

Duality. Strong Duality.

Seminar

Optimization for ML. Faculty of Computer Science. HSE University

Dual function

The **general mathematical programming problem** with functional constraints:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned}$$

And the Lagrangian, associated with this problem:

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

We assume $\mathcal{D} = \bigcap_{i=1}^m \text{dom } f_i \cap \bigcap_{i=1}^p \text{dom } h_i$ is nonempty. We define the Lagrange **dual function** (or just dual function) $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ as the minimum value of the Lagrangian over x : for $\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}^p$

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right)$$

Dual function. Summary

💡 Primal

Function:

$$f_0(x)$$

Variables:

$$x \in S \subseteq \mathbb{R}^n$$

Constraints:

$$f_i(x) \leq 0, i = 1, \dots, m$$

$$h_i(x) = 0, i = 1, \dots, p$$

💡 Dual

Function:

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

Variables

$$\lambda \in \mathbb{R}_+^m, \nu \in \mathbb{R}^p$$

Constraints:

$$\lambda_i \geq 0, \forall i \in \overline{1, m}$$

Strong Duality

It is common to name this relation between optimals of primal and dual problems as **weak duality**. For problem, we have:

$$d^* \leq p^*$$

While the difference between them is often called **duality gap**:

$$0 \leq p^* - d^*$$

Strong duality happens if duality gap is zero:

$$p^* = d^*$$

Slater's condition

If for a convex optimization problem (i.e., assuming minimization, f_0, f_i are convex and h_i are affine), there exists a point x such that $h(x) = 0$ and $f_i(x) < 0$ (existence of a **strictly feasible point**), then we have a zero duality gap and KKT conditions become necessary and sufficient.

Reminder of KKT statements

Suppose we have a **general optimization problem**

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \tag{1}$$

and **convex optimization problem**, where all equality constraints are affine:

$$h_i(x) = a_i^T x - b_i, i \in 1, \dots, p.$$

The **KKT system** is:

$$\begin{aligned} \nabla_x L(x^*, \lambda^*, \nu^*) &= 0 \\ \nabla_\nu L(x^*, \lambda^*, \nu^*) &= 0 \\ \lambda_i^* &\geq 0, i = 1, \dots, m \\ \lambda_i^* f_i(x^*) &= 0, i = 1, \dots, m \\ f_i(x^*) &\leq 0, i = 1, \dots, m \end{aligned} \tag{2}$$

KKT becomes necessary

If x^* is a solution of the original problem Equation 1, then if any of the following regularity conditions is satisfied:

- **Strong duality** If $f_1, \dots, f_m, h_1, \dots, h_p$ are differentiable functions and we have a problem Equation 1 with zero duality gap, then Equation 2 are necessary (i.e. any optimal set x^*, λ^*, ν^* should satisfy Equation 2)
- **LCQ** (Linearity constraint qualification). If $f_1, \dots, f_m, h_1, \dots, h_p$ are affine functions, then no other condition is needed.
- **LICQ** (Linear independence constraint qualification). The gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at x^*
- **SC** (Slater's condition) For a convex optimization problem (i.e., assuming minimization, f_i are convex and h_j is affine), there exists a point x such that $h_j(x) = 0$ and $g_i(x) < 0$.

Then it should satisfy Equation 2

KKT in convex case

If a convex optimization problem with differentiable objective and constraint functions satisfies Slater's condition, then the KKT conditions provide necessary and sufficient conditions for optimality: Slater's condition implies that the optimal duality gap is zero and the dual optimum is attained, so x^* is optimal if and only if there are (λ^*, ν^*) that, together with x^* , satisfy the KKT conditions.

Problem 1. Dual LP

Ensure, that the following standard form *Linear Programming* (LP):

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^\top x \\ \text{s.t.} \quad & Ax = b \\ & x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

Has the following dual:

$$\begin{aligned} \max_{y \in \mathbb{R}^n} \quad & b^\top y \\ \text{s.t.} \quad & A^\top y \preceq c \end{aligned}$$

Find the dual problem to the problem above (it should be the original LP).

Problem 2. Projection onto probability simplex

Find the Euclidean projection of $x \in \mathbb{R}^n$ onto probability simplex

$$\mathcal{P} = \{z \in \mathbb{R}^n \mid z \succeq 0, \mathbf{1}^\top z = 1\},$$

i.e. solve the following problem:

$$\begin{aligned} \frac{1}{2} \|y - x\|_2^2 &\rightarrow \min_{y \in \mathbb{R}^n \succeq 0} \\ \text{s.t. } \mathbf{1}^\top y &= 1 \end{aligned}$$

Problem 3. Shadow prices or tax interpretation

Consider an enterprise where x represents its operational strategy and $f_0(x)$ is the operating cost. Therefore, $-f_0(x)$ denotes the profit in dollars. Each constraint $f_i(x) \leq 0$ signifies a resource or regulatory limit. The goal is to maximize profit while adhering to these limits, which is equivalent to solving:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \end{aligned}$$

The optimal profit here is $-p^*$.

Problem 4. Norm regularized problems

Ensure, that the following normed regularized problem:

$$\min f(x) + \|Ax\|$$

has the following dual:

$$\begin{aligned} f^*(-A^\top y) &\rightarrow \min_y \\ \text{s.t. } \|y\|_* &\leq 1 \end{aligned}$$