Semidefinite programming
Special and general cases

Imre Pólik, PhD

Lehigh University
Department of Industrial and Systems Engineering

January 20, 2009

Outline

Semidefinite optimization - simplified notation

The unknown is a matrix:

$$\begin{aligned} \min\ & C \bullet X \\ \mathcal{A}X &= b \\ X &\succeq 0 \end{aligned} \qquad\qquad \begin{aligned} \max\ & b^T y \\ \mathcal{A}^* y + S &= C \\ S &\succeq 0 \end{aligned}$$

- C, X, S are $n \times n$ symmetric matrices
- $b, y \in \mathbb{R}^m$ are vectors
- $\mathcal{A} : \mathbb{R}^{n\times n} \to \mathbb{R}^m$ is a linear operator

Block-diagonal form – I

Multiple unknowns:

$$
\begin{aligned}
\min\ & C_1 \bullet X_1 + C_2 \bullet X_2 \\
& \mathcal{A}_1 X_1 + \mathcal{A}_2 X_2 = b \\
& X_1 \quad X_2 \succeq 0
\end{aligned}
\qquad
\begin{aligned}
\max\ & b^T y \\
& \mathcal{A}_1^* y + S_1 = C_1 \\
& \mathcal{A}_2^* y + S_2 = C_2 \\
& S_1, S_2 \succeq 0
\end{aligned}
$$

- $C_{1,2}, X_{1,2}, S_{1,2}$ are $n \times n$ symmetric matrices
- $b, y \in \mathbb{R}^m$ are vectors
- $\mathcal{A}_{1,2} : \mathbb{R}^{n\times n} \to \mathbb{R}^m$ is a linear operator

X_1 and X_2 can be of different size

Block-diagonal form – II

$$C_1 \bullet X_1 + C_2 \bullet X_2 = \underbrace{\begin{pmatrix} C_1 & \\ & C_2 \end{pmatrix}}_{C} \bullet \underbrace{\begin{pmatrix} X_1 & \\ & X_2 \end{pmatrix}}_{X}$$

$$\mathcal{A}_1 X_1 + \mathcal{A}_2 X_2 = \underbrace{\begin{pmatrix} \mathcal{A}_1 & \\ & \mathcal{A}_2 \end{pmatrix}}_{\mathcal{A}} \begin{pmatrix} X_1 & \\ & X_2 \end{pmatrix}$$

$$\begin{aligned} \min\ C \bullet X \\ \mathcal{A}X = b \\ X \succeq 0 \end{aligned} \qquad\qquad \begin{aligned} \max\ b^T y \\ \mathcal{A}^* y + S = C \\ S \succeq 0 \end{aligned}$$

Any number and size of PSD blocks can be converted into one block.

Linear programming

Assume C, X, S and A_i are diagonal:

$$X \succeq 0 \Leftrightarrow \operatorname{diag}(X) = x \geq 0$$
$$C \bullet X = \operatorname{diag}(C)^T \operatorname{diag}(X) = c^T x$$
$$A_i \bullet X = \operatorname{diag}(A_i)^T \operatorname{diag}(X) = a_i^T x$$

$$\begin{aligned} \min\ & c^T x \\ Ax &= b \\ x &\geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\geq 0 \end{aligned}$$

- Linear programming can be cast as SDP using diagonal matrices.
- Also, as a product of 1×1 SDPs.

The Lorentz cone

- (Essentially) equivalent formulations:

$$\mathbb{L}^n = \{u : u_1 \geq \|u_{2:n}\|_2\} \subset \mathbb{R}^n$$
$$\mathbb{L}^n_{\mathrm{r}} = \left\{u : u_1 u_2 \geq \|u_{3:n}\|_2^2, u_1, u_2 \geq 0\right\} \subset \mathbb{R}^n$$

- Standard scalar product: $u^T v$
- For $(u, v) \in \mathbb{L}$: $u^T v \geq 0$ (nontrivial)
- Self-dual: $(\mathbb{L}^n)^* = \mathbb{L}^n$
- Also called: second-order, ice-cream or quadratic cone
- SDP representation (arrowhead matrix):

$$u \in \mathbb{L}^n \Leftrightarrow \begin{pmatrix} u_1 & u_{2:n}^T \\ u_{2:n} & u_1 I \end{pmatrix} \succeq 0$$

Second-order cone programming (SOCP)

- One cone:

$$\min c^T x \qquad \max b^T y$$
$$Ax = b \qquad A^T y + s = c$$
$$x \in \mathbb{L}^n \qquad s \in \mathbb{L}^n$$

- Multiple (k) cones, different dimensions:

$$\min \sum_{j=1}^{k} {c^j}^T x^j \qquad \max b^T y$$
$$\sum_{j=1}^{k} A^j x^j = b \qquad {A^j}^T y + s^j = c^j,\ j = 1, \dots, k$$
$$x^j \in \mathbb{L}^{n_j} \qquad s^j \in \mathbb{L}^{n_j},\ j = 1, \dots, k$$

- Simplification? No!

Mixed LP/SOCP/SDP

$$\min \; c_\ell^T x_\ell + \sum_{j=1}^{k} {c_q^j}^T x_q^j + C \bullet X \qquad \max \; b^T y$$

$$A_\ell^T y + s\ell = c_\ell$$

$$A_\ell x_\ell + \sum_{j=1}^{k} A_q^j x_q^j + \mathcal{A} X = b \qquad {A^j}^T y + s^j = c^j, j = 1, \ldots, k$$

$$\mathcal{A}^* y + S = C$$

$$x_\ell \geq 0 \qquad s_\ell \geq 0$$

$$x_q^j \in \mathbb{L}^{n_j} \qquad s_q^j \in \mathbb{L}^{n_j}, j = 1, \ldots, k$$

$$X \succeq 0 \qquad S \succeq 0$$

All of this can be included in one (big) SDP