Chapter 4: Optimization for Data Science
Duality

TANN Chantara

Department of Applied Mathematics and Statistics
Institute of Technology of Cambodia

October 8, 2022

Table of Contents

1. The Lagrange Duel Function
2. The Dual Problem
3. Dual Cones
4. Generalized Inequalities
5. Summary

Lagrangian Duality

$$P : \begin{cases} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \forall i = 1, ..., m \\ & h_i(x) = 0, \forall i = 1, ..., p \end{cases}$$

- $x \in \mathbb{R}$ decision variable
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ objective function
- $f_i : \mathbb{R}^n \to \mathbb{R}$ inequality constraint functions
- $h_i : \mathbb{R}^n \to \mathbb{R}$ equality constraint functions

Assume for simplicity that function values cannot be ∞

Lagrangian

Definition: The Lagrangian L of the problem P is defined as the function $L : \mathbb{R}^n \times \mathbb{R}^n_+ \times \mathbb{R}^p \to \mathbb{R}$

$$L(x, \lambda, y) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

- Weighted sum of objective and constraint functions.
- λ_i is the Lagrange multiplier corresponding to $f_i(x) \leq 0$.
- μ_i is the Lagrange multiplier corresponding to $h_i(x) = 0$.

The Lagrangian is concave (affine) in (λ, μ) for any fixed x.
If P is convex optimization problem, the Lagrangian is convex in x for any fixed (λ, μ), i.e., L is saddle function.

Lagrangian Cont'd

The Lagrangian allows us to reexpress the optimization problem P as a min-max problem. Indeed, define

$$f(x) = \sup_{\lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p}$$

and note that

$$f(x) = \begin{cases} f_0(x), & \text{if } f_i(x) \leq 0 \forall i \text{ and } h_i(x) = 0 \forall i \\ \infty & \text{otherwise.} \end{cases}$$

Thus, we obtain

Lagrangian Reformulation:

$$\inf P = \inf_{x \in \mathbb{R}^n} f(x) = \inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p} L(x, \lambda, \mu)$$

The Dual Problem

Below we refer to P as the primal problem. Using the Lagrangian we can introduced a dual problem with objective function.

$$g(\lambda, \mu) = \infty_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$$

Dual Problem:

$$D : \sup_{\lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p} g(\lambda, \mu)$$

By construction, D is equivalent to a min-max problem.

$$\sup D = \sup_{\lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p} g(\lambda, \mu) = \sup_{\lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p} \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$$

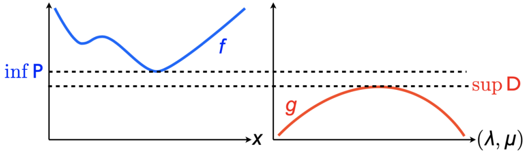
Weak Duality

Proposition: $g(\lambda, \mu) \geq f(x), \forall x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p$
Proof: By definition we have

$$g(\lambda, \mu) = \inf_{\bar{x} \in \mathbb{R}^n} L(\bar{x}, \lambda, \mu) \leq L(x, \lambda, \mu) \leq \sup_{\bar{\lambda}, \bar{\mu}} L(x, \bar{\lambda}, \bar{\mu}) = f(x)$$

Corollary (Weak Duality):

$$\sup D \leq \inf P$$



Thus if D is unbounded ($\sup D = \infty$), then P must be infeasible. If P is unbounded ($\inf P = -\infty$), then D must be infeasible.

Significance of Dual Solutions

Note: Every feasible solution of P (D) provides an upper (lower) bound on both $\inf \mathsf{P}$ and $\sup \mathsf{D}$.

Assume $\hat{x}$ is a (feasible) candidate solution for P. Its quality is quantified by $f_0(\hat{x}) - \inf \mathsf{P}$. However, $\inf \mathsf{P}$ is unknown!

A dual feasible solution (λ, μ) provides a proof or certificate that

$$\inf \mathsf{P} \geq g(\lambda, \mu) \Rightarrow f_0(\hat{x}) - \inf \mathsf{P} \leq f_0(\hat{x}) - g(\lambda, \mu)$$

Thus, if $f_0(\hat{x}) - g(\lambda, \mu) \leq \varepsilon$, then $\hat{x}$ is an ε-optimal solution.

Strong Duality

Definition: $\Delta = \inf P - \sup P$ is called the duality gap. By weak duality, we have that $\Delta \geq 0$. If $\Delta = 0$, we say that strong duality holds.

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

Strong Duality

- does not hold in general.
- always holds if P is a convex problem satisfying a constraint qualification.

Slater's constraint qualification holds if there exist x_S with

- $f_i(x_S) < 0, \forall i = 1, \dots, m$
- $h_i(x_S) = 0, \forall i = 1, \dots, p$

Least squares problem

Primal Problem

$$\begin{cases} \min_{x \in \mathbb{R}^n} & x^T x \\ \text{s.t.} & Ax = b \end{cases}$$

Lagrangian

$$L(x, \mu) = x^T x + \mu^T (Ax - b)$$

Dual Objective

$$\begin{aligned} &(\mu) = \min_{x \in \mathbb{R}^n} L(x, \mu) \\ &\Rightarrow \nabla_x L(x, \mu) = 2x + A^T \mu = 0 \\ &\Rightarrow x = -\frac{1}{2} A^T \mu \\ &\Rightarrow g(\mu) = -\frac{1}{4} \mu^T A A^T \mu - b^T \mu \end{aligned}$$

Dual Problem

$$\max_{\mu \in \mathbb{R}^m} -\frac{1}{4} \mu^T A A^T \mu - b^T \mu$$

Standard form Linear Program

Primal problem
$$\begin{cases} \min_{x \geq 0} & c^T x \\ \text{s.t.} & Ax = b \end{cases}$$

Lagrangian
$$L(x, \lambda, \mu) = c^T x - \lambda^T x + \mu^T (Ax - b)$$

Dual objective
$$g(\lambda, \mu) = \min_x L(x, \lambda, \mu)$$

$$= \begin{cases} -b^T \mu, & c - \lambda + A^T \mu = 0 \\ -\infty & \text{else} \end{cases}$$

Dual Problem
$$\begin{cases} \max_{\mu, \lambda \geq 0} & -b^T \mu \\ \text{s.t.} & c - \lambda + A^T \mu = 0 \end{cases}$$

$$\iff \begin{cases} \max_{\mu} & b^T \mu \\ \text{s.t.} & A^T \mu \leq c \end{cases}$$

Quadratic Program

Primal Problem $(P \succ 0)$
$$\begin{cases} \min_x & x^T P x \\ \text{s.t.} & Ax \le b \end{cases}$$

Lagrangian
$$L(x, \lambda) = x^T P x + \lambda^T (Ax - b)$$

Dual Objective
$$\begin{aligned} g(\lambda, \mu) &= \min_x L(x, \lambda, \mu) \\ &= -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \end{aligned}$$

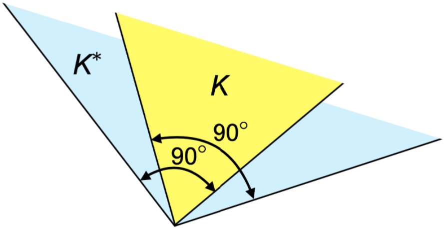
Dual Problem
$$\max_{\lambda \ge 0} -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda$$

Dual Cones

Definition: If K is a cone, then the set

$$K^* = \{y \in \mathbb{R}^n : x^T y \geq 0, \forall x \in K\}$$

is the dual cone of K.

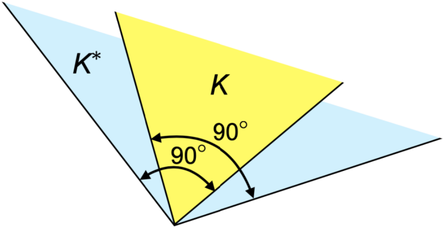


Note that K^* is a cone by construction.
Example: $(\mathbb{S}^n_+) = \{Y \in \mathbb{S}^n : \mathrm{tr}(X^T Y) \geq 0, \forall X \in \mathbb{S}^n_+\}$

Dual Cones Cont'd

Properties of the dual cone:

- K^* is closed and convex.
- $K_2 \subset K_1 \implies K_1^* \subseteq K_2^*$ (the smaller K, the larger K^*).
- $K^{**} = \mathrm{cl}(\mathrm{conv}(K))$, the smallest convex closed superset of K.
- If a convex cone K is proper, then K^* is proper and $K^{**} = K$



Definition: A cone K is called self-dual, if $K^* = K$.
Example of self-dual cones: $\mathbb{R}^n_+$, the second-order cone $\mathbb{S}^n_+$.

Problems with Generalized Inequality

$$P : \begin{cases} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) \preceq_{K_i} 0, \forall i = 1, \dots, m \\ & h_i(x) = 0, \forall i = 1, \dots, p \end{cases}$$

Where $K_i \subset \mathbb{R}^{r_i}$ is a proper convex cone.

- Assign a Lagrange multiplier $\lambda_i \in K_i^*$ to $f_i(x) \preceq_{K_i} 0$.
- Assign a Lagrange multiplier $\mu_i \in \mathbb{R}$ to $h_i(x) = 0$.

The Lagrangian $L : \mathbb{R}^n \times K_1^* \times K_m^* \times \mathbb{R}^p \to \mathbb{R}$ is defined as

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i^T f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

Problems with Generalized Inequality Cont'd

$$P : \inf_{x \in \mathbb{R}^n} \qquad f(x) = \sup_{\lambda \in K^*, \mu \in \mathbb{R}^p} L(x, \lambda, \mu)$$

$$D : \sup_{\lambda \in K^*, \mu \in \mathbb{R}^p} g(\lambda, \mu) \qquad g(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$$

- Weal duality holds because
  - $f_i(x) \in K_i \Rightarrow \forall \lambda_i \in K_i^* : \lambda_i^T f_i(x) \leq 0$
  - $f_i(x) \notin K_i \Rightarrow \forall \lambda_i \in K_i^* : \lambda_i^T f_i(x) > 0$
- Strong duality holds for covex problems satisfy a constraint qualification.
- Slater's constraint qualification holds if there exist x_S with
  - $f_i(x_S) \prec 0, \forall i = 1, ..., m$
  - $h_i(x_S) = 0, \forall i = 1, ..., p$

SOCP Duality

Primal Program:

$$P : \begin{cases} \min_{x \in \mathbb{R}^n} & f^T x \\ \text{s.t.} & ||A_i x - b_i||_2 \leq c_i^T x + d_i, i = 1, ..., m \end{cases}$$

Dual Program:

$$D : \begin{cases} \max_{V_i, \mu \in \mathbb{R}^m} & -\sum_{i=1}^m (b_i^T v_i + d_i \mu_i) \\ \text{s.t.} & \sum_{i=1}^m (A_i^T v_i + c_i \mu_i) = f \\ \text{s.t.} & ||V_i|| \leq \mu_i, i = 1, ..., m \end{cases}$$

Derivation: Keep as Homework.

SPD Duality

Primal Program:

$$\begin{cases} \min_{x\in\mathbb{R}^n} & c^T x \\ \text{s.t.} & F_1x_1 + \cdots + F_nx_x \preceq G \end{cases}$$

Dual Program:

$$\begin{cases} \min_{\wedge\succeq 0} & -\mathrm{tr}(\wedge^T G) \\ \text{s.t.} & \mathrm{tr}(\wedge^T F_i) = -c_i, \forall i = 1, ..., n \end{cases}$$

Derivation: Homework to proof.

Summary

- **Lagrangian:** weighted sum of objective and constraints; saddle function for convex problems; used to construct primal objective (partial maximum w.r.t. Lagrange multipliers) and dual objective (partial minimum w.r.t. decision variables).
- **Duality:** $\sup \mathsf{D}$ never exceeds $\inf \mathsf{P}$ (weak duality); $\sup \mathsf{D}$ equals $\inf \mathsf{P}$ for convex problems satisfying a constraint qualification (strong duality); all primal (dual) feasible solutions offer upper (lower) bounds on both $\inf \mathsf{P}$ and $\sup \mathsf{D}$.
- **Explicit dual problems:** The dual of an LP/QP/SOCP/SDP is also an LP/QP/SOCP/SDP, respectively.
- **Dual cone:** a proper convex cone coincides with the dual of its dual (bidual) cone; the nonnegative orthant, the second-order cone and the positive semidefinite cone are self-dual.