optimality Conditions

If we assume the strong duality holds and x^* is the primal optimal and (λ^*, ν^*) is the dual optimal, we have

$$\begin{aligned}f_0(x^*) = g(\lambda^*, \nu^*) &= \inf_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \right) \\ &\leq f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{p} \nu_i^* h_i(x^*) \\ &\leq f_0(x^*)\end{aligned} \tag{16}$$

Hence, the two inequalities must hold with qualities. Then, we can conclude that $\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0$. Since each term in the summation is non-positive, we have

$$\lambda_i^* f_i(x^*) = 0 \ for \ i = 1, 2, ..., m \tag{17}$$

This condition is called *Complementary Slackness*. It is equivalent to the following

$$\begin{aligned} \lambda^* > 0 \Rightarrow f_i(x^*) = 0 \\ f_i(x^*) < 0 \Rightarrow \lambda^* = 0 \end{aligned} \tag{18}$$

This means the i-th optimal Lagrange multiplier is zero unless the i-th constraint is active at the optimum.

3.1 KKT

For problems with differentiable f_i and h_i, the following conditions are called *KKT Conditions* (We still do not assume convexity). We assume x^*, λ^*, and ν^* is the optimal primal point and dual points with zero duality gap.

- Primal Constraints: $f_i(x) \leq 0 \ \ i = 1, 2, ..., m$, $h_i(x) = 0 \ \ i = 1, 2, ..., p$
- Dual Constraints: $\lambda \succeq 0$
- Complementary Slackness: $\lambda_i^* f_i(x^*) = 0 \ for \ i = 1, 2, ..., m$
- Gradient Vanishes: $\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0$

For any optimization problem with differentiable objective and constraint functions for which strong duality obtains, any pair of primal and dual optimal points must satisfy the KKT conditions. For the KKT condition, we have

- If x^* and λ^*, ν^* are primal and dual solutions, with zero duality gap, then x^*, λ^*, and ν^* satisfy the KKT conditions
- If x^* and λ^*, ν^* satisfy the KKT conditions, then x^* and λ^*, ν^* are primal and dual solutions

4 Reference

[1] https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf

[2] https://web.stanford.edu/~boyd/cvxbook/bv_cvxslides.pdf

[3] https://www.cs.cmu.edu/~ggordon/10725-F12/slides/16-kkt.pdf