Lagrangian Duality in 10 Minutes

David S. Rosenberg

New York University

February 13, 2018

Contents

A General Optimization Problem

General Optimization Problem: Standard Form

Inequality Constrained Optimization Problem: Standard Form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leqslant 0, \quad i = 1, \ldots, m \end{array}$$

where $x \in \mathbf{R}^n$ are the **optimization variables** and f_0 is the **objective function.**

- No assumptions on functions $f_0, \ldots, f_m$.
  - (In particular **no convexity assumptions**.)

The Primal and the Dual

- For any **primal form** optimization problem,

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leqslant 0, \quad i = 1, \ldots, m, \end{array}$$

there is a recipe for constructing a corresponding **Lagrangian dual problem:**

$$\begin{array}{ll} \text{maximize} & g(\lambda) \\ \text{subject to} & \lambda_i \geqslant 0, \quad i = 1, \ldots, m, \end{array}$$

where $\lambda = (\lambda_1, \ldots, \lambda_m)$ are called **Lagrange multipliers** or **dual variables**.

In this formulation, g may take the value $-\infty$. Can get rid of this with additional constraints.

The Dual is Always a Convex Problem

- For any primal problem (convex or not), g is a **concave function**.
- Thus the dual is a **concave maximization** problem:

$$\begin{array}{ll} \text{maximize} & g(\lambda) \\ \text{subject to} & \lambda_i \geqslant 0, \ \ i = 1, \ldots, m. \end{array}$$

- Switch sign of g and change max $\mapsto$ min to get a convex optimization problem.
- Because of the trivial equivalence to a convex optimization problem, concave maximization problems are also typically considered convex optimization problems.
- Can the dual problem help us solve the primal problem?

Lagrangian Duality

Primal and Dual Optimal Points (Definitions)

Primal problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leqslant 0, \ \ i = 1, \ldots, m, \end{array}$$

Dual problem

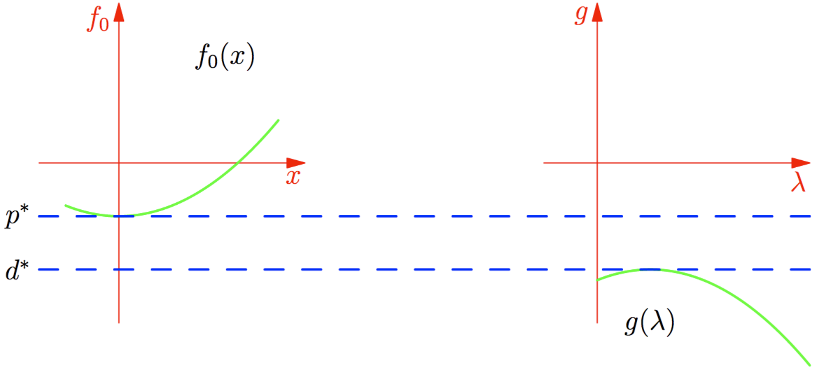
$$\begin{array}{ll} \text{maximize} & g(\lambda) \\ \text{subject to} & \lambda_i \geqslant 0, \ \ i = 1, \ldots, m, \end{array}$$

- The **primal optimal value** is $p^* = \inf\{f_0(x) \mid x \text{ satisfies all constraints}\}$.
- x^* is an **primal optimal point** if x^* is feasible and $f(x^*) = p^*$.
- The **dual optimal value** is $d^* = \sup\{g(\lambda) \mid \lambda_i \geqslant 0, \ i = 1, \ldots, m\}$.
- λ^* is a **dual optimal point** if $\lambda_i^* \geqslant 0, \ i = 1, \ldots, m$ and $g(\lambda^*) = d^*$.
  - λ_i^*'s are also called **optimal Lagrange multipliers**.

Weak Duality

- For any optimization problem, we have $p^* \geqslant d^*$.
- This is called **weak duality.**

Weak Duality – Illustrated



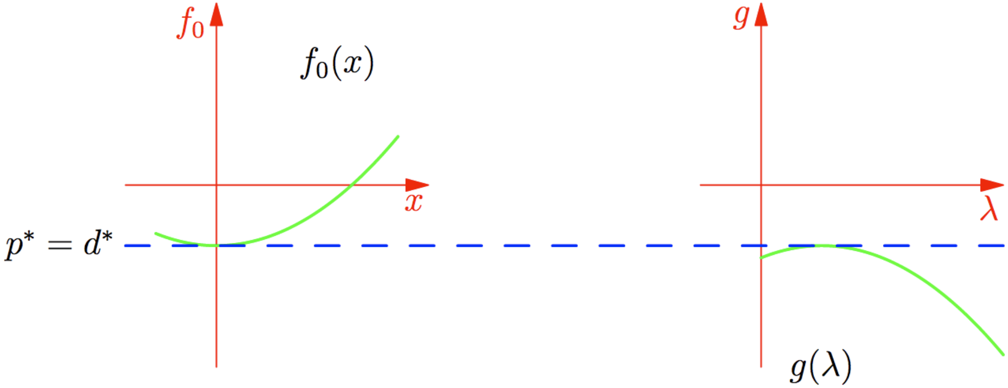
We **always** have **weak duality:** $p^* \geqslant d^*$.

Plot courtesy of Brett Bernstein.

Strong Duality

- For some problems, we have **strong duality:** $p^* = d^*$.
- For ***convex*** problems, **strong duality** is fairly typical.

Strong Duality – Illustrated



Under certain conditions, we have **strong duality:** $p^* = d^*$.

Plot courtesy of Brett Bernstein.

From Dual Solution to Primal?

- Suppose λ^* is the dual optimal solution.
- Does this help us find x^*, the primal optimal solution?

- In general, it may not be easy to go from λ^* to x^*.
- It depends on the form of the primal problem.
- For SVMs, we'll see that it's easy to go from dual to primal solution.

Convex Optimization

Convex Optimization Problem: Standard Form

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leqslant 0, \;\; i = 1, \ldots, m
\end{array}
$$

where $f_0, \ldots, f_m$ are convex functions.

Slater's Constraint Qualifications for Strong Duality

- For a **convex** optimization problem over domain $\mathbf{R}^n$,
- a **sufficient** condition for **strong duality** is

$$\exists x \in \mathbf{R}^d \text{ such that } f_i(x) < 0 \text{ for } i = 1, \ldots, m.$$

- Such an x is called a **strictly feasible** point.

Consequences of Strong Duality

Complementary Slackness

- If we have **strong duality**, we get an interesting relationship between
  - the optimal Lagrange multiplier λ_i^* and
  - the ith constraint at the optimum: $f_i(x^*)$
- Relationship is called "**complementary slackness**":

$$\lambda_i^* f_i(x^*) = 0$$

- Implies that at optimum, at least one of the following is satisfied:

$$\begin{aligned} \lambda_i^* &= 0 \\ f_i(x^*) &= 0 \text{ (constraint is "}\textbf{active}\text{")} \end{aligned}$$