Part IV Optimality

Wenxun Xing

Department of Mathematical Sciences
Tsinghua University
Tel: 62787945
Email: wxing@tsinghua.edu.cn
Office hour: 4:00-5:00pm, Thursday
Office: The New Science Building, A416

Nov., 2019

Optimality and dual problems

Content

- Optimality conditions based on differentiation.
- Constraint qualifications.
- Lagrangian dual problems.
- Linear conic optimization problems.

Optimization problem

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) \le 0 \\ & x \in \mathbb{R}^n. \end{aligned}$$

- Feasible set: $\mathcal{F} = \{x \in \mathbb{R}^n \mid g_i(x) \le 0, i = 1, 2, \ldots, m\}$.
- Unconstrained optimization problem: $\mathcal{F} = \mathbb{R}^n$.
- Feasible problem: $\mathcal{F} \neq \emptyset$.
- Bounded below: The problem is bounded below.
- Attainable: There exists an $x^* \in \mathcal{F}$ such that $f(x^*)$ reaches the optimal value.
- Solvable: Feasible, bounded below and attainable.

- Local minimizer: For a given $x^* \in \mathcal{F}$, there exists a $\delta > 0$ such that

$$f(x^*) \leq f(x), \forall x \in N(x^*, \delta) \cap \mathcal{F}.$$

- Global minimizer: For a given $x^* \in \mathcal{F}$,

$$f(x^*) \leq f(x), \forall x \in \mathcal{F}.$$

- Strictly local/global optimizer: $\leq$ is replaced by $<$.
- An example of not attainable:

$$\begin{array}{ll} \min & x_1^2 \\ \text{s.t.} & x_1 x_2 = 1 \\ & x_1, x_2 \in \mathbb{R}, \end{array}$$

Theorem

If $\mathcal{F} \neq \emptyset$ is convex and $f(x)$ is a convex function over $\mathcal{F}$, then any local minimizer is a global minimizer.

The first order optimality conditions

- For a given $\bar{x} \in \mathcal{F}$ and $\delta > 0$,

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + o(\|x - \bar{x}\|), \; \forall\, x \in N(\bar{x}, \delta) \cap \mathcal{F}.$$

- Checking rule: $\bar{x}$ is not a local minimizer if there exists a $d \in \mathbb{R}^n$ such that $\nabla f(\bar{x})^T d < 0$ for $\forall 0 < \delta \leq \delta_0$ and $x = \bar{x} + \delta d \in \mathcal{F}$.
- Set of feasible directions:

$$\mathcal{D}(x) = \left\{ d \in \mathbb{R}^n \mid \exists \delta_0 > 0, \text{ such that } x + \delta d \in \mathcal{F}, \forall\, 0 < \delta \leq \delta_0 \right\}.$$

Theorem

If $\bar{x} \in \mathcal{F}$ is a local minimizer, then

$$\nabla f(\bar{x})^T d \geq 0, \forall d \in \mathcal{D}(\bar{x}).$$

Discussions on $\mathcal{D}(x)$

- Not sufficient.

$$\begin{array}{ll} \min & -x^4 \\ \text{s.t.} & -1 \le x \le 1. \end{array}$$

  Consider the point $\bar{x} = 0$. $\frac{df(\bar{x})}{dx} = -4\bar{x}^3 = 0$, but $\bar{x} = 0$ is not a local minimizer.

- $\mathcal{D}(x)$ is a cone, but it may not be closed or convex.
  Let

$$\mathcal{F}_1 = \left\{(x_1, x_2)^T \in \mathbb{R}^2 \mid (x_1 - 1)^2 + x_2^2 \le 1\right\}$$

  and consider $\bar{x} = (0, 0)^T$.

$$\mathcal{D}(\bar{x}) = \left\{(d_1, d_2)^T \in \mathbb{R}^2 \mid d_1 > 0\right\}$$

  is not closed.

Let

$$\mathcal{F}_2 = \left\{(x_1, x_2)^T \in \mathbb{R}^2 \mid 2x_1 - x_2 \leq 0, x_1 \geq 0, x_2 \geq 0\right\}$$
$$\cup \left\{(x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 - 2x_2 \geq 0, x_1 \geq 0, x_2 \geq 0\right\}.$$

and consider $\bar{x} = (0, 0)^T$.

$$\mathcal{D}(\bar{x}) = \left\{(d_1, d_2)^T \in \mathbb{R}^2 \mid 2d_1 - d_2 \leq 0, d_1 \geq 0, d_2 \geq 0\right\}$$
$$\cup \left\{(d_1, d_2)^T \in \mathbb{R}^2 \mid d_1 - 2d_2 \geq 0, d_1 \geq 0, d_2 \geq 0\right\},$$

is not convex.

Theorem

Suppose $\mathcal{F}$ be nonempty and convex, $f : \mathbb{R}^n$ be convex. Then $\bar{x} \in \mathcal{F}$ is a local minimizer if and only if

$$\nabla f(\bar{x})^T d \geq 0, \forall d \in \mathcal{D}(\bar{x}).$$

Constraints $g_i(x) \leq 0, i = 1, 2, \ldots, m$

- Active constraint set

$$\mathcal{I}(x) = \{i \mid g_i(x) = 0\}.$$

- Set of locally constrained directions

$$\mathcal{L}(x) = \left\{d \in \mathbb{R}^n \mid \nabla g_i(x)^T d \leq 0, \forall i \in \mathcal{I}(x)\right\}.$$

Lemma

If the optimization problem is feasible and all its constraints are continuous differential, then $\mathcal{L}(x)$ is a nonempty convex cone for any $x \in \mathcal{F}$ and

$$\mathcal{D}(x) \subseteq \mathcal{L}(x).$$

Theorem

(Karush-Kuhn-Tucker Theorem) Suppose all functions in the optimization problem be continuous differential and $\bar{x} \in \mathcal{F}$ be a local minimizer. If $\mathcal{L}(\bar{x}) \subseteq \mathrm{cl}(\mathrm{conv}(\mathcal{D}(\bar{x})))$, then there exists a $\bar{\lambda} \in \mathbb{R}^m_+$ such that

$$\begin{array}{l} \nabla f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla g_i(\bar{x}) = 0, \\ \bar{\lambda}_i g_i(\bar{x}) = 0, \quad i = 1, 2, \ldots, m. \end{array}$$

- Constraint qualifications: $\mathcal{L}(\bar{x}) \subseteq \mathrm{cl}(\mathrm{conv}(\mathcal{D}(\bar{x})))$.
- An example

$$\begin{array}{rl} \min & f(x) = x_1 \\ \text{s.t.} & g_1(x) = x_2 - x_1^3 \leq 0 \\ & g_2(x) = -x_2 \leq 0 \\ & x = (x_1, x_2)^T \in \mathbb{R}^2, \end{array}$$

$\bar{x} = (0, 0)^T$ is a global (local) minimizer, but violates

$$\nabla f(\bar{x}) + \lambda_1 \nabla g_1(\bar{x}) + \lambda_2 \nabla g_2(\bar{x}) = 0.$$

The second-order optimality conditions

Theorem

Suppose $\bar{x} \in \mathbb{R}^n$ be a KKT point and $f(x)$ be twice differential continuous at this point. Then $\nabla^2 f(\bar{x}) \in \mathcal{S}_+^n$ if $\bar{x}$ is a local minimizer. Moreover $\bar{x}$ is a local minimizer if $\nabla^2 f(\bar{x}) \in \mathcal{S}_{++}^n$.

Let $(\bar{x}, \bar{\lambda})$ be a KKT point. $\bar{\lambda}_i g_i(\bar{x}) = 0$ implies $g_i(\bar{x}) = 0$ if $\bar{\lambda}_i > 0$. Define

$$\overline{\mathcal{I}}(\bar{x}) = \{i \mid i \in \mathcal{I}(\bar{x}), \bar{\lambda}_i > 0\}.$$

- Locally constrained directions.

$$\overline{\mathcal{L}}(\bar{x}) = \left\{d \in \mathbb{R}^n \mid \nabla g_i(\bar{x})^T d = 0, i \in \overline{\mathcal{I}}(\bar{x}); \nabla g_i(\bar{x})^T d \leq 0, i \in \mathcal{I}(\bar{x}) \setminus \overline{\mathcal{I}}(\bar{x})\right\}.$$

Theorem

Suppose $f(x)$ and $g(x)$ be the second-order differential functions, $(\bar{x}, \bar{\lambda})$ be a KKT point. Define

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

$\bar{x}$ is a strictly local minimizer if

$$d^T \nabla_x^2 L(\bar{x}, \bar{\lambda}) d > 0, \ \forall d \in \overline{\mathcal{L}}(\bar{x}), d \neq 0.$$

- $\bar{x}$ is not a local minimizer: there exists a series $\{x^k\}_{k=1}^{+\infty} \subseteq \mathcal{F}$ or a direction $d \in \mathcal{D}(\bar{x})$ satisfying $f(\bar{x} + \delta d) \leq f(\bar{x})$ for all $0 < \delta \leq \delta_0$.
- An example. Consider a point $x = 0$ for the following function

$$f(x_1, x_2) = (x_1 - x_2^2)(x_1 - 3x_2^2).$$

For a direction $d = (0, 1)^T$, $f(\delta d) = 3\delta^4 > 0$. For a direction $d = (1, k)^T, k \in \mathbb{R}$, $f(\delta d) = (\delta - k^2\delta^2)(\delta - 3k^2\delta^2) > 0$ if $0 < \delta < \frac{1}{3k^2}$. But $f(x_1, x_2) = -x_2^4 < 0$ if $x_1 = 2x_2^2$.

- Tangent directions
  For a given x and $x^k = x + \theta_k d^k \in \mathcal{F}$ with $\{d^k\}_{k=1}^{+\infty} \subseteq \mathbb{R}^n$ and $\{\theta_k\}_{k=1}^{+\infty} \subseteq \mathbb{R}_+$, if $d^k \to d$ and $\theta_k \to 0$ when $k \to +\infty$, then d is called a tangent direction at x.
- Set of tangent directions

$$\mathcal{T}(x) = \{d \in \mathbb{R}^n \mid d \text{ is a tangent direction at } x\}.$$

Lemma

$\mathcal{T}(x)$ is a closed cone.

Theorem

Suppose $\bar{x} \in \mathcal{F}$ be a local minimizer, $f(x), g_i(x), i = 1, 2, \ldots, m$ be continuously differential at $\bar{x}$. Then

$$\nabla f(\bar{x})^T d \geq 0, \forall d \in \mathcal{T}(\bar{x}),$$

i.e., $\nabla f(\bar{x}) \in \mathcal{T}^*(\bar{x})$.

Constraint qualifications

- KKT theorem is true under a condition: $\mathcal{L}(x) \subseteq \mathrm{cl}(\mathrm{conv}(\mathcal{D}(x)))$.
- Set of interior directions

$$\mathcal{L}^0(x) = \left\{ d \in \mathbb{R}^n \mid \nabla g_i(x)^T d < 0, \forall i \in \mathcal{I}(x) \right\}.$$

- Attainable directions at x: $d = \lim_{\delta \to 0} \frac{r(\delta) - r(0)}{\delta}$ with a $0 < \delta_0$ and a continuous curve $r(\delta)$ such that $r(0) = x$ and $r(\delta) \in \mathcal{F}, \forall 0 \leq \delta \leq \delta_0$
- Set of attainable directions:
  $\mathcal{A}(x) = \{ d \in \mathbb{R}^n \mid d \text{ is an attainable direction at } x \}$.

Theorem

For any feasible solution x of the problem, we have

$$\mathcal{L}^0(x) \subseteq \mathcal{D}(x) \subseteq \mathcal{A}(x) \subseteq \mathcal{T}(x) \subseteq \mathcal{L}(x).$$

Some CQs

- LICQ: $\{\nabla g_i(x), i \in \mathcal{I}(x)\}$ are linearly independent.
- Slater CQ: $g_i(x), i \in \mathcal{I}(x)$ are convex and there exists an x^0 such that $g_i(x^0) < 0,\ i = 1, 2, \ldots, m.$
- Cottle CQ: there exists a d such that $\nabla g_i(x)^T d < 0, \forall i \in \mathcal{I}(x)$.
- Zangwill CQ: $\mathcal{L}(x) \subseteq \mathrm{cl}(\mathcal{D}(x))$.
- Feasible direction CQ: $\mathcal{L}(x) \subseteq \mathrm{cl}(\mathrm{conv}(D(x)))$.
- Kuhn-Tucker CQ: $\mathcal{L}(x) \subseteq \mathrm{cl}(A(x))$.
- Attainable direction CQ: $\mathcal{L}(x) \subseteq \mathrm{cl}(\mathrm{conv}(A(x)))$.
- Abadie CQ: $\mathcal{L}(x) \subseteq \mathcal{T}(x)$.
- Guignard CQ: $\mathcal{L}(x) \subseteq \mathrm{cl}(\mathrm{conv}(\mathcal{T}(x)))$.

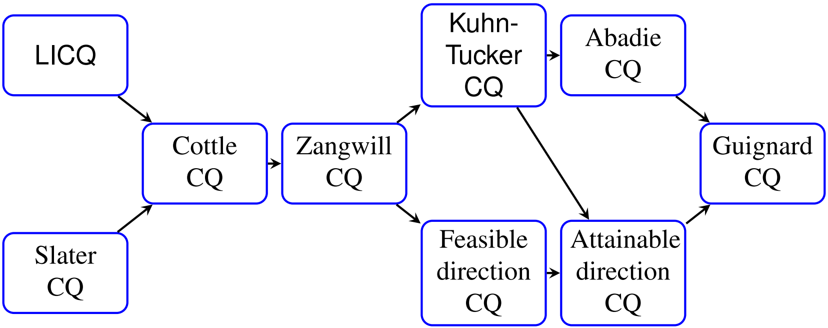

Figure: Relationship charts

Lagrangian duality

- The Lagrangian function. For a given $\lambda \in \mathbb{R}^m_+$,

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

- Saddle points and lower bounds

$$\max_{\lambda \in \mathbb{R}^m_+} L(x, \lambda) = \begin{cases} f(x), & x \in \mathcal{F} \\ +\infty, & x \notin \mathcal{F}. \end{cases}$$

$$v_P = \min_{x \in \mathcal{F}} f(x) = \min_{x \in \mathbb{R}^n} \max_{\lambda \in \mathbb{R}^m_+} L(x, \lambda).$$

$$v_D = \max_{\lambda \in \mathbb{R}^m_+} \min_{x \in \mathbb{R}^n} L(x, \lambda) \leq \min_{x \in \mathbb{R}^n} \max_{\lambda \in \mathbb{R}^m_+} L(x, \lambda).$$

- Lagrangian relaxation. For any given $\lambda \in \mathbb{R}^m_+$

$$v(\lambda) = \min_{x \in \mathbb{R}^n} L(x, \lambda) \leq \min_{x \in \mathcal{F}} L(x, \lambda) \leq \min_{x \in \mathcal{F}} f(x).$$

- Lagrangian dual problem

$$v_D = \max_{\lambda \in \mathbb{R}^m_+} \min_{x \in \mathbb{R}^n} L(x, \lambda) = \max_{\lambda \in \mathbb{R}^m_+} v(\lambda).$$

- Weak duality property. $v_P \geq v_D$.
- A sufficient condition.
  For a given $\bar{\lambda} \geq 0$ and $\bar{x}$ being an optimal solution of the Lagrangian relaxation problem, $\bar{x}$ is an optimal solution of the primal optimization problem if $\bar{\lambda}_i g_i(\bar{x}) = 0, i = 1, 2, \ldots, m$ (complementary condition) and $\bar{x} \in \mathcal{F}$.
- A sufficient condition for a global minimizer.
  Suppose $f(x), g_i(x), i = 1, 2, \ldots, m$ are convex, and $(\bar{x}, \bar{\lambda})$ is a KKT point. Then $\bar{x}$ is a global minimizer.

Example 1: linear programming

- Linear programming.

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \geq b \\ & x \in \mathbb{R}^n_+ \end{array}$$

- Lagrangian multipliers:

$$Ax \geq b \longleftrightarrow \lambda \in \mathbb{R}^m_+,$$

$$x \geq 0 \longleftrightarrow \beta \in \mathbb{R}^n_+.$$

- Lagrangian function

$$L(x, \lambda, \beta) = (c - A^T\lambda - \beta)^T x + \lambda^T b.$$

- Duality

$$
\begin{aligned}
& \max_{\lambda \in \mathbb{R}^m_+, \beta \in \mathbb{R}^n_+} \min_{x \in \mathbb{R}^n} L(x, \lambda, \beta) \\
= & \max_{\lambda \in \mathbb{R}^m_+, \beta \in \mathbb{R}^n_+} \min_{x \in \mathbb{R}^n} \{(c - A^T\lambda - \beta)^T x + \lambda^T b\} \\
= & \max_{\lambda \in \mathbb{R}^m_+, \beta \in \mathbb{R}^n_+} \begin{cases} \lambda^T b, & c - A^T\lambda - \beta = 0 \\ -\infty, & \text{otherwise.} \end{cases} \\
= & \max_{\lambda \in \mathbb{R}^m_+} \begin{cases} \lambda^T b, & c - A^T\lambda \geq 0 \\ -\infty, & \text{otherwise.} \end{cases} \\
= & \max_{\{\lambda \in \mathbb{R}^m_+ | A^T\lambda \leq c\}} b^T\lambda.
\end{aligned}
$$

- Dual problem.

$$
\begin{aligned}
\max \quad & b^T\lambda \\
\text{s.t.} \quad & A^T\lambda \leq c \\
& \lambda \in \mathbb{R}^m_+.
\end{aligned}
$$

Example 2: 1-QCQP

- 1-QCQP

$$\begin{array}{ll} \min & \frac{1}{2}x^T Ax \\ \text{s.t.} & \frac{1}{2}x^T Bx \leq 1 \\ & x \in \mathbb{R}^n, \end{array}$$

- Lagrangian multipliers.

$$\frac{1}{2}x^T Bx \leq 1 \longleftrightarrow \sigma.$$

- Lagrangian function

$$L(x, \sigma) = \frac{1}{2}x^T (A + \sigma B)x - \sigma.$$

- Duality

$$
\begin{aligned}
& \max_{\sigma \geq 0} \min_{x \in \mathbb{R}^n} L(x, \sigma) \\
= \quad & \max_{\sigma \geq 0} \begin{cases} -\sigma, & A + \sigma B \in \mathcal{S}_+^n \\ -\infty, & A + \sigma B \notin \mathcal{S}_+^n. \end{cases} \\
= \quad & \max_{\{\sigma \geq 0 \mid A + \sigma B \in \mathcal{S}_+^n\}} -\sigma.
\end{aligned}
$$

- Dual problem

$$
\begin{aligned}
\max \quad & -\sigma \\
\text{s.t.} \quad & A + \sigma B \in \mathcal{S}_+^n \\
& \sigma \geq 0.
\end{aligned}
$$

Extended Lagrangian duality

- A specified set $\mathcal{G}$: $\mathcal{F} \subseteq \mathcal{G}$.
- The extended Lagrangian function

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), x \in \mathcal{G}.$$

- Lagrangian relaxation problem

$$v(\lambda, \mathcal{G}) = \min_{x \in \mathcal{G}} L(x, \lambda),$$

- Lagrangian dual problem

$$v_d(\mathcal{G}) = \max_{\lambda \in \mathbb{R}^m_+} v(\lambda, \mathcal{G}).$$

Theorem

The extended Lagrangian dual problem has

(i) *Duality.* $v_p \geq v_d(\mathcal{G}), \forall \mathcal{F} \subseteq \mathcal{G}$.

(ii) *Approximation. Suppose* $\mathcal{F} \subseteq \mathcal{G}_1 \subseteq \mathcal{G}_2$. *We have* $v_p \geq v_d(\mathcal{G}_1) \geq v_d(\mathcal{G}_2)$.

(iii) *Strong duality. Suppose* $\mathcal{F} = \mathcal{G}$. *We have* $v_p = v_d(\mathcal{G})$.

Theorem

For a given $\bar{\lambda} \geq 0$, *suppose* $\bar{x}$ *be an optimal solution of the extended Lagrangian relaxation problem and* $(\bar{x}, \bar{\lambda})$ *satisfy the complementary slackness* $\bar{\lambda}_i g_i(\bar{x}) = 0, i = 1, 2, \ldots, m$. *Then* $\bar{x}$ *is a global optimal solution of the primal problem if* $\bar{x} \in \mathcal{F}$.

- A semi-infinite programming problem

$$\begin{array}{rl} \max & \sigma \\ \text{s.t.} & L(x, \lambda) \geq \sigma, \forall x \in \mathcal{G} \\ & \lambda \in \mathbb{R}^m_+, \sigma \in \mathbb{R}. \end{array}$$

An example: linear programming

- Lagrange function $L(x, \lambda, \beta) = (c - A^T\lambda)^T x + \lambda^T b, x \in \mathcal{G}$.
- Dual

$$
\begin{aligned}
& \max_{\lambda \in \mathbb{R}^m_+} \min_{x \in \mathcal{G}} L(x, \lambda) \\
= & \max_{\lambda \in \mathbb{R}^m_+} \min_{x \in \mathbb{R}^n_+} \{(c - A^T\lambda)^T x + \lambda^T b\} \\
= & \max_{\lambda \in \mathbb{R}^m_+} \begin{cases} \lambda^T b, & c - A^T\lambda \geq 0 \\ -\infty, & \text{otherwise.} \end{cases} \\
= & \max_{\{\lambda \in \mathbb{R}^m_+ | A^T\lambda \leq c\}} b^T\lambda.
\end{aligned}
$$

- Dual problem

$$
\begin{aligned}
\max \quad & b^T\lambda \\
\text{s.t.} \quad & A^T\lambda \leq c \\
& \lambda \in \mathbb{R}^m_+,
\end{aligned}
$$

QCQP and its dual

- QCQP

$$\begin{array}{rl} \min & f(x) = \frac{1}{2}x^T Q_0 x + (q^0)^T x + c_0 \\ \text{s.t.} & g_i(x) = \frac{1}{2}x^T Q_i x + (q^i)^T x + c_i \leq 0, i = 1, 2, \ldots, m \\ & x \in \mathbb{R}^n, \end{array}$$

- Feasible set

$$\mathcal{F} = \left\{ x \in \mathbb{R}^n \mid g_i(x) = \frac{1}{2}x^T Q_i x + (q^i)^T x + c_i \leq 0, i = 1, 2, \ldots, m \right\}.$$

- Extended set $\mathcal{G} \supseteq \mathcal{F}$.
- Extended Lagrange function

$$L(x, \lambda) = \frac{1}{2}x^T (Q_0 + \sum_{i=1}^m \lambda_i Q_i) x + (q^0 + \sum_{i=1}^m \lambda_i q^i)^T x + c_0 + \sum_{i=1}^m \lambda_i c_i, x \in \mathcal{G}.$$

- Extended Lagrange dual problem

$$\begin{array}{ll} \max & \sigma \\ \text{s.t.} & \begin{pmatrix} 1 \\ x \end{pmatrix}^T U \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0, \ \forall x \in \mathcal{G} \\ & \sigma \in \mathbb{R}, \ \lambda \in \mathbb{R}^m_+, \end{array}$$

where

$$U = \begin{pmatrix} -2(\sigma - c_0 - \sum_{i=1}^m \lambda_i c_i) & (q_0 + \sum_{i=1}^m \lambda_i q^i)^T \\ q_0 + \sum_{i=1}^m \lambda_i q^i & Q_0 + \sum_{i=1}^m \lambda_i Q_i \end{pmatrix}.$$

- An equivalent form: linear conic programming problem

$$\begin{array}{ll} \max & \sigma \\ \text{s.t.} & \begin{pmatrix} -2\sigma + 2c_0 + 2\sum_{i=1}^m \lambda_i c_i & (q^0 + \sum_{i=1}^m \lambda_i q^i)^T \\ q^0 + \sum_{i=1}^m \lambda_i q^i & Q_0 + \sum_{i=1}^m \lambda_i Q_i \end{pmatrix} \in \mathcal{D}_{\mathcal{G}} \\ & \sigma \in \mathbb{R}, \ \lambda \in \mathbb{R}^m_+, \end{array}$$

where

$$\mathcal{D}_{\mathcal{G}} = \left\{ U \in \mathcal{S}^{n+1} \mid \begin{pmatrix} 1 \\ x \end{pmatrix}^T U \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0, \forall x \in \mathcal{G} \right\}.$$

Theorem

(i) *If $\mathcal{G} \supseteq \mathcal{F}$, the optimal value of the extended Lagragian dual problem is a lower bound of the primal rpoblem.*

(ii) *If $\mathcal{G} = \mathcal{F}$, the optimal values of the extended Lagragian dual problem and the primal rpoblem are the same.*

Conjugate duality theory

Conjugate Program

$$\begin{array}{ll} \inf & f(x) \\ s.t. & x \in \mathcal{X} \cap K \end{array} \qquad \text{(CP)}$$

where $f : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ and K is a cone in $\mathbb{R}^n$.

Conjugate dual

$$\begin{array}{ll} \inf & f^*(y) \\ s.t. & y \in \mathcal{Y} \cap K^* \end{array} \qquad \text{(CD)}$$

where $f^* : \mathcal{Y}$ is the conjugate transform of $f : \mathcal{X}$ and K^* is the dual cone of K.

- feas(*) denotes the feasible domain of problem (*)
- opt(*) denotes the optimal solution set of problem (*)
- v(*) denotes the optimal value of problem (*)

How to get the dual?–LP

$$\begin{array}{rl} \min & c^T x \\ s.t. & Ax = b \\ & x \in \mathbb{R}^n_+ \end{array}$$

$$\mathcal{X} = \{x \in \mathbb{R}^n | Ax = b\}, \mathcal{K} = \mathbb{R}^n_+.$$

$$Ax = b \Leftrightarrow (B, N) \begin{pmatrix} x_B \\ x_N \end{pmatrix} = b \Leftrightarrow x_B = B^{-1}(b - N x_N).$$

$$\begin{aligned} f^*(z) = & \sup_{x \in \mathcal{X}} (x^T z - c^T x) = \sup_{x \in \mathcal{X}} (z - c)_B^T x_B + (z - c)_N^T x_N \\ = & \sup_{x_N \in \mathbb{R}^{n-m}} (z - c)_B^T B^{-1} b + [(z - c)_N - N^T (B^{-1})^T (z - c)_B]^T x_N \\ = & \begin{cases} (z - c)_B^T B^{-1} b, & (z - c)_N - N^T (B^{-1})^T (z - c)_B = 0 \\ +\infty, & otherwise \end{cases} \end{aligned}$$

$$\mathcal{Y} = \{z \in \mathbb{R}^n | (z - c)_N - N^T (B^{-1})^T (z - c)_B = 0\}, \mathcal{K}^* = \mathbb{R}^n_+.$$

$$(z-c)_N - N^T (B^{-1})^T (z-c)_B = 0 \Leftrightarrow \begin{pmatrix} (z - c)_B \\ (z - c)_N \end{pmatrix} - (B, N)^T (B^{-1})^T (z-c)_B = 0.$$

How to get the dual?–LP

Let $w = (B^{-1})^T(z-c)_B$.

$$\begin{array}{ll} \inf & b^T w \\ s.t. & w = (B^{-1})^T(z-c)_B \\ & z - c - A^T w = 0 \\ & z \in \mathbb{R}^n_+, w \in \mathbb{R}^m. \end{array}$$

$w = (B^{-1})^T(z-c)_B$ is redundant. Let $y = -w$.

$$\begin{array}{rl} -\max & b^T y \\ s.t. & A^T y + z = c \\ & z \in \mathbb{R}^n_+, y \in \mathbb{R}^m. \end{array}$$

Note: There is a negative sign in the dual problem.

Conjugate duality theory

Theorem (Conjugate duality theorem/KKT duality theorem)

If $x \in \text{feas}(\text{CP})$ and $y \in \text{feas}(\text{CD})$, then

$$0 \leq x \bullet y \leq f(x) + f^*(y)$$

with the equality holding if and only if

$$x \bullet y = 0 \text{ and } y \in \partial f(x),$$

in which case

$$x \in \text{opt}(\text{CP}) \text{ and } y \in \text{opt}(\text{CD}).$$

Conjugate duality theory

Theorem (Weak duality theorem)

If both CP and CD are feasible, then

(i) $v(\mathrm{CP})$ is finite and

$$v(\mathrm{CP}) + f^*(y) \geq 0, \forall\, y \in \mathrm{feas}(\mathrm{CD});$$

(ii) $v(\mathrm{CD})$ is finite and

$$v(\mathrm{CP}) + v(\mathrm{CD}) \geq 0.$$

Conjugate duality theory

Theorem (Fenchel's theorem/Strong duality theorem)

Suppose that $f : \mathcal{X}$ and K are closed and convex. If $v(\mathrm{CD})$ is finite and one of the following conditions holds:

(i) $\mathrm{ri}(K^*) \cap \mathrm{ri}(\mathcal{Y}) \neq \emptyset$,

(ii) both K^* and $\mathcal{Y}$ are polytopes,

then

$$v(\mathrm{CP}) + v(\mathrm{CD}) = 0 \text{ and } \mathrm{opt}(\mathrm{CP}) \neq \emptyset.$$

Similarly, if $v(\mathrm{CP})$ is finite and one of the following conditions holds:

(i) $\mathrm{ri}(K) \cap \mathrm{ri}(\mathcal{X}) \neq \emptyset$,

(ii) both K and $\mathcal{X}$ are polytopes,

then

$$v(\mathrm{CP}) + v(\mathrm{CD}) = 0 \text{ and } \mathrm{opt}(\mathrm{CD}) \neq \emptyset.$$

Conjugate and Lagrangian duality

- Conjugate models

$$\varphi : x \in \mathbb{R}^n \mapsto (-g_1(x), \ldots, -g_m(x), f(x))^T \in \mathbb{R}^{m+1}.$$

Denote

$$\mathcal{X} = \left\{u \in \mathbb{R}^{m+1} \mid u = \varphi(x), x \in \mathbb{R}^n\right\}$$

and

$$\mathcal{K} = \left\{u \in \mathbb{R}^{m+1} \mid u_i \geq 0, i = 1, 2, \ldots, m\right\}.$$

$$\begin{array}{ll} \min & h(u) = u_{m+1} \\ \text{s.t.} & u \in \mathcal{X} \cap \mathcal{K}. \end{array}$$

- Conjugate function. For any $\lambda \in \mathbb{R}^{m+1}$,

$$h^*(\lambda) = \max_{u\in\mathcal{X}}\{\lambda^T u - u_{m+1}\} = -\min_{x\in\mathbb{R}^n}\left\{(1-\lambda_{m+1})f(x) + \sum_{i=1}^m \lambda_i g_i(x)\right\}.$$

- Well-defined set.

$$\mathcal{Y} = \{\lambda \in \mathbb{R}^{m+1} \mid h^*(\lambda) < +\infty\}.$$

- Dual cone.

$$\begin{aligned}\mathcal{K}^* &= \left\{\lambda \in \mathbb{R}^{m+1} \mid u^T\lambda \geq 0, \forall u \in \mathcal{K}\right\} \\ &= \left\{\lambda \in \mathbb{R}^{m+1} \mid \lambda_i \geq 0, i = 1, 2, \ldots, m; \lambda_{m+1} = 0\right\}.\end{aligned}$$

- Conjugate problem

$$\begin{aligned}&\min_{\lambda\in\mathcal{Y}\cap\mathcal{K}^*} h^*(\lambda) \\ = &\min_{\lambda_i\geq 0, 1\leq i\leq m; \lambda_{m+1}=0}\left[-\min_{x\in\mathbb{R}^n}\{f(x) + \textstyle\sum_{i=1}^m \lambda_i g_i(x)\}\right] \\ = &-\max_{\lambda_i\geq 0, 1\leq i\leq m}\min_{x\in\mathbb{R}^n}\{f(x) + \textstyle\sum_{i=1}^m \lambda_i g_i(x)\}.\end{aligned}$$

Relationships

- One negative sign.
- Strong duality condition—complementary slackness condition. Suppose $u^* \in \mathcal{X} \cap \mathcal{K}$ and $\lambda^* \in \mathcal{Y} \cap \mathcal{K}^*$.

$$u^{*T}\lambda^* = \sum_{i=1}^{m+1} u_i^*\lambda_i^* = \sum_{i=1}^{m} u_i^*\lambda_i^* = 0,$$

$$u_i^*\lambda_i^* = -g_i(x^*)\lambda_i^* = 0, \ \ i = 1, 2, \ldots, m,$$

- Strong duality condition—subgradient condition: $\lambda^* \in \partial u_{m+1}^*$.

$$u_{m+1} \geq u_{m+1}^* + (\lambda^*)^T(u - u^*), \ \forall u \in \mathcal{X}.$$

Denote $u^* = \varphi(x^*)$.

$$f(x) \geq f(x^*) - \sum_{i=1}^{m} \lambda_i^* (g_i(x) - g_i(x^*)),\ \forall x \in \mathbb{R}^n,$$

$$f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*),\ \forall x \in \mathbb{R}^n.$$

Theorem

If there exists (λ^*, x^*) such that x^* is a feasible solution of the primal problem and $\lambda^* \in \mathbb{R}_+^m$, then x^* and λ^* are the optimal solution of the primal and the Lagrangian dual if and only if

$$f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*),\ \forall x \in \mathbb{R}^n,$$

and

$$\lambda_i^* g_i(x^*) = 0, i = 1, 2, \ldots, m.$$

Deriving LCoD from LCoP: standard form

LCoP: standard form

$$
\begin{aligned}
\min \quad & c \bullet x \\
s.t. \quad & a^i \bullet x = b_i, i = 1, \ldots, m \qquad\qquad \text{(LCoP)} \\
& x \in K
\end{aligned}
$$

Deriving LCoD in the framework of conjugate program.

Deriving LCoD from LCoP

LCoP as CP

Variables: $u^T = (u_0, u_1, \ldots, u_m) \in \mathbb{R}^{m+1}$;

$f(u) = u_0$;

$\mathcal{X} = \{u \in \mathbb{R}^{m+1} | u_i = b_i, i = 1, \ldots, m\}$;

$K_0 = \{u \in \mathbb{R}^{m+1} | u_0 = c \bullet x, u_i = a^i \bullet x, x \in K, i = 1, \ldots, m\}$.

$$\begin{array}{ll} \inf & f(u) \\ s.t. & u \in \mathcal{X} \cap K_0 \end{array}$$

Deriving LCoD from LCoP

Corresponding CD

Variables: $v^T = (v_0, v_1, \ldots, v_m) \in \mathbb{R}^{m+1}$;

$$\begin{aligned} f^*(v) &= \sup_{u \in \mathcal{X}} \{u \bullet v - f(u)\} < +\infty \\ &= \sup_{u_0 \in \mathbb{R}} \{(v_0 - 1)u_0 + \sum_{i=1}^{m} b_i v_i\} \end{aligned}$$

Hence

$f^*(v) = \sum_{i=1}^m b_i v_i$;

$\mathcal{Y} = \{v \in \mathbb{R}^{m+1} | v_0 = 1\}$;

Deriving LCoD from LCoP

Corresponding CD

Moreover,

$$\begin{aligned} K_0^* &= \{v \in \mathbb{R}^{m+1} | v \bullet u \geq 0, \forall u \in K_0\} \\ &= \{v \in \mathbb{R}^{m+1} | (v_0 c + \textstyle\sum_{i=1}^m v_i a^i) \bullet x \geq 0, \forall x \in K\} \\ &= \{v \in \mathbb{R}^{m+1} | v_0 c + \textstyle\sum_{i=1}^m v_i a^i \in K^*\}. \end{aligned}$$

$$\mathcal{Y} \cap K_0^* = \{v \in \mathbb{R}^{m+1} | c + \sum_{i=1}^m v_i a^i = s, s \in K^*\}.$$

$$\begin{array}{ll} \inf & \sum_{i=1}^m b_i v_i \\ s.t. & c + \sum_{i=1}^m v_i a^i = s \\ & s \in K^* \end{array}$$

Deriving LCoD from LCoP

CD to LCoD

Define variables: $y = -(v_1, \ldots, v_m)^T$, we have

$$\begin{array}{ll} \max & b^T y \\ s.t. & \sum_{i=1}^m y_i a^i + s = c \\ & s \in K^*, \; y \in \mathbb{R}^m \end{array} \qquad \text{(LCoD)}$$

Therefore, the duality theorems of conjugate programs may apply to LCoP.

Conic duality theorems for LCoP

Theorem (Weak duality theorem)

If both LCoP and LCoD are feasible, then

$$c \bullet x \geq b^T y, \forall x \in \text{feas}(\text{LCoP}),\ (y, s) \in \text{feas}(\text{LCoD}).$$

Theorem (Strong duality theorem)

(i) If $\text{feas}(\text{LCoP}) \cap \text{int}(K) \neq \emptyset$ and $v(\text{LCoP})$ is finite, then there exists $(y^*, s^*) \in \text{feas}(\text{LCoD})$ such that $b^T y^* = v(\text{LCoP})$.

(ii) If $\text{feas}(\text{LCoD}) \cap \text{int}(K^*) \neq \emptyset$ and $v(\text{LCoD})$ is finite, then there exists $x^* \in \text{feas}(\text{LCoP})$ such that $c \bullet x = v(\text{LCoD})$.

Conic duality theorems for LCoP

Theorem (KKT duality theorem)

If $\mathrm{feas(LCoP)}$ and $\mathrm{feas(LCoD)}$ are both nonempty and $\mathrm{feas(LCoP)} \cap \mathrm{int}(K) \neq \emptyset$, then x^* is optimal for LCoP if and only if the following conditions hold:

(i) $x^* \in \mathrm{feas(LCoP)}$;

(ii) There exists $(y^*, s^*) \in \mathrm{feas(LCoD)}$;

(iii) $c \bullet x^* = b^T y^*$ (or equivalently $x^* \bullet s^* = c \bullet x^* - b^T y^* = 0$).

Inequality models

- Inequality models

$$
\begin{array}{rl}
\min & c \bullet x \\
\text{s.t.} & a^i \bullet x \geq b_i, i = 1, \ldots, m \\
& x \in \mathcal{K}.
\end{array}
$$

- Dual problems

$$
\begin{array}{rl}
\max & b^T y \\
\text{s.t.} & \sum_{i=1}^m y_i a^i + s = c \\
& s \in \mathcal{K}^*, y \in \mathbb{R}_+^m.
\end{array}
$$

- Weak duality theorem: the same to the standard form.
- Strong duality theorem

Theorem

If there exists $x^0 \in \mathbb{E}$ satisfying $a^i \bullet x^0 > b_i,\ i = 1, 2, \ldots, m$, $x^0 \in \mathrm{ri}(\mathcal{K})$ and the inequality model is bounded below, then there exists an optimal solution (y^*, s^*) of its dual problem satisfying $b^T y^*$ reaches its minimal optimal value of the inequality problem.
Symmetrically, if there exists $s^0 \in \mathrm{ri}(\mathcal{K}^*)$ and $y^0 \in \mathbb{R}^m_{++}$ satisfying $\sum_{i=1}^m y_i^0 a^i + s^0 = c$ and the dual problem is upper-bounded, then there exists an x^* with $c \bullet x^*$ reaches the optimal value of the dual problem.

Theorem

Suppose the $\mathcal{K}$ be closed convex cone. Then the dual of (LCD) is (LCP).