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
- Sensitivity analysis
- Generalized inequality

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

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

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 $f_i(x) \leq 0, i = 1, \dots, m$
- 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*): $g: \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

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

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 $h_i(x') = 0$:

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

$g(\lambda, \nu) \leq f_0(x')$ holds for any feasible x'. Thus, $g(\lambda, \nu) \leq f_0(x^\star) = p^\star$.

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*):

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

- 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

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

Dual problem

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

Examples

primal problem (standard form LP)

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

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

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

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

Examples

primal problem (inequality form LP)

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

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

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

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

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

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

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

Dual problem

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

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

Dual problem: $\max\limits_{\lambda \ge 0, \nu} \min\limits_{x} L(x, \lambda, \nu)$

Primal problem: $\min\limits_{x} \max\limits_{\lambda \ge 0, \nu} L(x, \lambda, \nu)$ $\Rightarrow$

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

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

Weak duality

Primal problem

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

Dual problem

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

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

Lower bound property: for any $\lambda \geq 0$ 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

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

Dual problem

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

- $p^\star$: 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

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

Dual problem

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

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

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

Slater's condition

One simple constraint qualification: convex optimization problem + *Slater's condition*

Convex optimization problem

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

Slater's condition*:* there exists an $\boxed{x'} \in \mathbf{int}\,\mathcal{D}$ such that $f_i(x') < 0, i = 1, \dots, m, Ax = b.$

strictly feasible point

Slater's condition (weak form)*:* if some inequality constraint functions are affine, e.g., $f_1, \dots, f_k$ are affine: there exists an x' such that

$$f_i(x') \leq 0, i = 1, \dots, k, f_i(x') < 0, i = k + 1, \dots, m, Ax = b.$$

If the problem is a convex optimization problem and *Slater's condition* holds, then strong duality holds.

Complementary slackness

What can we learn from strong duality?

Suppose strong duality holds. Let $x^\star$ and $(\lambda^\star, \nu^\star)$ be primal and dual optimal, respectively.

$$g(\lambda^\star, \nu^\star) = f_0(x^\star)$$

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

$$\boxed{\begin{aligned} &\le f_0(x^\star) + \textstyle\sum_{i=1}^{m} \lambda_i^\star f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star h_i(x^\star) \\ &\le f_0(x^\star) \end{aligned}}$$

Equality holds: $\lambda_i^\star f_i(x^\star) = 0, \, i = 1, \dots, m$

Complementary slackness: $\lambda_i^\star f_i(x^\star) = 0, \, i = 1, \dots, m$

$$\lambda_i^\star > 0 \Rightarrow f_i(x^\star) = 0$$
$$f_i(x^\star) < 0 \Rightarrow \lambda_i^\star = 0$$

Lagrangian optimality

What can we learn from strong duality?

Suppose strong duality holds. Let $x^\star$ and $(\lambda^\star, \nu^\star)$ be primal and dual optimal, respectively.

$$g(\lambda^\star, \nu^\star) = f_0(x^\star)$$

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

$$\le f_0(x^\star)$$

Equality holds: $\inf_x \, L(x, \lambda^\star, \nu^\star) = L(x^\star, \lambda^\star, \nu^\star)$

Lagrangian optimality: $\nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(x^\star) = 0$

Karush-Kuhn-Tucker (KKT) conditions

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

Suppose f_i and h_i are differentiable.

KKT conditions

- **Complementary slackness:** $\lambda_i^\star f_i(x^\star) = 0, i = 1, \dots, m$
- **Lagrangian optimality:** $\nabla f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^p \nu_i^\star \nabla h_i(x^\star) = 0$
- **Primal feasibility:** $f_i(x^\star) \le 0, i = 1, \dots, m, h_i(x^\star) = 0, i = 1, \dots, p$
- **Dual feasibility**: $\lambda_i^\star \ge 0, i = 1, \dots, m$

Strong duality holds. $x^\star$ and $(\lambda^\star, \nu^\star)$ are primal and dual optimal, respectively. $\Rightarrow$ $x^\star$ and $(\lambda^\star, \nu^\star)$ satisfy KKT conditions.

KKT conditions are necessary conditions for strong duality and optimality

Karush-Kuhn-Tucker (KKT) conditions

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

Suppose f_i and h_i are differentiable, and the problem is a convex optimization problem.

KKT conditions

- **Complementary slackness:** $\lambda_i^\star f_i(x^\star) = 0, i = 1, \dots, m$
- **Lagrangian optimality:** $\nabla f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^p \nu_i^\star \nabla h_i(x^\star) = 0$
- **Primal feasibility:** $f_i(x^\star) \leq 0, i = 1, \dots, m, h_i(x^\star) = 0, i = 1, \dots, p$
- **Dual feasibility**: $\lambda_i^\star \geq 0, i = 1, \dots, m$

$x^\star$ and $(\lambda^\star, \nu^\star)$ satisfy KKT conditions.

→

Strong duality holds. $x^\star$ and $(\lambda^\star, \nu^\star)$ are primal and dual optimal, respectively.

KKT conditions are sufficient conditions for strong duality and optimality of a convex optimization problem.

Karush-Kuhn-Tucker (KKT) conditions

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

Suppose f_i and h_i are differentiable, the problem is a convex optimization problem and satisfies Slater's conditions.

KKT conditions

- **Complementary slackness:** $\lambda_i^\star f_i(x^\star) = 0, i = 1, \dots, m$
- **Lagrangian optimality:** $\nabla f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^p \nu_i^\star \nabla h_i(x^\star) = 0$
- **Primal feasibility:** $f_i(x^\star) \leq 0, i = 1, \dots, m, h_i(x^\star) = 0, i = 1, \dots, p$
- **Dual feasibility**: $\lambda_i^\star \geq 0, i = 1, \dots, m$

$x^\star$ and $(\lambda^\star, \nu^\star)$ satisfy KKT conditions. $\Longleftrightarrow$ $x^\star$ and $(\lambda^\star, \nu^\star)$ are primal and dual optimal, respectively.

For a convex optimization that satisfies Slater' conditions, KKT conditions are sufficient and necessary conditions for strong duality and optimality.

Example

Water-filling.

- To allocate power to a set of n communication channels to maximize total communication rate.
- $\log(\alpha_i + x_i)$ is the communication rates of channel i under power x_i and context-related parameter α_i

$$\begin{aligned} \text{min} \quad & -\textstyle\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ \text{s.t.} \quad & x \geq 0, \\ & \textstyle\sum_{i=1}^{n} x_i = 1 \end{aligned}$$

KKT conditions

- **Complementary slackness:** $\lambda_i^\star x_i^\star = 0, i = 1, \dots, n$
- **Lagragian optimality:** $-\frac{1}{\alpha_i + x_i^\star} - \lambda_i^\star + \nu^\star = 0, i = 1, \dots, n$ **(1)**
- **Primal feasibility:** $x_i^\star \geq 0, i = 1, \dots, n, \sum_{i=1}^{n} x_i^\star = 1$
- **Dual feasibility**: $\lambda_i^\star \geq 0, i = 1, \dots, m$

➢ If $x_i^\star > 0$, then $\lambda_i^\star = 0$. $\lambda_i^\star = 0$ and (1) give $x_i^\star = 1/\nu^\star - \alpha_i$

➢ If $\lambda_i^\star > 0$, then $x_i^\star = 0$. $\lambda_i^\star > 0$ and (1) give $\nu^\star \geq 1/\alpha_i$

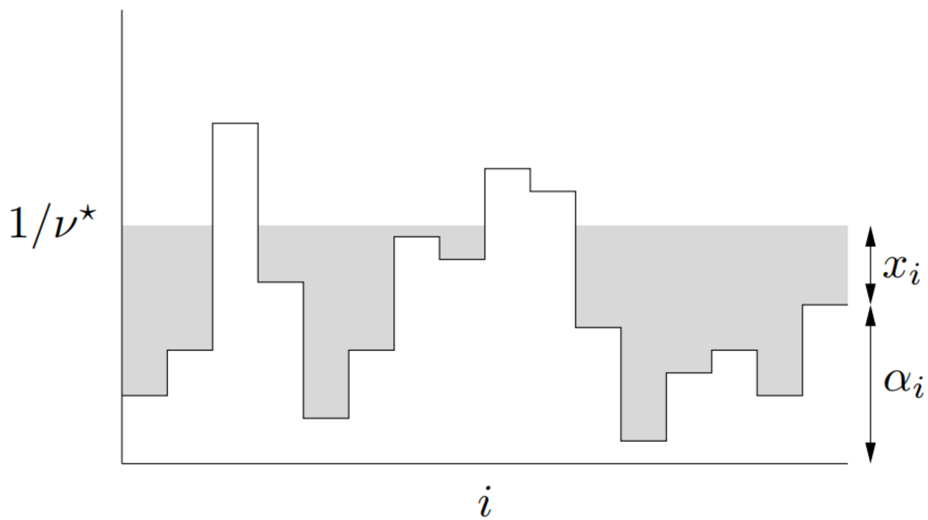
$$\Rightarrow \quad x_i^\star = \begin{cases} \frac{1}{\nu^\star} - \alpha_i, & \nu^\star < \frac{1}{\alpha_i} \\ 0, & \nu^\star \geq \frac{1}{\alpha_i} \end{cases}$$

Example

$$\begin{aligned} &\min \ -\textstyle\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ &\text{s.t.} \ \ x \geq 0, \\ &\qquad \textstyle\sum_{i=1}^{n} x_i = 1 \end{aligned}$$

Solution $\quad x_i^\star = \begin{cases} 1/\nu^\star - \alpha_i, & \nu^\star < 1/\alpha_i \\ 0, & \nu^\star \geq 1/\alpha_i \end{cases} \quad$ or $\quad x_i^\star = \max\{0, 1/\nu^\star - \alpha_i\}$,

where $\nu^\star$ is such that $\sum_{i=1}^{n} \max\{0, 1/\nu^\star - \alpha_i\} = 1$



Saddle point

Primal problem

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

Dual problem

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

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

Primal problem:

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

Dual problem:

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

(x', λ', ν') where $\lambda' \geq 0$ is a saddle point of the Lagrangian function L if

$$L(x', \lambda, \nu) \leq L(x', \lambda', \nu') \leq L(x, \lambda', \nu')$$

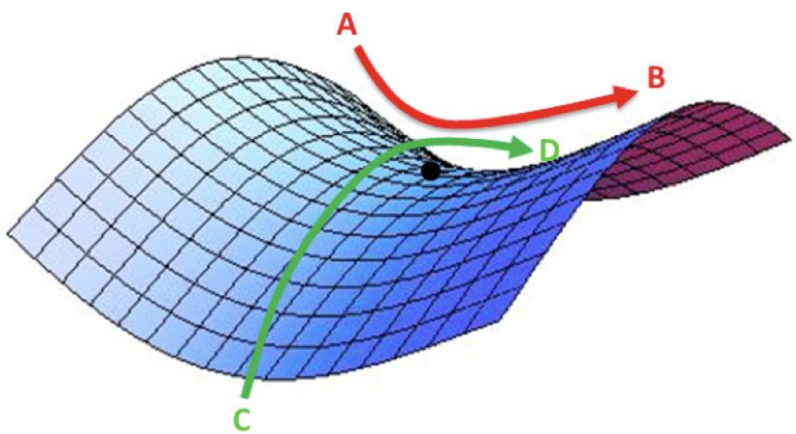
- x' minimizes L when (λ, ν) is fixed at (λ', ν'), i.e., $L(x', \lambda', \nu') = \min_{x} L(x, \lambda', \nu')$.
- (λ', ν') maximize L when x is fixed at x', i.e., $L(x', \lambda', \nu') = \max_{\lambda \geq 0, \nu} L(x', \lambda, \nu)$

Saddle point

(x', λ', ν') where $\lambda' \geq 0$ is a saddle point of the Lagrangian function L if

$$L(x', \lambda, \nu) \leq L(x', \lambda', \nu') \leq L(x, \lambda', \nu')$$

- x' minimizes L when (λ, ν) is fixed at (λ', ν'), i.e., $L(x', \lambda', \nu') = \min_x L(x, \lambda', \nu')$.
- (λ', ν') maximize L when x is fixed at x', i.e., $L(x', \lambda', \nu') = \max_{\lambda \geq 0, \nu} L(x', \lambda, \nu)$



(x', λ', ν') is a saddle point of L if and only if x' and (λ', ν') are primal and dual optimal, respectively, and strong duality holds.

Shadow price Interpretation

$$\begin{aligned} \max \quad & c^T x \\ \text{subject to} \quad & Ax \le b \\ & x \ge 0 \end{aligned} \quad \xrightarrow{\text{dual}} \quad \begin{aligned} \min \quad & b^T \lambda \\ \text{s.t.} \quad & A^T \lambda \ge c \\ & \lambda \ge 0 \end{aligned}$$

Production planning (primal problem).

- To determine the quantities of n products to maximize total profit s.t. resource constraints.
- c: profit; A: resource consumption; b: available resource.

Lagrange: $L(x, \lambda, \alpha) = -c^T x + \boxed{\lambda^T}(Ax - b) - \alpha^T x$

Shadow price for each resource:
Have unused resource: sell
Need more resource: buy

$$g(\lambda) = \min_x L(x, \lambda, \alpha) = -b^T \lambda + \min_x x^T (A^T \lambda - c - \alpha)$$

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

$$\max \quad g(\lambda, v) = \begin{cases} -b^T \lambda, & A^T \lambda - c - \alpha = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

$$\text{s.t.} \quad \lambda \ge 0, \alpha \ge 0$$

Weak duality:
profit of product $\le$ value of resources

Resource purchase (dual problem).

- To determine the prices of m resources to minimize total cost.
- Constraint: for each product, payment of selling the resources $\ge$ profit of selling the product

Sensitivity analysis

Primal optimization problem and its dual

Primal problem

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

Dual problem

$$\max \ g(\lambda, \nu)$$
$$\text{s.t.} \ \ \lambda \geq 0$$

Perturbed optimization problem and its dual

Primal problem

$$p^\star(\mu, \upsilon) = \min \ f_0(x)$$
$$\text{s.t.} \ \ f_i(x) \leq \mu_i, i = 1, \dots, m$$
$$h_i(x) = \upsilon_i, i = 1, \dots, p$$

Dual problem

$$\max \ g(\lambda, \nu) - \lambda^T \mu - \nu^T \upsilon$$
$$\text{s.t.} \ \ \lambda \geq 0$$

- μ and υ are the parameters that perturb the inequality and equality constraints.
- $\mu_i > 0$: constraint becomes relaxed; $\mu_i < 0$: constraint becomes tightened.
- What is the connection between $p^\star(\mu, \upsilon)$ and $p^\star$?

Sensitivity analysis

Assume strong duality holds. Let $\lambda^\star, \nu^\star$ be optimal dual variable for unperturbed problem.

Connection between $p^\star\ (\mu, \upsilon)$ and $p^\star$:

$$p^\star\ (\mu, \upsilon) \geq p^\star - \mu^T \lambda^\star - \upsilon^T \nu^\star$$

- If $\mu_i < 0$ and $\lambda_i^\star$ is large , then $p^\star\ (\mu, \upsilon)$ will increase greatly.
- If $\mu_i > 0$ and $\lambda_i^\star$ is small, then $p^\star\ (\mu, \upsilon)$ will not decrease greatly.
- If $\upsilon_i < 0$ and $\nu_i^\star$ is large and positive, $p^\star\ (\mu, \upsilon)$ will increase greatly.

  If $\upsilon_i > 0$ and $\nu_i^\star$ is large and negative, $p^\star\ (\mu, \upsilon)$ will increase greatly.
- If $\upsilon_i > 0$ and $\nu_i^\star$ is small and positive, $p^\star\ (\mu, \upsilon)$ will not decrease greatly.

  If $\upsilon_i < 0$ and $\nu_i^\star$ is small and negative, $p^\star\ (\mu, \upsilon)$ will not decrease greatly.

Example: shadow price

Production planning: to determine the quantities of n products to maximize total profit subject to resource constraints.

$$\begin{aligned} \max \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

- λ that is associated with inequality constraint: shadow price of each resource.
- $\lambda_i^\star$ tells how much more profit the firm could make, for a small increase of resource i.
- If $\lambda_i^\star$ is larger, more available resource makes the firm earn more profit.

  If $\lambda_i^\star$ is smaller, more available resource does not make the firm earn more profit.

Generalized inequality

Optimization problem with general inequality constraints

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

- $K_i \subseteq \mathbb{R}^{k_i}$ is a proper cone; $\preccurlyeq_{K_i}$ is general inequality on $\mathbb{R}^{k_i}$.

Generalized inequality: Lagrangian

Primal problem

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

Lagrangian: $L: \mathbb{R}^n \times \mathbb{R}^{k_i} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \to \mathbb{R}$

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

- $\lambda_i \subseteq \mathbb{R}^{k_i}$ is Lagrange multiplier associated with $f_i(x) \preccurlyeq_{K_i} 0, i = 1, \dots, m$
- $\nu_i \subseteq \mathbb{R}$ is Lagrange multiplier associated with $h_i(x) = 0, i = 1, \dots, p$
- Lagrangian: objective function + weighted sum of constraint functions

Generalized inequality: dual function

Dual function: $g: \mathbb{R}^{k_i} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \to \mathbb{R}$

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu)$$

$$= \inf_x \; f_0(x) + \sum_{i=1}^{m} \lambda_i^T f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

Scalar version of lower bound property:
for any $\lambda \geq 0$, we have

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

Lower bound property: for any $\lambda_i \succcurlyeq_{K_i^*} 0$, we have

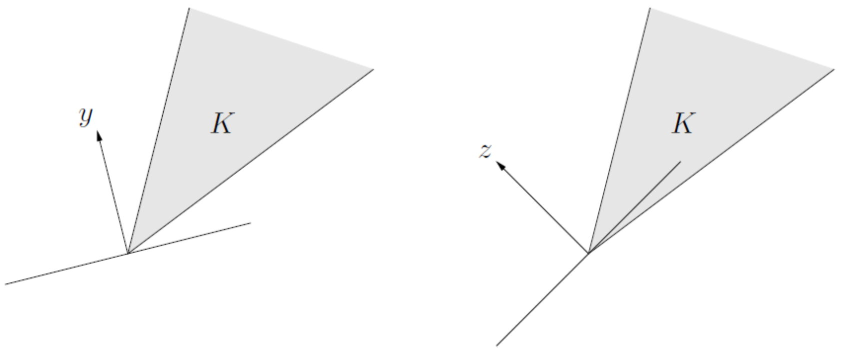
$$g(\lambda, \nu) \leq p^\star$$

- K_i^* is dual cone of K_i.

Dual cone

Dual cone of K:

$$K^* = \{y \mid y^T x \geq 0 \text{ for all } x \in K\}$$



$y \in K^*$ $\qquad$ $z \notin K^*$

Examples:

- $K = \mathbb{R}^n_+$: $K^* = \mathbb{R}^n_+$
- $K = \mathbf{S}^n_+$: $K^* = \mathbf{S}^n_+$
- $K = \{(x,t) \mid \|x\|_2 \leq t\}$: $K^* = \{(x,t) \mid \|x\|_2 \leq t\}$

Dual cones of proper cones are proper, hence define generalized inequality:

$$y \succcurlyeq_{K^*} 0 \quad \Longleftrightarrow \quad y^T x \geq 0 \text{ for all } x \succcurlyeq_K 0$$

Generalized inequality: dual function

Dual function: $g: \mathbb{R}^{k_i} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \to \mathbb{R}$

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

Scalar version of lower bound property:

for any $\lambda \geq 0$, we have

$$g(\lambda, \nu) \leq f_0(x') + \sum_{i=1}^m \underbrace{\boxed{\lambda_i f_i(x')}}_{\leq\, 0} + \sum_{i=1}^p \nu_i h_i(x') \leq f_0(x')$$

Lower bound property:

for any $\lambda_i \succcurlyeq_{K_i^*} 0$, we have

$$\begin{aligned} g(\lambda, \nu) &\leq f_0(x') + \sum_{i=1}^m \underbrace{\underbrace{\boxed{\lambda_i^T}}_{\succcurlyeq_{K_i^*} 0} \underbrace{\boxed{f_i(x')}}_{\preccurlyeq_{K_i} 0}}_{\leq\, 0} + \sum_{i=1}^p \nu_i h_i(x') \\ &\leq f_0(x') \end{aligned}$$

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

Generalized inequality: dual problem

$$\begin{aligned} \max\ & g(\lambda, \nu) \\ \text{s.t.}\ & \lambda_i \succcurlyeq_{K_i^*} 0, i = 1, \dots, m \end{aligned}$$

- $p^\star$: optimal value of primal problem; $d^\star$: optimal value of dual problem
- Weak duality: $d^\star \leq p^\star$
- Strong duality: $d^\star = p^\star$
- Strong duality holds when the primal problem is convex and satisfies Slater's condition

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

***Slater's condition**:* there exists an $x' \in \mathbf{int}\, \mathcal{D}$ such that $f_i(x') \prec_{K_i} 0, i = 1, \dots, m, Ax = b.$

Conic programming and its dual

Standard form conic programming

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

Lagrangian

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

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

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

Example

Standard form linear programming

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

Lagrangian

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

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

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

$\Rightarrow$

$$\begin{aligned} \text{max} \quad & b^T \nu \\ \text{s.t.} \quad & \lambda \geq 0 \\ & c - A^T \nu - \lambda = 0 \end{aligned}$$

$$\begin{aligned} \text{max} \quad & b^T \nu \\ \text{s.t.} \quad & A^T \nu \leq c \end{aligned}$$

Dual cone of non-negative orthant is itself

Example

Standard form second-order cone programming

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Ax = b \\ & x \succcurlyeq_{Q^{n+1}} 0 \end{array}$$

$Q^{n+1} = \{(x,t) | \ \|x\|_2 \leq t\}$:

Lagrangian

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

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

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

$$\text{s.t. } \lambda \succcurlyeq_{Q^{n+1}} 0$$

$$\Rightarrow \quad \begin{array}{rl} \max & b^T \nu \\ \text{s.t.} & \lambda \succcurlyeq_{Q^{n+1}} 0 \\ & c - A^T \nu - \lambda = 0 \end{array}$$

$$\begin{array}{rl} \max & b^T \nu \\ \text{s.t.} & A^T \nu \preccurlyeq_{Q^{n+1}} c \end{array}$$

Dual cone of second-order cone is itself

Example

Standard form semidefinite programming

$$\begin{aligned} \text{min} \quad & \mathrm{Tr}(CX) \\ \text{subject to} \quad & \mathrm{Tr}(A_i X) = b_i, \quad \text{for } i = 1, \dots, p \\ & X \succcurlyeq_{\mathbf{S}_+^n} \mathbf{0} \end{aligned}$$

Lagrangian

$$\begin{aligned} L(X, Z, \nu) &= \mathrm{Tr}(CX) - \mathrm{Tr}(ZX) + \textstyle\sum_{i=1}^p \nu_i \left(b_i - \mathrm{Tr}(A_i X)\right) \\ &= b^T \nu + \mathrm{Tr}\left(\left(C - Z - \textstyle\sum_{i=1}^p \nu_i A_i\right) X\right) \end{aligned}$$

Dual function

$$\begin{aligned} g(Z, \nu) = \inf_x L(X, Z, \nu) &= b^T \nu + \inf_x \mathrm{Tr}\left(\left(C - Z - \textstyle\sum_{i=1}^p \nu_i A_i\right) X\right) \\ &= \begin{cases} b^T \nu, & C - Z - \sum_{i=1}^p \nu_i A_i = 0 \\ -\infty, & \text{otherwise} \end{cases} \end{aligned}$$

Dual problem

$$\text{max} \quad g(Z, \nu) = \begin{cases} b^T \nu, & C - Z - \sum_{i=1}^p \nu_i A_i = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

$$\text{s.t.} \quad Z \succcurlyeq_{\mathbf{S}_+^n} 0$$

Dual cone of positive semidefinite matrix is itself

$$\Rightarrow \quad \begin{aligned} \text{max} \quad & b^T \nu \\ \text{s.t.} \quad & Z \succcurlyeq_{\mathbf{S}_+^n} 0 \\ & C - Z - \sum_{i=1}^p \nu_i A_i = 0 \end{aligned}$$

$$\begin{aligned} \text{max} \quad & b^T \nu \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^p \nu_i A_i \preccurlyeq_{\mathbf{S}_+^n} C \end{aligned}$$

KKT conditions

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

Suppose f_i and h_i are differentiable, the problem is a convex optimization problem and satisfies Slater's conditions.

KKT conditions

- **Complementary slackness:** $\lambda_i^{\star T} f_i(x^\star) = 0, i = 1, \dots, m$
- **Lagrangian optimality:** $\nabla f_0(x^\star) + \sum_{i=1}^m \lambda_i^{\star T} D f_i(x^\star) + \sum_{i=1}^p \nu_i^\star \nabla h_i(x^\star) = 0$
- **Primal feasibility:** $f_i(x^\star) \preccurlyeq_{K_i} 0, i = 1, \dots, m, h_i(x^\star) = 0, i = 1, \dots, p$
- **Dual feasibility**: $\lambda_i^\star \succcurlyeq_{K^*} 0, i = 1, \dots, m$

$x^\star$ and $(\lambda^\star, \nu^\star)$ satisfy KKT conditions.

$\Longleftrightarrow$

$x^\star$ and $(\lambda^\star, \nu^\star)$ are primal and dual optimal, respectively.

For a convex optimization that satisfies Slater' conditions, KKT conditions are sufficient and necessary conditions for strong duality and optimality.