IE 521 Convex Optimization

Lecture 13: Conic Duality

Niao He

16th April 2019

Outline

What's the dual to a conic program?

Recall the LP duality:

$$(LP) \quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \geq b \end{array} \qquad\qquad (LD) \quad \begin{array}{ll} \max & b^T y \\ \text{s.t.} & A^T y = c \\ & y \geq 0 \end{array}$$

Now consider the conic program

$$(CP) \quad \begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \geq_{\mathcal{K}} b \end{array} \qquad\qquad (CD) \quad ?$$

Q. Now that $Ax \geq_{\mathcal{K}} b \Rightarrow y^T(Ax) \geq y^T b$ for which y?

Dual Cone

Definition. The dual cone of a nonempty cone $\mathcal{K}$ is

$$\mathcal{K}_* = \left\{ y : y^T x \geq 0, \forall x \in \mathcal{K} \right\}$$

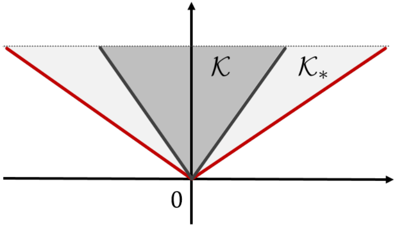

Figure: Dual Cone

Remark. Dual cone is always a closed cone.

Properties of Dual Cone

Proposition. Let $\mathcal{K}$ be a closed cone and $\mathcal{K}_*$ be its dual.

(a) $(\mathcal{K}_*)_* = \mathcal{K}$

(b) $\mathcal{K}$ is pointed iff $\mathcal{K}_*$ has non-empty interior

(c) $\mathcal{K}$ is a regular cone iff $\mathcal{K}_*$ is a regular cone

Proof: Self-exercise.

Self-dual Cone

Definition. If $\mathcal{K} = \mathcal{K}_*$, we call it a self-dual cone.

Remark. Nonnegative orthant, second order cone, and positive semidefinite cone are all self-dual:

- $(\mathbb{R}^m_+)_* = \mathbb{R}^m_+$
- $(L^n)_* = L^n$
- $(S^n_+)_* = S^n_+$

Self-dual Cone

Proposition. L^n is self-dual, i.e. $(L^n)^* = L^n$.

Proof

(i) $L^n \subseteq (L^n)^*$: Suppose $y \in L^n$, we show that $\forall x \in L^n$,

$$y^T x = y_1 x_1 + \ldots + y_n x_n \geq -\sqrt{\sum_{i=1}^{n-1} y_i^2}\sqrt{\sum_{i=1}^{n-1} x_i^2} + y_n x_n \geq 0$$

due to Cauchy-Schwarz inequality.

(ii) $(L^n)^* \subseteq L^n$: Suppose $y \in (L^n)^*$, we have $y^T x \geq 0, \forall x \in L^n$
If $(y_1, ..., y_{n-1}) = 0$, let $x = [0, ..., 0, 1] \in L^n$, we get

$$y^T x = y_n \geq 0, \Rightarrow y \in L^n.$$

Otherwise, let $x = [-y_1, ..., -y_{n-1}, \sqrt{\sum_{i=1}^{n-1} y_i^2}] \in L^n$,

$$y^T x = -\sum_{i=1}^{n-1} y_i^2 + y_n \sqrt{\sum_{i=1}^{n-1} y_i^2} \geq 0 \Rightarrow y \in L^n.$$

Dual of Conic Program

Primal Conic Program:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq_{\mathcal{K}} b \end{aligned} \tag{CP}$$

Dual Conic Program:

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y = c \\ & y \geq_{\mathcal{K}_*} 0 \end{aligned} \tag{CD}$$

Conic Duality

Theorem.

- **(Weak Conic Duality)**: $\text{Opt}(CD) \leq \text{Opt}(CP)$
- **(Strong Conic Duality)**: If (CP) is bounded below and strictly feasible, i.e.,

$$\exists x_0, \text{ s.t. } Ax_0 >_{\mathcal{K}} b,$$

then (CD) is solvable and $\text{Opt}(CD) = \text{Opt}(CP)$.

Corollary. If (CD) is bounded above and strictly feasible,

$$\text{i.e. } \exists y >_{\mathcal{K}_*} 0, \text{ s.t. } A^T y = c$$

then (CP) is solvable and $\text{Opt}(CD) = \text{Opt}(CP)$.

Proof of Conic Duality

Denote $p^* = \text{Opt}(CP)$.
Sufficient to show that $\exists y^*$ feasible to (CD), s.t., $b^T y^* \geq p^*$.
When $c = 0$, simply set $y^* = 0$. Now consider $c \neq 0$. Define

$$M = \left\{Ax - b : c^T x \leq p^*\right\}.$$

- $M \cap \text{int}(\mathcal{K}) = \emptyset$ (why?)
- By seperation theorem, $\exists y \neq 0$, s.t.

$$\sup_{z \in M} y^T z \leq \inf_{z \in \text{int}(\mathcal{K})} y^T z$$

- It must hold that $y \in \mathcal{K}_*$ and $\sup_{x: c^T x \leq p^*} y^T (Ax - b) \leq 0$.
- Hence, $\lambda c = A^T y$ for some $\lambda \geq 0$.
- By strictly feasibility of (CP), we further hanve $\lambda > 0$ (why?).
- Setting $y^* = \frac{y}{\lambda}$, we have $y^* \in \mathcal{K}_*, A^T y^* = c$ and $p^* \leq b^T y^*$.

Optimality Conditions

Theorem. Suppose *at least one* of (CP) and (CD) is bounded and strictly feasible, then the feasible primal-dual pair (x^*, y^*) is a pair of optimal primal-dual solutions iff

- (Zero duality gap) $c^T x^* - b^T y^* = 0$
- (Complementary slackness) $(Ax^* - b)^T y^* = 0$

Observe that

$$\begin{aligned} c^T x^* - b^T y^* = & \underbrace{c^T x^* - \mathrm{Opt}(CP)}_{\geq 0} \\ & + \underbrace{\mathrm{Opt}(CD) - b^T y^*}_{\geq 0} \\ & + \underbrace{\mathrm{Opt}(CP) - \mathrm{Opt}(CD)}_{\geq 0} \end{aligned}$$

Discussions

- In the case of LP, strict feasibility is not required for strong duality nor solvability of the program.
- In general case of CP, strict feasibility is required.

Example

A conic problem can be strictly feasible and bounded, but NOT solvable.

$$\begin{aligned} \min_{x_1,x_2} \quad & x_1 \\ \text{s.t.} \quad & \begin{bmatrix} x_1 - x_2 \\ 1 \\ x_1 + x_2 \end{bmatrix} \geq_{L^3} 0 \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \min_{x_1,x_2} \quad & x_1 \\ \text{s.t.} \quad & 4x_1x_2 \geq 1 \\ & x_1 + x_2 > 0 \end{aligned}$$

Example

A conic problem can be solvable yet not strictly feasible, and the dual is infeasible.

$$\begin{aligned} \min_{x_1,x_2} \quad & x_2 \\ \text{s.t.} \quad & \begin{bmatrix} x_1 \\ x_2 \\ x_1 \end{bmatrix} \geq_{L^3} 0 \end{aligned} \qquad \Longleftrightarrow \qquad \begin{aligned} \max_{\lambda} \quad & 0 \\ \text{s.t.} \quad & \begin{bmatrix} \lambda_1 + \lambda_3 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ & \lambda \geq_{L^3} 0 \end{aligned}$$

SOCP Duality

Primal SOCP:

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.} \quad & \|A_i x - b_i\|_2 \leq d_i^T x - e_i, i = 1, ..., m \end{aligned} \qquad \text{(SOCP–P)}$$

Dual SOCP:

$$\begin{aligned} \max_{\substack{\lambda \in \mathbb{R}^m \\ u_i \in \mathbb{R}^{n_i - 1}, i = 1, ..., m}} \quad & \sum_{i=1}^m b_i^T u_i + e^T \lambda \\ \text{s.t.} \quad & \sum_{i=1}^m (A_i^T u_i + d_i \lambda_i) = c \\ & \|u_i\|_2 \leq \lambda_i, \qquad i = 1, ..., m \end{aligned} \qquad \text{(SOCP–D)}$$

SDP Duality

Primal SDP:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \mathcal{A}x - B = \sum_{i=1}^{n} x_i A_i - B \succeq 0 \end{aligned} \tag{SDP–P}$$

Dual SDP:

$$\begin{aligned} \max_{Y} \quad & \operatorname{tr}(BY) \\ \text{s.t.} \quad & \operatorname{tr}(A_i Y) = c_i \quad i = 1, ..., n \\ & Y \succeq 0 \end{aligned} \tag{SDP–D}$$

SDP Optimality Conditions

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \sum_{i=1}^n x_i A_i - B \succeq 0 \end{aligned} \qquad \begin{aligned} \max_Y \quad & \operatorname{tr}(BY) \\ \text{s.t.} \quad & \operatorname{tr}(A_i Y) = c_i, i = 1, ..., n \\ & Y \succeq 0 \end{aligned}$$

Remark. (x^*, Y^*) is optimal primal-dual pair iff

1. $\sum_{i=1}^n x_i^* A_i \succeq B$ (primal feasibility)
2. $Y^* \succcurlyeq 0, \operatorname{tr}(A_i Y^*) = c_i, i = 1, ..., m$ (dual feasibility)
3. $Y^* (\sum_{i=1}^n x_i^* A_i - B) = 0$ (complementary slackness)

Application of SDP Duality

Example . Use SDP duality to show that for any $B \in S_+^n$:

$$\lambda_{max}(B) = \max_{x \in \mathbb{R}^n} \left\{ x^T B x : \|x\|_2 = 1 \right\}$$

$$\begin{array}{lll} \max_{x} & \mathrm{tr}(Bxx^T) & \\ \text{s.t.} & \mathrm{tr}(xx^T) = 1 & \quad (P) \end{array}$$

$$\begin{array}{lll} \max_{X} & \mathrm{tr}(BX) & \\ \text{s.t.} & \mathrm{tr}(X) = 1 & \quad (P') \\ & X \succcurlyeq 0 & \end{array}$$

$$\Updownarrow$$

$$\begin{array}{lll} \min_{x} & \lambda & \\ \text{s.t.} & \lambda I - B \succeq 0 & \quad (D) \end{array}$$

(P)=(P'), why?

SDP Relaxation of Nonconvex QCQP

Quadratic constrained quadratic programming:

$$\begin{aligned} \min\ & x^T Q_0 x + 2q_0^T x + c_0 \\ \text{s.t. } & x_i^T Q_i x_i + 2q_i^T x + c_i \leq 0,\ 1 \leq i \leq m \end{aligned} \qquad \text{(QCQP)}$$

Rank-1 reformulation:

$$\begin{aligned} \min_{x,X}\quad & \mathrm{tr}(A_0 X) \\ \text{s.t.}\quad & \mathrm{tr}(A_i X) \leq 0,\ 1 \leq i \leq m \\ & X = \begin{bmatrix} xx^T & x \\ x^T & 1 \end{bmatrix} \end{aligned} \qquad \text{(QCQP')}$$

Here $A_i = \begin{bmatrix} Q_i & q_i \\ q_i^T & c_i \end{bmatrix}, i = 0, 1, ..., m$

SDP Relaxation of Nonconvex QCQP

SDP relaxation:

$$\begin{aligned} \min_{X} \quad & \operatorname{tr}(A_0 X) \\ \text{s.t.} \quad & \operatorname{tr}(A_i X) \leq 0,\ 1 \leq i \leq m \\ & X \succeq 0 \\ & X_{n+1,n+1} = 1 \end{aligned} \tag{SDP-r}$$

Dual of SDP relaxation:

$$\begin{aligned} \max_{\lambda \geq 0, t} \quad & t \\ \text{s.t.} \quad & A_0 + \sum \lambda_i A_i - \begin{bmatrix} 0 & 0 \\ 0 & t \end{bmatrix} \succeq 0 \end{aligned} \tag{SDP-d}$$

Remark. Opt(SDP-d) $\leq$ Opt(SDP-r) $\leq$ Opt(QCQP)

References

- Ben-Tal & Nemirovski (2013), Chapters 1 -3