IE 521: Convex Optimization **Spring 2017, UIUC**

Lecture 15: Applications of Conic Duality – March 13

Instructor: Niao He *Scribe: Shuanglong Wang*

Courtesy warning: These notes do not necessarily cover everything discussed in the class. Please email TA (swang157@illinois.edu) if you find any typos or mistakes.

In this lecture, we cover the following topics

- SOCP Duality
- SDP Duality
- SDP Relaxations for Non-Convex Quadratic Programs

References: Ben-Tal & Nemirovski. *Lectures on Modern Convex Optimization*, Chapter 2.1, 3.1

15.1 Recall: Conic Duality

Let K be a regular cone, and $K_* = \left\{y : y^T x \geq 0, \forall x \in K\right\}$ be its dual cone

$$p^* = \min_x \quad c^T x \qquad \text{s.t.} \quad Ax \geq_K b \qquad \text{(CP)}$$

$$d^* = \max_y \quad b^T y \qquad \text{s.t.} \quad A^T y = c, \quad y \geq_{K_*} 0 \qquad \text{(CD)}$$

- Weak Duality: $p^* \geq d^*$
- Strong Duality: If one of (CP), (CD) is bounded and strictly feasible, then $p^* = d^*$.
- Optimality Condition: (x^*, y^*) is optimal iff
  (i) primal feasibility: $Ax^* \geq_K b$
  (ii) dual feasibility: $A^T y^* = c, y^* \geq_{K_*} 0$
  (iii) zero-duality gap: $c^T x^* - b^T y^* = 0$

15.2 SOCP Duality

Second-order Cone Program (SOCP): when K is the Cartesian product of Lorentz cones, namely, $K = L^{n_1} \times ... \times L^{n_m}$.
The general form of a second-order cone program is

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

where $c \in \mathbf{R}^n, A_i \in \mathbf{R}^{(n_i-1)\times n}, \mathbf{R}_i \in \mathbf{R}^{(n_i-1)\times 1}, e_i \in \mathbf{R}$.

The conic form is

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \tilde{A}_i x \geq_{L^{n_i}} \tilde{b}_i, \quad i = 1, ..., m \end{aligned}$$

where $\tilde{A}_1 = \begin{bmatrix} A_i \\ d_i^T \end{bmatrix}, \tilde{b_i} = \begin{bmatrix} b_i \\ e_i \end{bmatrix}$

Proposition 15.1 *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 $\forall x \in L^n$,

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

where the first inequality is due to Cauchy-Schwarz inequality and the second inequality is because $y_n \geq \sqrt{\sum_{i=1}^{n-1} y_i^2}$ and $x_n \geq \sqrt{\sum_{i=1}^{n-1} x_i^2}$.

(ii) $(L^n)^* \subseteq L^n$
Suppose $y \in (L^n)^*$, we have $y^T x \geq 0, \forall x \in L^n$
If $(y_1, ..., y_{n-1}) = 0$, setting $x = [0, ..., 0, 1] \in L^n$, we get $y^T x = y_n \geq 0$, so $y \in L^n$
If $(y_1, ..., y_{n-1}) \neq 0$, setting $x = [-y_1, ..., -y_{n-1}, \sqrt{\sum_{i=1}^{n-1} y_i^2}] \in L^n$, we get

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

■

Hence, the dual of the SOCP is

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

Remark: The same dual can be derived from Lagrange duality. The Lagrange function is

$$L(x, \lambda) = c^T x + \sum_{i=1}^m \lambda_i \Big(\| A_i x - b_i \|_2 - (d_i^T x - e_i) \Big)$$

The Lagrange dual is given by

$$\begin{aligned} & \max_{\lambda \geq 0} \min_x L(x, \lambda) \\ = & \max_{\lambda \geq 0} \min_x \max_{\|u_i\|_2 \leq \lambda_i, i=1,...,m} c^T x - \sum_{i=1}^m u_i^T (A_i x_i - b_i) - \sum_{i=1}^m \lambda_i (d_i^T x - e_i) \\ = & \max_{\lambda \geq 0, \|u_i\|_2 \leq \lambda_i} \min_x (c - \sum_{i=1}^m (A_i^T u_i + d_i \lambda_i))^T x + \sum_{i=1}^m b_i^T u_i + e^T \lambda \\ = & \max_{\lambda, u_1, ..., u_m} \quad \sum_{i=1}^m b_i^T u_i + e^T \lambda \\ & \text{s.t.} \quad \sum_{i=1}^m (A_i^T u_i + d_i \lambda_i) = c \\ & \qquad \| u_i \|_2 \leq \lambda_i, i = 1, ..., m \end{aligned}$$

which is the same as dual derived from conic duality.

15.3 SDP Duality

Semidefinite Program (SDP): when K is the positive semidefinite cone.
The general form of a semidefinite program is

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

The constraint type $x_1 A_1 + ... + x_n A_n - B \succcurlyeq 0$ is called Linear Matrix Inequality.

Proposition 15.2 *S_+^n is self-dual, i.e. $(S_+^n)^* = S_+^n$*

Proof:

(i) $S_+^n \subseteq (S_+^n)^*$
Suppose $Y \succcurlyeq 0$. we have $Y = U\Lambda U^T = \sum_{i=1}^n \lambda_i u_i u_i^T$, where $\lambda \geq 0, i = 1, ..., m$

$$\forall X \succcurlyeq 0, \langle X, Y\rangle = \mathrm{tr}(X^TY) = \mathrm{tr}(XY) = \mathrm{tr}(\sum_{i=1}^n \lambda_i u_i u_i^T Y) = \sum_{i=1}^n \lambda_i \mathrm{tr}(u_i^T Y u_i) \geq 0$$

Hence, $Y \in (S_+^n)^*$.

(ii) $(S_+^n)^* \subseteq S_+^n$
Suppose $Y \in (S_+^n)^*$, i.e. $\mathrm{tr}(XY) \geq 0, \forall x \in S_+^n$. For any $x \in \mathbf{R}^n$, let $X = xx^T$, we have $\mathrm{tr}(XY) = \mathrm{tr}(xx^TY) = x^TYx \geq 0$. Hence, $Y \in S_+^n$.

■

The dual of SDP is:

$$\begin{aligned} \max_Y \quad & \mathrm{tr}(BY) \\ \text{s.t.} \quad & \mathrm{tr}(A_iY) = c_i \quad i = 1, ..., n \\ & Y \succcurlyeq 0 \end{aligned} \tag{SDP–D}$$

Remark: Based on conic duality theorem (x^*, y^*) is optimal primal-dual pair iff

1. $\sum_{i=1}^n x_i A_i \succcurlyeq B$ (primal feasibility)
2. $Y \succcurlyeq 0, \mathrm{tr}(A_iY) = c_i, i = 1, ..., m$ (dual feasibility)
3. $\mathrm{tr}(Y(\sum_{i=1}^n x_iA_i - B)) = 0$, i.e. $Y(\sum_{i=1}^n x_iA_i - B) = 0$

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

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

Note that the right hand side can be rewritten as

$$\begin{aligned} \max_x \quad & \mathrm{tr}(Bxx^T) \\ \text{s.t.} \quad & \mathrm{tr}(xx^T) = 1 \end{aligned}$$

We show this is equivalent to solve the SDP relaxation.

$$\begin{aligned} \max_X \quad & \mathrm{tr}(BX) \\ \text{s.t.} \quad & \mathrm{tr}(X) = 1 \\ & X \succcurlyeq 0 \end{aligned} \tag{P}$$

Let $p = \max_x \left\{ \mathrm{tr}(Bxx^T) : \mathrm{tr}(xx^T) = 1 \right\}$ and $\tilde{p} = \max_X \{\mathrm{tr}(BX) : \mathrm{tr}(X) = 1, X \succcurlyeq 0\}$. Since the latter is a relaxation, so $\tilde{p} \geq p$. On the other hand, for any $X \succcurlyeq 0$ with $\mathrm{tr}(X) = 1$,

$$X = \sum_{i=1}^{n} \lambda_i x_i x_i^T, \text{ with } \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, i = 1, ..., n, \text{ and } \parallel x_i \parallel_2 = 1, \text{i.e. } \mathrm{tr}(x_i x_i^T) = 1$$

$$\mathrm{tr}(BX) = \mathrm{tr}(\sum_{i=1}^{n} \lambda_i x_i x_i^T B) = \sum_{i=1}^{n} \lambda_i \mathrm{tr}(Bx_i x_i^T) \leq \max_{i=1,..,n} \mathrm{tr}(Bx_i x_i^T) \leq p$$

Hence, $\tilde{p} \leq p$. Therefore, $\tilde{p} = p$.

By the SDP duality, the dual of (P) is

$$\begin{aligned} \min_x \quad & \lambda \\ \text{s.t.} \quad & \lambda I - B \succcurlyeq 0 \end{aligned} \tag{D}$$

Note that $\lambda I - B \succcurlyeq 0$ is equivalent to $\lambda \geq \lambda_{max}(B)$.

Since (P) is strictly feasible, $X = n^{-1} I$ satisfies $\mathrm{tr}(X) = 1$ and $X \succ 0$, strong duality holds. Therefore,

$$\max_x \left\{ x^T B x : \parallel x \parallel_2 = 1 \right\} = \mathrm{Opt}(P) = \mathrm{Opt}(D) = \lambda_{max}(B)$$

15.4 SDP Relaxations of Non-convex Quadratic Programming

Consider the quadratic constrained quadratic programming:

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

Let $X = \begin{bmatrix} xx^T & x \\ x^T & 1 \end{bmatrix} = \begin{bmatrix} x \\ 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}^T$ and $A_i = \begin{bmatrix} Q_i & q_i \\ q_i^T & c_i \end{bmatrix}, i = 0, 1, ..., m$

We can write (QCQP) as

$$\begin{aligned} \min_{x,X} \quad & \mathrm{tr}(A_0 X) \\ \text{s.t.} \quad & \mathrm{tr}(A_i X) \leq 0, \quad i = 1, ..., m \\ & X = \begin{bmatrix} xx^T & x \\ x^T & 1 \end{bmatrix} \end{aligned}$$

SDP Relaxation. The following SDP is a relaxation of (QCQP)

$$\begin{aligned} \text{Opt}^{\text{SDP}} = \min_X \quad & \text{tr}(A_0 X) \\ \text{s.t.} \quad & \text{tr}(A_i X) \leq 0, \quad i = 1, ..., m \\ & X \succcurlyeq 0 \\ & X_{n+1,n+1} = 1 \end{aligned} \tag{SDP-relaxation}$$

Therefore, $\text{Opt}^{\text{SDP}} \leq \text{Opt}$.

Lagrange Relaxation. Another relaxation of QCQP is based on Lagrange duality. The Lagrange dual is given by

$$\begin{aligned} & \max_{\lambda \geq 0} \inf_x \quad L(x, \lambda) \\ = & \max_{\lambda \geq 0} \inf_x \quad x^T Q_0 x + 2q_0^T x + c_0 + x^T (\textstyle\sum_{i=1}^m \lambda_i Q_i) x + 2(\sum_{i=1}^m \lambda_i q_i)^T x + \sum_{i=1}^m \lambda_i c_i \\ = & \max_{\lambda \geq 0} \{t : \quad x^T (Q_0 + \textstyle\sum_{i=1}^m \lambda_i Q_i) x + 2(q_0 + \sum_{i=1}^m \lambda_i q_i)^T x + (\sum_{i=1}^m \lambda_i c_i + c_0) \geq t, \forall x\} \end{aligned}$$

Note that $\forall x, x^T A x \geq 0$ is equivalent to $A \succcurlyeq 0$. Hence, the Lagrange relaxation is

$$\begin{aligned} \text{Opt}^{\text{Lagrange}} = \max_{\lambda \geq 0} \quad & t \\ \text{s.t.} \quad & \begin{bmatrix} Q_0 + \sum_{i=1}^m \lambda_i Q_i & q_0 + \sum_{i=1}^m \lambda_i q_i \\ (q_0 + \sum_{i=1}^m \lambda_i q_i)^T & \sum_{i=1}^m \lambda_i c_i + c_0 - t \end{bmatrix} \succcurlyeq 0 \end{aligned} \tag{Lagrange relaxation}$$

By weak duality, we know that $\text{Opt}^{\text{Lagrange}} \leq \text{Opt}$.

Indeed, one can show that the Lagrange relaxation is the SDP dual to the SDP relaxation.

- If either of them is strictly feasible, then $\text{Opt}^{\text{Lagrange}} = \text{Opt}^{\text{SDP}}$.
- Otherwise, we always have $\text{Opt}^{\text{Lagrange}} \leq \text{Opt}^{\text{SDP}} \leq \text{Opt}$.

Therefore, SDP relaxation is always tighter than Lagrange relaxation.