IE 521 Convex Optimization

Lecture 14: SDP Relaxation and Applications

Niao He

16th April 2019

Outline

Conic Duality

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:

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

Theorem. (Strong Conic Duality) If (CP) is bounded below and strictly feasible, i.e., $\exists x_0$, s.t. $Ax_0 >_{\mathcal{K}} b$, then (CD) is solvable and $\text{Opt}(CD) = \text{Opt}(CP)$.

Example: LP Duality

Primal LP:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \end{aligned} \tag{LP-P}$$

Dual LP:

$$\begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y = c \\ & y \geq 0 \end{aligned} \tag{LP-D}$$

Example: SOCP Duality

Primal SOCP:

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

Dual SOCP:

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

Example: SDP Duality

Primal SDP:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \sum_{i=1}^n x_i A_i - B \succeq 0 \end{aligned} \qquad \text{(SDP–P)}$$

Dual SDP:

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

Example: Variant of SDP Duality

Primal SDP:

$$\begin{aligned} \min_{Y} \quad & B \bullet Y \\ \text{s.t.} \quad & A_i \bullet Y = c_i \quad i = 1, ..., n \\ & Y \succeq 0 \end{aligned} \qquad \text{(SDP–P')}$$

Dual SDP:

$$\begin{aligned} \max_{x} \quad & c^T x \\ \text{s.t.} \quad & B - \sum_{i=1}^{n} x_i A_i \succeq 0 \end{aligned} \qquad \text{(SDP–D')}$$

Application 1: Maximal Eigenvalue

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

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

- Reformulation with rank-1 constraint:

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

- SDP Relaxation:

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

- Exact Recovery: Opt(SDP-r) $=$ Opt(P)

Application 1: Maximal Eigenvalue

- SDP Relaxation:

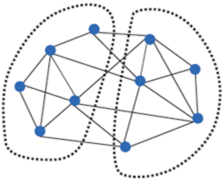
$$\begin{aligned} \max_{X} \quad & \mathrm{tr}(BX) \\ \text{s.t.} \quad & \mathrm{tr}(X) = 1 \\ & X \succeq 0 \end{aligned} \qquad \text{(SDP-r)}$$

- Dual to SDP relaxation:

$$\begin{aligned} \lambda_{\max}(B) = \min_{x} \quad & \lambda \\ \text{s.t.} \quad & \lambda I - B \succeq 0 \end{aligned} \qquad \text{(SDP-d)}$$

Application 2: MAX CUT Problem

Consider undirected weighted graph $\mathcal{G} = (V, E, W)$.
Here $|V| = n, , W = \{w_{ij}\}_{(i,j)\in E}$ with $w_{ij} \geq 0$.



MAX CUT Problem (NP-Hard):

$$\max_{X} \quad \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(1 - x_i x_j)$$
$$\text{s.t.} \quad x_i \in \{-1, 1\}, i = 1, \ldots, n \qquad \text{(MAXCUT)}$$

Application 2: MAX CUT Problem

Reformulation with Rank-1 Constraint:

$$\begin{aligned} \max_x \quad & \frac{1}{4}\sum_{i=1}^n \sum_{j=1}^n w_{ij} - \frac{1}{4}\mathrm{tr}(WX) \\ \text{s.t.} \quad & X_{ii} = 1, i = 1, \ldots, n \\ & X = xx^T \end{aligned} \qquad \text{(MAXCUT')}$$

SDP Relaxation:

$$\begin{aligned} \max_x \quad & \frac{1}{4}\sum_{i=1}^n \sum_{j=1}^n w_{ij} - \frac{1}{4}\mathrm{tr}(WX) \\ \text{s.t.} \quad & X_{ii} = 1, i = 1, \ldots, n \\ & X \succeq 0 \end{aligned} \qquad \text{(SDP-r)}$$

GW Theorem. [Goemans & Williamson, 1995]

$$\text{MAXCUT} \leq \text{Opt(SDP-r)} \leq 1.1383 \cdot \text{MAXCUT}$$

Nesterov's $\frac{\pi}{2}$ Theorem

Consider any QP with $Q \succeq 0$ in the form

$$\begin{aligned} \max_x \quad & x^T Q x \\ \text{s.t.} \quad & x_i \in \{-1, 1\}, i = 1, \ldots, n \end{aligned} \tag{QP}$$

and its SDP relaxation

$$\begin{aligned} \max_X \quad & \mathrm{tr}(QX) \\ \text{s.t.} \quad & X_{ii} = 1, i = 1, \ldots, n \\ & X \succeq 0 \end{aligned} \tag{SDP-r}$$

Nesterov's $\frac{\pi}{2}$ Theorem. [Nesterov, 1998] If $Q \succeq 0$,

$$\mathrm{Opt(QP)} \leq \mathrm{Opt(SDP\text{-}r)} \leq \frac{\pi}{2}\mathrm{Opt(QP)}$$

Nesterov's $\frac{\pi}{2}$ Theorem

Remark. MAXCUT is a special case:

$$Q_{ij} = -w_{ij}, i \neq j, \text{ and } Q_{ii} = \sum_{j=1}^{n} w_{ij}.$$

Proof Sketch:

- Let $\xi \sim \mathcal{N}(0, X^*)$, where X^* is optimal to (SDP-r).
- Let $\zeta = \text{sign}(\xi)$, $\zeta \in \{-1, 1\}^n$.
- $\text{Opt(QP)} \geq \mathbb{E}[\zeta^T Q \zeta] = \text{tr}\left(Q \frac{2}{\pi} \text{arcsin}(X^*)\right)$.
- Note $\text{arcsin}(X^*) \succeq X^*$, where arcsin is inverse of sine.
- Hence, $\text{Opt(QP)} \geq \frac{2}{\pi}\text{Opt(SDP-r)}$

Q. What about when Q is indefinite?

(Nemirovski, Roos, Terlaky, 1998)

$$\text{Opt(SDP-r)} \leq O(1) \ln(n) \text{Opt(QP)}$$

Application 3: Nonconvex QCQP

Quadratic constrained quadratic programming:

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

Reformulation with rank-1 constraint:

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

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

Application 3: Nonconvex QCQP

SDP relaxation:

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

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)

$\mathcal{S}$-Lemma

$\mathcal{S}$-Lemma. Suppose $A, B \in \mathbb{S}^n$ and $x_0^T A x_0 > 0$ for some x_0. Then

$$x^T B x \geq 0, \forall x : x^T A x \geq 0$$

holds true *if and only if*

$$\exists \lambda \geq 0 : B \succeq \lambda A.$$

Remark.

- Note Farkas' Lemma only applies to convex functions. Here the quadratic functions are not necessarily convex.
- Can not generalize to more than one constraint:

$$x^T B x \geq 0, \forall x : x^T A_i x \geq 0, i = 1, 2$$
$$\not\Rightarrow \exists \lambda_1 \geq 0, \lambda_2 \geq 0, B \succeq \lambda_1 A_1 + \lambda_2 A_2.$$

Proof of $\mathcal{S}$-Lemma

- First prove that $x^T Bx \geq 0, \forall x : x^T Ax \geq 0$ implies that

$$\mathrm{tr}(BX) \geq 0, \forall X \succeq 0 : \mathrm{tr}(AX) \geq 0. \quad \text{(why?)}$$

- Equivalently, $\mathrm{Opt}(P) = 0$

$$\begin{array}{rl} \min & \mathrm{tr}(BX) \\ \text{s.t.} & \mathrm{tr}(AX) \geq 0 \\ & X \succeq 0 \end{array} \qquad \text{(P)}$$

- This is guaranteed if and only if the dual is feasible:

$$\begin{array}{rl} \max\limits_{\lambda, Y} & 0 \\ \text{s.t.} & B = \lambda A + Y \\ & \lambda \geq 0, Y \succeq 0 \end{array} \qquad \text{(D)}$$

Application 4: Stability of Dynamical Systems

- Consider the linear dynamical system:

$$\frac{dx}{dt} = Ax(t) + Bu(t), x(0) = x_0$$
$$y(t) = Cx(t)$$

with the *sector constraint* ($\alpha < \beta$):

$$\sigma(y(t), u(t)) = 2(\beta y(t) - u(t))^T(u(t) - \alpha y(t)) \geq 0.$$

- The system is stable iff $\exists P \in \mathbb{S}^n$ such that the Lyapunov function $V(t) = x(t)^T Px(t)$ is non-increasing, i.e.,

$$\frac{dV(t)}{dt} < 0, \forall x(t), u(t) : \sigma(Cx(t), u(t)) \geq 0.$$

Application 4: Stability of Dynamical Systems

Note

$$\frac{dV(t)}{dt} = \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} A^TP + PA & PB \\ B^TP & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$$

$$\sigma(Cx(t), u(t)) = \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} -2\alpha\beta C^TC & (\alpha+\beta)C^T \\ (\alpha+\beta)C & -2 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$$

The system is stable iff the LMI is feasible:

$$\begin{bmatrix} A^TP + PA - 2\alpha\beta C^TC & PB + (\alpha+\beta)C^T \\ B^TP + (\alpha+\beta)C & -2 \end{bmatrix} \preceq 0$$

Application 4: Stability of Dynamical Systems

Now consider the linear dynamical system:

$$\frac{dx}{dt} = Ax(t) + Bu(t), x(0) = x_0$$
$$y(t) = Cx(t)$$

with the *unity-bounded constraint* ($\alpha < \beta$):

$$|u_i(t)| \leq |y_i(t)|, i = 1, 2, \ldots, p.$$

The system is stable if the LMI is feasible:

$$\exists P \in \mathbb{S}^n, D = \mathrm{diag}(\lambda_1, \ldots, \lambda_p) \text{ such that}$$
$$\begin{bmatrix} A^T P + PA + C^T DC & PB \\ B^T P & -D \end{bmatrix} \preceq 0.$$

More Applications of SDP Relaxation

Machine Learning

- Low-rank matrix factorization;
- k-means for clustering;
- Graphical lasso for estimating covariance matrix;

Optimization

- Robust optimization and chance constraint programs;
- Trust region methods;
- Polynomial optimization;
- Optimal control;

Signal Processing

- MIMO detection in signal processing;
- Stochastic block models for community detection;

References

- Ben-Tal & Nemirovski (2013), Chapters 3.1-3.6