Optimization Theory and Algorithms

Instructor: Prof. LIAO, Guocheng （廖国成）

Email: liaogch6@mail.sysu.edu.cn

School of Software Engineering
Sun Yat-sen University

Outline

- Lagrange dual problem
- Weak duality and strong duality
- KKT conditions
- Saddle point

Motivation of duality theory

- Helps analyze and even solve the original difficult problem from an easier dual problem
- Obtain some properties of the original problem by analyzing dual problem
- Sensitivity analysis

Lagrangian

Standard form optimization problem

$$\begin{aligned} \text{min} \ \ & f_0(x) \\ \text{s.t.} \ \ & f_i(x) \le 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{aligned}$$

variable $x \in \mathbb{R}^n$; optimal value $p^\star$; not necessarily convex

Lagrangian: $L: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$

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

- λ_i is Lagrange multiplier associated with $f_i(x) \le 0, i = 1, \dots, m$
- ν_i is Lagrange multiplier associated with $h_i(x) = 0, i = 1, \dots, p$
- Lagrangian: objective function + weighted sum of constraint functions

Lagrangian dual function

Lagrange dual function (or just *dual function*): $g: \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$

$$
\begin{aligned}
g(\lambda, \nu) &= \inf_x L(x, \lambda, \nu) \\
&= \inf_x \; f_0(x) + \textstyle\sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)
\end{aligned}
$$

Dual function is the pointwise infimum of affine functions of (λ, ν), so it is concave.

Lower bound property: for any $\lambda \geq 0$ and any ν, we have

$$g(\lambda, \nu) \leq p^\star$$

Proof: let x' is feasible, i.e., $f_i(x') \leq 0$ and $h_i(x') = 0$:

$$
\begin{aligned}
g(\lambda, \nu) &= \inf_x \; f_0(x) + \textstyle\sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \\
&\leq f_0(x') + \textstyle\sum_{i=1}^m \lambda_i f_i(x') + \sum_{i=1}^p \nu_i h_i(x') \\
&\leq f_0(x') \\
&\leq p^\star
\end{aligned}
$$

If $\lambda < 0$ otherwise, $g(\lambda, \nu) = -\infty$, which is meaningless.

Lagrange dual problem

Motivation: to make the lower bound $g(\lambda, \nu)$ of $p^\star$ as large as possible

Lagrange dual problem (or just *dual problem*):

$$\begin{aligned} \text{max} \;\; & g(\lambda, \nu) \\ \text{s.t.} \;\; & \lambda \geq 0 \end{aligned}$$

- Dual problem is a convex problem (concave function maximization subject to convex constraint function)
- (λ, ν) is dual feasible if $\lambda \geq 0$ and $g(\lambda, \nu) > -\infty$
- $(\lambda^\star, \nu^\star)$ is dual optimal (or optimal Lagrange multipliers) if they are optimal for the dual problem

Primal problem

$$\begin{aligned} \text{min} \;\; & f_0(x) \\ \text{s.t.} \;\; & f_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{aligned}$$

Dual problem

$$\begin{aligned} \text{max} \;\; & g(\lambda, \nu) \\ \text{s.t.} \;\; & \lambda \geq 0 \end{aligned}$$

Examples

primal problem (standard form LP)

$$\begin{aligned} \min\ & c^T x \\ \text{subject to}\ & Ax = b \\ & x \succcurlyeq 0 \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \min\ & c^T x \\ \text{subject to}\ & b - Ax = 0 \\ & -x \preccurlyeq 0 \end{aligned}$$

Lagrangian

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x - \lambda^T x + \nu^T (b - Ax) \\ &= b^T \nu + (c - A^T \nu - \lambda)^T x \end{aligned}$$

- λ_i is associated with inequality constraint $f_i(x) = -x_i \leq 0,\ i = 1, \dots, n$
- ν_i is associated with equality constraint $f_i(x) = b_i - a_i^T x,\ i = 1, \dots, m$

Dual function

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = b^T \nu + \inf_x\ (c - A^T \nu - \lambda)^T x = \begin{cases} b^T \nu, & c - A^T \nu - \lambda = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

Dual problem

$$\begin{aligned} \max\ & g(\lambda, \nu) = \begin{cases} b^T \nu, & c - A^T \nu - \lambda = 0 \\ -\infty, & \text{otherwise} \end{cases} \\ \text{s.t.}\ & \lambda \geq 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} \max\ & b^T \nu \\ \text{s.t.}\ & \lambda \geq 0 \\ & c - A^T \nu - \lambda = 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} \max\ & b^T \nu \\ \text{s.t.}\ & A^T \nu \leq c \end{aligned}$$

Examples

primal problem (inequality form LP)

$$\begin{aligned} &\min \ c^T x \\ &\text{subject to} \quad Ax \preccurlyeq b \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} &\min \ c^T x \\ &\text{subject to} \quad Ax - b \preccurlyeq 0 \end{aligned}$$

Lagrangian

$$\begin{aligned} L(x, \lambda) &= c^T x + \lambda^T (Ax - b) \\ &= -b^T \lambda + (c + A^T \lambda)^T x \end{aligned}$$

Dual function

$$g(\lambda) = \inf_x L(x, \lambda) = -b^T \lambda + \inf_x \ (c + A^T \lambda)^T x = \begin{cases} -b^T \lambda, & c + A^T \lambda = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

Dual problem

$$\begin{aligned} &\max \ g(\lambda, \nu) = \begin{cases} -b^T \lambda, & c + A^T \lambda = 0 \\ -\infty, & \text{otherwise} \end{cases} \\ &\text{s.t.} \ \ \lambda \geq 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} &\max \ -b^T \lambda \\ &\text{s.t.} \quad c + A^T \lambda = 0 \\ &\qquad\ \ \lambda \geq 0 \end{aligned}$$

Examples

primal problem （quadratic programming）

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$

Lagrangian

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

Dual function

- Take the gradient with respect to x, and set the gradient equal to zero

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \Longrightarrow \quad x = -(1/2) A^T \nu$$

- Plug in L to get g:

$$g(\nu) = L((-1/2)A^T \nu, \nu) = -\frac{1}{4} \nu^T A A^T \nu - b^T \nu$$

Dual problem

$$\max \ -\frac{1}{4} \nu^T A A^T \nu - b^T \nu$$

Examples

primal problem (non-convex)

$$
\begin{array}{ll}
\text{minimize} & x^T W x \\
\text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n
\end{array}
$$

Lagrangian

$$
\begin{aligned}
L(x, \nu) &= x^T W x + \sum_{i=1}^{n} \nu_i (x_i^2 - 1) \\
&= x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu.
\end{aligned}
$$

Dual function

$$
\begin{aligned}
g(\nu) &= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\
&= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise,} \end{cases}
\end{aligned}
$$

The infimum of a quadratic form is either zero (positive semidefinite) or $-\infty$ (not positive semidefinite)

Dual problem

$$
\begin{array}{ll}
\text{maximize} & -\mathbf{1}^T \nu \\
\text{subject to} & W + \mathbf{diag}(\nu) \succeq 0
\end{array}
$$

Take $\nu = -\lambda_{min}(W)\mathbf{1}$, we get a lower bound $p^\star \geq n\lambda_{min}(W)$

Primal problem v.s. dual problem

Primal problem

$$\begin{aligned} \text{min} \;\; & f_0(x) \\ \text{s.t.} \;\; & f_i(x) \leq 0, i = 1, \ldots, m \\ & h_i(x) = 0, i = 1, \ldots, p \end{aligned}$$

Dual problem

$$\begin{aligned} \text{max} \;\; & g(\lambda, \nu) \\ \text{s.t.} \;\; & \lambda \geq 0 \end{aligned}$$

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

Dual problem: $\max_{\lambda \geq 0, \nu} \min_{x} L(x, \lambda, \nu)$

Primal problem: $\min_{x} \max_{\lambda \geq 0, \nu} L(x, \lambda, \nu)$ $\Rightarrow$

$$\begin{aligned} \text{min} \;\; & f_0(x) \\ \text{s.t.} \;\; & f_i(x) \leq 0, i = 1, \ldots, m \\ & h_i(x) = 0, i = 1, \ldots, p \end{aligned}$$

$$\max_{\lambda \geq 0, \nu} L(x, \lambda, \nu) = \begin{cases} f_0(x), & \text{if } f_i(x) \leq 0, h_i(x) = 0 \\ +\infty, & \text{otherwise} \end{cases}$$

Weak duality

Primal problem

$$
\begin{aligned}
p^\star = \min \;\; & f_0(x) \\
\text{s.t.} \;\; & f_i(x) \le 0, i = 1, \dots, m \\
& h_i(x) = 0, i = 1, \dots, p
\end{aligned}
$$

Dual problem

$$
\begin{aligned}
d^\star = \max \;\; & g(\lambda, \nu) \\
\text{s.t.} \;\; & \lambda \ge 0
\end{aligned}
$$

- $p^\star$: optimal value of primal problem; $d^\star$: optimal value of dual problem

Lower bound property: for any $\lambda \ge 0$ and any ν, we have

$$g(\lambda, \nu) \le p^\star$$

$$\Downarrow$$

Weak duality: $d^\star \le p^\star$

$$\max_{\lambda \ge 0, \nu} \min_x L(x, \lambda, \nu) \le \min_x \max_{\lambda \ge 0, \nu} L(x, \lambda, \nu)$$

Duality gap: $p^\star - d^\star$

Weak duality

Primal problem

$$
\begin{aligned}
p^\star = \min\ & f_0(x) \\
\text{s.t.}\ & f_i(x) \le 0, i = 1, \dots, m \\
& h_i(x) = 0, i = 1, \dots, p
\end{aligned}
$$

Dual problem

$$
\begin{aligned}
d^\star = \max\ & g(\lambda, \nu) \\
\text{s.t.}\ & \lambda \ge 0
\end{aligned}
$$

- $p^\star$: optimal value of primal problem; $d^\star$: optimal value of dual problem

Weak duality: $d^\star \le p^\star$

- $p^\star = -\infty \Longrightarrow d^\star = -\infty$ (If the primal problem is unbounded below, dual problem is infeasible)
- $d^\star = \infty \Longrightarrow p^\star = \infty$ (If the dual problem is unbounded above, primal problem is infeasible)

Strong duality

Primal problem

$$\begin{aligned} p^\star = \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \le 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{aligned}$$

Dual problem

$$\begin{aligned} d^\star = \max\ & g(\lambda, \nu) \\ \text{s.t.}\ & \lambda \ge 0 \end{aligned}$$

Strong duality: $d^\star = p^\star$

- The best bound obtained from dual function is tight.
- Does not hold in general
- Sufficient conditions for strong duality are called **constraint qualifications**
- Strong duality usually holds for convex optimization

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \le 0, i = 1, \dots, m \\ & a_i^T x = b_i, i = 1, \dots, p \end{aligned}$$