Linear Conic Optimization Part IV

Wenxun Xing

Department of Mathematical Sciences
Tsinghua University
wxing@tsinghua.edu.cn
Office hour: Thur. 4:00-5:00.

May, 2023

Conic Representations

Content

- SOC representable
- LMI forms or SDP representable
- Interior point method

SOC Representable

- SOC representable set
  For a given $\mathcal{X}$, if there exist $n_i \times (n+p)$ matrix A_i, second-order cone $\mathcal{L}^{n_i}$, $b_i \in \mathbb{R}^{n_i}$ for $i = 1, 2, \ldots, r$ and $u \in \mathbb{R}^p$, such that

$$\mathcal{X} = \left\{ x \in \mathbb{R}^n \mid A_i \begin{pmatrix} x \\ u \end{pmatrix} \geq_{\mathcal{L}^{n_i}} b_i, i = 1, 2, \ldots, r \right\},$$

  then $\mathcal{X}$ is called a second-order cone representable set.
- SOC representable function
  For a given $f(x)$, if:

$$\text{epi } f = \left\{ \begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{R}^{n+1} \mid f(x) \leq t \right\}$$

  is a second-order cone representable set, then $f(x)$ is called a second-order cone representable function.

The necessary of the variable u

$$\mathcal{X} = \left\{(x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid \sqrt{x_1 x_2} \geq x_3, x_1 \geq 0, x_2 \geq 0\right\}.$$

Whenever $x_3 \geq 0$,

$$\sqrt{x_1 x_2} \geq x_3 \Leftrightarrow x_1 x_2 \geq x_3^2,$$

$$\Leftrightarrow (\frac{x_1 + x_2}{2})^2 \geq x_3^2 + (\frac{x_1 - x_2}{2})^2 \Leftrightarrow \sqrt{x_3^2 + (\frac{x_1 - x_2}{2})^2} \leq \frac{x_1 + x_2}{2}.$$

Then

$$\begin{aligned} \mathcal{Y} \quad &= \mathcal{X} \cap \{x_3 \geq 0\} = \left\{(x_1, x_2, x_3)^T \in \mathbb{R}^3_+ \mid \sqrt{x_1 x_2} \geq x_3\right\} \\ &= \left\{(x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid A_1 x \geq_{\mathcal{L}^3} 0, A_2 x \geq_{\mathcal{L}^1} 0, A_3 x \geq_{\mathcal{L}^1} 0, A_4 x \geq_{\mathcal{L}^1} 0\right\}, \end{aligned}$$

where,

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, A_2 = (1, 0, 0), A_3 = (0, 1, 0), A_4 = (0, 0, 1).$$

Add a variable u!

Whenever $x_3 < 0$,

$$\sqrt{x_1 x_2} \geq x_3, x_1 \geq 0, x_2 \geq 0 \Leftrightarrow \sqrt{x_1 x_2} \geq u, x_1 \geq 0, x_2 \geq 0, u \geq x_3, u \geq 0$$

$$\Leftrightarrow (\frac{x_1 + x_2}{2})^2 \geq u^2 + (\frac{x_1 - x_2}{2})^2, x_1 \geq 0, x_2 \geq 0, u \geq x_3, u \geq 0$$

$$\Leftrightarrow \sqrt{u^2 + (\frac{x_1 - x_2}{2})^2} \leq \frac{x_1 + x_2}{2}, x_1 \geq 0, x_2 \geq 0, u \geq x_3,\ u \geq 0.$$

$$\mathcal{X} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid \begin{array}{c} A_1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u \end{pmatrix} \geq_{\mathcal{L}^3} 0,\ A_2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u \end{pmatrix} \geq_{\mathcal{L}^1} 0 \\ A_3 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u \end{pmatrix} \geq_{\mathcal{L}^1} 0,\ A_4 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u \end{pmatrix} \geq_{\mathcal{L}^1} 0,\ A_5 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ u \end{pmatrix} \geq_{\mathcal{L}^1} 0 \end{array} \right\}.$$

where,

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix}, A_2 = (1, 0, 0, 0), A_3 = (0, 1, 0, 0), A_4 = (0, 0, -1, 1), A_5 = (0, 0, 0, 1).$$

Some useful results for SOC-R sets

Theorem

If $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_k \subseteq \mathbb{R}^n$ are SOC-representable, then (i) $\alpha\mathcal{X}_1$ for any $\alpha > 0$, (ii) $\mathcal{X}_1 \cap \mathcal{X}_2 \cap \cdots \cap \mathcal{X}_k$, (iii) $\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_k$ and (iv) $\mathcal{X}_1 + \mathcal{X}_1 + \cdots + \mathcal{X}_k$ are SOC-representable.

Theorem

Let $B \in \mathcal{M}(m,n)$, $d \in \mathbb{R}^m$ and linear transformation

$$x \in \mathcal{X} \subseteq \mathbb{R}^n \mapsto y = Bx + d \in \mathbb{R}^m$$

and denote

$$\mathcal{Y} = \{y \in \mathbb{R}^m \mid y = Bx + d, x \in \mathcal{X}\}.$$

If $\mathcal{X}$ is SOC-representable, so is $\mathcal{Y}$.

Some useful results for SOC-R functions

Theorem

If $f_1(x), f_2(x), \ldots, f_k(x)$ *are SOC-representable functions in* $\mathbb{R}^n$*, then (i)* $\alpha f_1(x)$ *for any* $\alpha > 0$*, (ii)* $\max\{f_1(x), f_2(x), \ldots, f_k(x)\}$*, and (iii)* $f_1(x) + f_2(x) + \cdots + f_k(x)$ *are SOC-representable.*

Theorem

If $f_1(y)$ *and* $f_2(x)$ *are convex and SOC-representable functions,* $f_1(y)$ *is monotonic nondecreasing, then* $f_1(f_2(x))$ *is convex and SOC-representable.*

Simple SOC Representable Sets/Functions

- $g(x) \equiv c$.
  Its epigraph is $\left\{ \begin{pmatrix} x \\ t \end{pmatrix} \mid c \leq t \right\}$. Let $A = (0)_{m \times n}$, then $\|Ax\| \leq t - c$, i.e., $\begin{pmatrix} Ax \\ t - c \end{pmatrix} \in \mathcal{L}^{m+1}$.
- Linear function $g(x) = Ax + b, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.
  For simple case $g(x) = a^T x + b, a \in \mathbb{R}^n, b \in \mathbb{R}$, there exists a $C = (0)_{p \times n}$ such that

  $$\left\{ \begin{pmatrix} x \\ t \end{pmatrix} \mid a^T x + b \leq t \right\}$$

  is represneted by $\|Cx\| \leq t - a^T x - b$.

Simple SOC Representable Sets/Functions

- $g(x) = \sqrt{x^T A x}, A \in \mathcal{S}_+^n$.
  Epigraph: $\left\{ \begin{pmatrix} x \\ t \end{pmatrix} \mid \sqrt{x^T A x} \le t \right\}$.
  As $A = B^T B$, let $y = Bx$. Then $\sqrt{y^T y} \le t$.
- $g(x) = x^T A x + b^T x + c, A \in \mathcal{S}_+^n$.
  Epigraph: $\left\{ \begin{pmatrix} x \\ t \end{pmatrix} \mid x^T A x + b^T x + c \le t \right\}$.
  Denote $A = B^T B$,

$$x^T A x + b^T x + c \le t \Leftrightarrow x^T A x \le t - b^T x - c$$
$$\Leftrightarrow$$
$$\sqrt{(Bx)^T Bx + \frac{(t - b^T x - c - 1)^2}{4}} \le \frac{t - b^T x - c + 1}{2}.$$

Let $y = Bx, z_1 = \frac{t - b^T x - c - 1}{2}, z_2 = \frac{t - b^T x - c + 1}{2}$. Then $\sqrt{y^T y + z_1^2} \le z_2$.

Simple SOC Representable Sets/Functions

- $g(x,s) = \begin{cases} \frac{x^TAx}{s}, & s > 0 \\ 0, & x^TAx = 0, s = 0 \\ +\infty, & \text{otherwise} \end{cases}$, where $A \in \mathcal{S}^n_+$.

  Epigraph:

$$\left\{ \begin{pmatrix} x \\ s \\ t \end{pmatrix} \mid g(x,s) \le t \right\}.$$

  By

$$\begin{aligned} g(x,s) \le t &\Leftrightarrow x^TAx \le st, s \ge 0, t \ge 0 \\ &\Leftrightarrow x^TAx + \tfrac{(t-s)^2}{4} \le \tfrac{(t+s)^2}{4}, s \ge 0, t \ge 0 \\ &\Leftrightarrow \sqrt{(Bx)^TBx + \tfrac{(t-s)^2}{4}} \le \tfrac{t+s}{2}, s \ge 0, t \ge 0. \end{aligned}$$

  Let $y = Bx, z_1 = \frac{t-s}{2}, z_2 = \frac{t+s}{2}$, then $\sqrt{y^Ty + z_1^2} \le z_2, s, t \ge 0$.

Simple SOC Representable Sets/Functions

- $g(x) = \frac{x^T Ax}{c^T x}, x \in \mathcal{X}$, where $c^T x \geq \alpha > 0, x \in \mathcal{X}$, $A \in \mathcal{S}_+^n$ and $\mathcal{X}$ is SOC representable.
  Epigraph:
$$\left\{ \left(\begin{array}{c} x \\ t \end{array} \right) \mid g(x) \leq t, x \in \mathcal{X} \right\}.$$
  By
$$\begin{aligned} g(x) \leq t, x \in \mathcal{X} &\Leftrightarrow x^T Ax \leq c^T xt, c^T x \geq 0, t \geq 0, x \in \mathcal{X} \\ &\Leftrightarrow x^T Ax + \tfrac{(t-c^T x)^2}{4} \leq \tfrac{(t+c^T x)^2}{4}, c^T x \geq 0, t \geq 0, x \in \mathcal{X} \\ &\Leftrightarrow \sqrt{(Bx)^T Bx + \tfrac{(t-c^T x)^2}{4}} \leq \tfrac{t+c^T x}{2}, s \geq 0, t \geq 0, x \in \mathcal{X}. \end{aligned}$$
  Let $y = Bx, z_1 = \frac{t-c^T x}{2}, z_2 = \frac{t+c^x}{2}$, then $\sqrt{y^T y + z_1^2} \leq z_2, s, t \geq 0, x \in \mathcal{X}$.

Simple SOC Representable Sets/Functions

- Hyperbola $g(x) = \frac{1}{x}, x > 0$. Epigraph:

$$\left\{ \begin{pmatrix} x \\ t \end{pmatrix} \mid g(x) \leq t, x > 0 \right\}.$$

Then

$$g(x) \leq t, x > 0 \Leftrightarrow xt \geq 1, x \geq 0 \Leftrightarrow \frac{(x+t)^2}{4} \geq \frac{(x-t)^2}{4} + 1, x \geq 0$$
$$\Leftrightarrow \sqrt{\frac{(x-t)^2}{4} + 1} \leq \frac{x+t}{2}, x \geq 0.$$

Let $y = \frac{x-t}{2}, z_1 = 1, z_2 = \frac{x+t}{2}$, we have $\sqrt{y^T y + z_1^2} \leq z_2, x \geq 0$.

SOC Representable Sets/Functions

- $\mathcal{K}_+^3 = \left\{(x_1, x_2, x_3)^T \in \mathbb{R}_+^3 \mid \sqrt{x_1 x_2} \geq x_3\right\}.$

$$\sqrt{x_1 x_2} \geq x_3, x \in \mathbb{R}_+^3 \Leftrightarrow x_1 x_2 \geq x_3^2, x \in \mathbb{R}_+^3$$
$$\Leftrightarrow (\tfrac{x_1+x_2}{2})^2 - (\tfrac{x_1-x_2}{2})^2 \geq x_3^2 \Leftrightarrow \tfrac{x_1+x_2}{2} \geq \sqrt{(\tfrac{x_1-x_2}{2})^2 + x_3^2}, x \in \mathbb{R}_+^3.$$

- $\mathcal{K}^3 = \left\{(x_1, x_2, x_3)^T \in \mathbb{R}_+^2 \times \mathbb{R} \mid \sqrt{x_1 x_2} \geq x_3\right\}$

$$\sqrt{x_1 x_2} \geq x_3, (x_1, x_2, x_3)^T \in \mathbb{R}_+^2 \times \mathbb{R}$$
$$\Leftrightarrow \sqrt{x_1 x_2} \geq s \geq 0, s \geq x_3, (x_1, x_2, x_3)^T \in \mathbb{R}_+^2 \times \mathbb{R}, s \in \mathbb{R}_+.$$

SOC Representable Sets/Functions

- $\mathcal{K}_+^{2^n+1} = \left\{(x_1, \ldots, x_{2^n}, t)^T \in \mathbb{R}_+^{2^n+1} \mid (x_1 \cdots x_{2^n})^{\frac{1}{2^n}} \geq t\right\}$

$$(x_1 \cdots x_{2^n})^{\frac{1}{2^n}} \geq t, (x_1, \ldots, x_{2^n}, t)^T \in \mathbb{R}_+^{2^n+1}$$

is equivalent to

$$x_{01} = x_1, x_{02} = x_2, \ldots, x_{02^n} = x_{2^n}, (x_1, \ldots, x_{2^n}, t)^T \in \mathbb{R}_+^{2^n+1}$$

$$0 \leq x_{11} \leq \sqrt{x_{01}x_{02}}, 0 \leq x_{12} \leq \sqrt{x_{03}x_{04}}, \ldots, 0 \leq x_{12^{n-1}} \leq \sqrt{x_{0(2^n-1)}x_{02^n}},$$

$$0 \leq x_{21} \leq \sqrt{x_{11}x_{12}}, 0 \leq x_{22} \leq \sqrt{x_{13}x_{14}}, \ldots, 0 \leq x_{22^{n-2}} \leq \sqrt{x_{1(2^{n-1}-1)}x_{12^{n-1}}},$$

$$\cdots\cdots\cdots$$

$$0 \leq x_{(n-1)1} \leq \sqrt{x_{(n-2)1}x_{(n-2)2}}, \;\; 0 \leq x_{(n-1)2} \leq \sqrt{x_{(n-2)3}x_{(n-2)4}}$$

$$t \leq \sqrt{x_{(n-1)1}x_{(n-1)2}}.$$

SOC Representable Sets/Functions

- $f(x_1, x_2, \ldots, x_n) = (x_1 x_2 \cdots x_n)^{-q}$, $x \in \mathbb{R}^n_{++}$, $q > 0$ is a rational number.

$$\text{epi}(f) = \left\{ \begin{pmatrix} x \\ t \end{pmatrix} \mid x \in \mathbb{R}^n_+, t \in \mathbb{R}_+, (x_1 x_2 \cdots x_n)^{-q} \le t \right\}.$$

$(x_1 x_2 \cdots x_n)^{-q} \le t, x \in \mathbb{R}^n_+,\ t \ge 0 \Rightarrow x \in \mathbb{R}^n_{++}$.

Let $q = \frac{r}{p}$, where r, p are integers. Choose the smallest l such that $nr + p \le 2^l$.

Consider

$$\mathcal{K}_+^{2^l+1} = \left\{ (y, s) \in \mathbb{R}^n_+ \times \mathbb{R}_+ \mid (y_1 y_2 \cdots y_{2^l})^{\frac{1}{2^l}} \ge s \right\}.$$

SOC Representable Sets/Functions

$$y_1 = y_2 = \cdots = y_r = x_1, \quad y_{r+1} = y_{r+2} = \cdots = y_{2r} = x_2,$$

$$\cdots\cdots$$

$$y_{(n-1)r+1} = y_{(n-1)r+2} = \cdots = y_{nr} = x_n, \quad y_{nr+1} = y_{nr+2} = \cdots = y_{nr+p} = t,$$

$$y_{nr+p+1} = y_{nr+p+2} = \cdots = y_{2^l} = s = 1.$$

Then $(y_1 y_2 \cdots y_{2^l})^{\frac{1}{2^l}} \geq s$ implies

$$(x_1 x_2 \cdots x_n)^{\frac{r}{2^l}} t^{\frac{p}{2^l}} \geq 1.$$

So

$$t^{\frac{p}{2^l}} \geq (x_1 x_2 \cdots x_n)^{-\frac{r}{2^l}},$$

i.e.

$$t \geq (x_1 x_2 \cdots x_n)^{-\frac{r}{p}} = (x_1 x_2 \cdots x_n)^{-q}.$$

Convex quadratically constrained quadratic programming

$$\begin{array}{ll} \min & \frac{1}{2}x^T Q_0 x + f_0^T x \\ \text{s.t.} & \frac{1}{2}x^T Q_i x + f_i^T x \le c_i,\ i = 1, 2, \cdots, m \\ & x \in \mathbb{R}^n, \end{array}$$

where $Q_i \in \mathcal{S}_+^n$, $i = 0, 1, \cdots, m$.

- An equivalent form

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \frac{1}{2}x^T Q_0 x \le t - f_0^T x \\ & \frac{1}{2}x^T Q_i x \le c_i - f_i^T x,\ i = 1, 2, \cdots, m \\ & x \in \mathbb{R}^n. \end{array}$$

Convex quadratically constrained quadratic programming

Let

$$\begin{cases} u^0 = P_0 x, \quad v_0 = \frac{1-t+f_0^T x}{\sqrt{2}}, \quad w_0 = \frac{1+t-f_0^T x}{\sqrt{2}} \\ u^i = P_i x, \quad v_i = \frac{1-c_i+f_i^T x}{\sqrt{2}}, \quad w_i = \frac{1+c_i-f_i^T x}{\sqrt{2}}, i = 1, 2, \ldots, m. \end{cases}$$

A second-order conic programming problem

$$\begin{array}{ll} \min & t \\ \text{s.t.} & u^0 = P_0 x,\ v_0 = \frac{1-t+f_0^T x}{\sqrt{2}},\ w_0 = \frac{1+t-f_0^T x}{\sqrt{2}} \\ & u^i = P_i x,\ v_i = \frac{1-c_i+f_i^T x}{\sqrt{2}},\ w_i = \frac{1+c_i-f_i^T x}{\sqrt{2}}, i = 1, 2, \ldots, m \\ & \begin{pmatrix} u^0 \\ v_0 \\ w_0 \end{pmatrix} \in \mathcal{L}^{n+2}; \begin{pmatrix} u^i \\ v_i \\ w_i \end{pmatrix} \in \mathcal{L}^{n+2}, i = 1, 2, \ldots, m; x \in \mathbb{R}^n; t \in \mathbb{R}. \end{array}$$

Robust linear programming

- Linear Programming

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

- Uncertainty $(c, A, b) \in \mathcal{U}$.

$$A^T = (A_1, A_2, \cdots, A_m),\ b = (b_1, b_2, \cdots, b_m)^T,\ A_i \in \mathbb{R}^n,$$

$$\mathcal{U} = \{A, b, c \mid c = c^* + P_0 u_0, \begin{pmatrix} A_i \\ b_i \end{pmatrix} = \begin{pmatrix} A_i^* \\ b_i^* \end{pmatrix} + P_i u_i, i = 1, 2, \cdots, m\},$$

$$u_i^T u_i \leq 1, i = 0, 1, 2, \cdots, m.$$

Robust linear programming

- Robust model

$$\begin{aligned} \min_{(c,A,b)\in\mathcal{U}} \quad & t \\ \text{s.t.} \quad & c^T x \le t \\ & Ax \ge b \\ & x \in \mathbb{R}^n_+. \end{aligned}$$

- Constraints

$$\begin{aligned} 0 \quad & \le \min_{u_i^T u_i \le 1} \left\{ A_i^T(u)x - b_i(u) \mid \begin{pmatrix} A_i \\ b_i \end{pmatrix} = \begin{pmatrix} A_i^* \\ b_i^* \end{pmatrix} + P_i u_i \right\} \\ & = (A_i^*)^T x - b_i^* + \min_{u_i^T u_i \le 1} u_i^T P_i^T \begin{pmatrix} x \\ -1 \end{pmatrix} \\ & = (A_i^*)^T x - b_i^* - \left\| P_i^T \begin{pmatrix} x \\ -1 \end{pmatrix} \right\|. \end{aligned}$$

Robust linear programming

- Second-order conic model

$$
\begin{array}{ll}
\min & t \\
\text{s.t.} & \left\| P_0^T x \right\| + c^{*T} x \le t \\
& \left\| P_i^T \begin{pmatrix} x \\ -1 \end{pmatrix} \right\| - (A_i^*)^T x \le -b_i^*, i = 1, 2, \cdots m \\
& x \in \mathbb{R}_+^n, t \in \mathbb{R}.
\end{array}
$$

LMI-linear matrix inequality

- $A \bullet X + a^T x \leq b$,
  where $A \in \mathcal{S}^n, a \in \mathbb{R}^r, b \in \mathbb{R}$ are given, $x \in \mathbb{R}^r$, $X \in \mathcal{S}^n_+$ are decision variables.

- $\sum_{j=1}^r x_j C_j - D \in \mathcal{S}^s_+$,
  where $C_j, D \in \mathcal{S}^s, j = 1, 2, \ldots, r$ are given and $x \in \mathbb{R}^r$ is a decision variable.

- LMI representable set: $\mathcal{X}$ is represented by LMIs. LMI representable function: its epigraph is LMI representable.

LMI Representable Examples

- $\mathbb{R}^n_+ = \left\{(x_1, x_2, \ldots, x_n)^T \mid x_i \geq 0, i = 1, 2, \ldots, n\right\}$.

$$(x_1, x_2, \ldots, x_n)^T \geq 0 \Leftrightarrow X = (x_{ij}) \in \mathcal{S}^n_+, x_{ii} - x_i = 0, x_{ij} = 0, i \neq j$$

- $\mathcal{L}^n$.

$$x \in \mathcal{L}^n \Leftrightarrow \begin{pmatrix} x_n I_{n-1} & x_{1:n-1} \\ x_{1:n-1}^T & x_n \end{pmatrix} \in \mathcal{S}^n_+,$$

where $x_{1:n-1} = (x_1, x_2, \ldots, x_{n-1})^T$.

LMI Representable Examples

- For a given $X \in \mathcal{S}^n$, its maximum eigenvalue $\lambda_{max}(X)$ is LMI representable function.

$$\{(X,t) \in \mathcal{S}^n \times \mathbb{R} \mid \lambda_{max}(X) \leq t\} = \{(X,t) \in \mathcal{S}^n \times \mathbb{R} \mid tI - X \in \mathcal{S}^n_+\}$$

- To get the maximum eigenvalue of a given matrix X.

$$\begin{array}{rl} \min & t \\ \text{s.t.} & tI - X \in \mathcal{S}^n_+ \\ & t \in \mathbb{R}. \end{array}$$

- For a given $X \in \mathcal{S}^n$, the maximum of the absolute eigenvalues is LMIr.

$$\begin{aligned} &\{(X,t) \in \mathcal{S}^n \times \mathbb{R} \mid |\lambda(X)|_{max} \leq t\} \\ &= \left\{(X,t) \in \mathcal{S}^n \times \mathbb{R} \mid tI - X \in \mathcal{S}^n_+, tI + X \in \mathcal{S}^n_+\right\}. \end{aligned}$$

LMI Representable Examples

- $f(X) = \begin{cases} \det(X)^{-q}, & X \in \mathcal{S}^n_{++} \\ +\infty, & \text{otherwise,} \end{cases}$ is LMIr function, where $q > 0$ is a rational number.

$$\mathrm{epi}(f) = \left\{(X,t) \in \mathcal{S}^n \times \mathbb{R} \mid \det(X)^{-q} \le t, X \in \mathcal{S}^n_+\right\}$$

and

$$\mathcal{Y} = \left\{ (X,t) \in \mathcal{S}^n \times \mathbb{R} \mid \begin{array}{l} \begin{pmatrix} X & \Delta \\ \Delta^T & D(\Delta) \end{pmatrix} \in \mathcal{S}^{2n}_+, \Delta \text{ lower triangular} \\ D(\Delta) = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n) \text{ diagonal of } \Delta \\ (\delta_1 \delta_2 \cdots \delta_n)^{-q} \le t \end{array} \right\}$$

$\mathcal{Y} \subseteq \text{epi}(f)$

For any $(X, t) \in \mathcal{Y}$, $\begin{pmatrix} X & \Delta \\ \Delta^T & D(\Delta) \end{pmatrix} \in \mathcal{S}_+^{2n}$ implies $D(\Delta) \in \mathcal{S}_+^n$, and $\delta_i \geq 0, i = 1, 2, \ldots, n$. With $(\delta_1 \delta_2 \cdots \delta_n)^{-q} \leq t$, we have $\delta_i > 0, i = 1, 2, \ldots, n$.

Together with $\begin{pmatrix} X & \Delta \\ \Delta^T & D(\Delta) \end{pmatrix} \in \mathcal{S}_+^{2n}$ and Shur Theorem, we have $X - \Delta D^{-1}(\Delta)\Delta^T \in \mathcal{S}_+^n$.

As the diagonal elements of Δ are positive, $\Delta D^{-1}(\Delta)\Delta^T \in \mathcal{S}_{++}^n$ and $X \in \mathcal{S}_{++}^n$. Then there exists an invertible P such that $P^T X P = I$ and $P^T \Delta D^{-1}(\Delta)\Delta^T P = \text{diag}(d_1, d_2, \ldots, d_n)$.

Then $0 \leq d_1 d_2 \cdots d_n \leq 1$ and $\det(P^T X P) \geq \det(P^T \Delta D^{-1}(\Delta)\Delta^T P)$.

$$\det(X) \geq \det(\Delta D^{-1}(\Delta)\Delta^T) = \delta_1 \delta_2 \cdots \delta_n,$$

$$\det(X)^{-q} \leq (\delta_1 \delta_2 \cdots \delta_n)^{-q} \leq t.$$

So $\mathcal{Y} \subseteq \text{epi}(f)$.

$\mathcal{Y} \supseteq \mathrm{epi}(f)$

For any $(X,t) \in \mathrm{epi}(f)$, $X \in \mathcal{S}^n_+$ and $\det(X)^{-q} \leq t$, we have $X \in \mathcal{S}^n_{++}$.
By $X \in \mathcal{S}^n_{++}$ and Cholesky decomposition, there exists a lower triangular matrix L with positive diagonal elements such that $X = LL^T$.
Denote the diagonal elements of L as $a_1, a_2, \ldots, a_n$. Let $\Delta = L\mathrm{diag}(a_1, a_2, \ldots, a_n)$. We have

$$D(\Delta) = \mathrm{diag}(a_1^2, a_2^2, \ldots, a_n^2) = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n),$$

$$X - \Delta D^{-1}(\Delta)\Delta^T = X - LL^T = 0.$$

Thus

$$\begin{pmatrix} X & \Delta \\ \Delta^T & D(\Delta) \end{pmatrix} \in \mathcal{S}^{2n}_+$$

$$\det(X) = \det(LL^T) = a_1^2 a_2^2 \cdots a_n^2 = \det(D(\Delta)) = \delta_1 \delta_2 \cdots \delta_n.$$

We get $(X,t) \in \mathcal{Y}$. So $\mathrm{epi}(f) = \mathcal{Y}$.

SDP relaxation

- QCQP

$$
\begin{aligned}
v_{QP} = \min \quad & f(x) = \tfrac{1}{2}x^T Q_0 x + q_0^T x + c_0 \\
\text{s.t.} \quad & g_i(x) = \tfrac{1}{2}x^T Q_i x + q_i^T x + c_i \leq 0, i = 1, 2, \ldots, m \\
& x \in \mathbb{R}^n.
\end{aligned}
$$

- Relaxation

$$
\begin{aligned}
v_{RP} = \min \quad & \frac{1}{2}\begin{pmatrix} 2c_0 & q_0^T \\ q_0 & Q_0 \end{pmatrix} \bullet X \\
\text{s.t.} \quad & \frac{1}{2}\begin{pmatrix} 2c_i & q_i^T \\ q_i & Q_i \end{pmatrix} \bullet X \leq 0, i = 1, 2, \cdots, m \\
& x_{11} = 1 \\
& X \in \mathcal{S}_+^{n+1}.
\end{aligned}
$$

Rank-one decomposition

Theorem

Let $X \succeq 0$ of rank r. Let G be a given matrix. Then $G \bullet X \geq 0$ if and only if there exist $p_i \in R^n$, $i = 1, 2, \cdots, r$, such that

$$X = \sum_{i=1}^{r} p_i p_i^T \quad \text{and} \quad p_i^T G p_i \geq 0.$$

Procedure

- Input: $X \succeq 0$, G be a given matrix such that $G \cdot X \geq 0$.
- Output: A vector y with $0 \leq y^T G y \leq G \cdot X$ such that $X - yy^T$ is semi-definite positive of rank $r - 1$.

Rank-one decomposition algorithm

Step 0 Compute $p_1, p_2, \cdots, p_r$ such that $X = \sum_{i=1}^r p_i p_i^T$.

Step 1 If $(p_1^T G p_1)(p_i^T G p_i) \geq 0$ for all $i = 2, 3, \cdots, r$ then return $y = p_1$. Otherwise let j be the one (any) such that $(p_1^T G p_1)(p_j^T G p_j) < 0$.

Step 2 Determine α such that $(p_1 + \alpha p_j)^T G (p_1 + \alpha p_j) = 0$. Return $y = (p_1 + \alpha p_j)/\sqrt{1 + \alpha^2}$.

Trust region model–an example

- Trust region model

$$\begin{array}{ll} \min & \frac{1}{2}x^T Ax + f^T x \\ s.t. & \frac{1}{2}x^T Bx \leq \mu, \end{array}$$

  where A, B are $n \times n$ symmetric matrices, B is positive definite, $\mu > 0$.

- SDP relaxation model

$$\begin{array}{rl} Z_R = \min & \frac{1}{2}\begin{bmatrix} 0 & f^T \\ f & A \end{bmatrix} \cdot X \\ s.t. & \frac{1}{2}\begin{bmatrix} 2\mu & 0 \\ 0 & -B \end{bmatrix} \cdot X \geq 0, \\ & X_{11} = 1, \\ & X \succeq 0. \end{array}$$

Optimality

Theorem

For any feasible solution X of the relaxation of the SDP relaxation model, it can be decomposed into

$$X = \sum_{i=1}^{r} p_i p_i^T,$$

such that $(p_i)_1 \neq 0$, $p_i^T \begin{bmatrix} 2\mu & 0 \\ 0 & -B \end{bmatrix} p_i \geq 0$ *and* $\sum_{i=1}^{r} (p_i)_1^2 = 1$, *in which* $(p_i)_1$ *denotes the first component of* p_i*.*

Optimality

Let $y_i = p_i/(p_i)_1$. Then $(y_i)_{2:n+1}$ is a feasible solution of the trust region problem.

$$\begin{aligned}
& \tfrac{1}{2}\begin{bmatrix} 0 & f^T \\ f & A \end{bmatrix} \cdot X \\
= \quad & \tfrac{1}{2}\sum_{i=1}^{r} p_i^T \begin{bmatrix} 0 & f^T \\ f & A \end{bmatrix} p_i \\
= \quad & \tfrac{1}{2}\sum_{i=1}^{r} (p_i)_1^2 \begin{pmatrix} 1 \\ (y_i)_{2:n+1} \end{pmatrix}^T \begin{bmatrix} 0 & f^T \\ f & A \end{bmatrix} \begin{pmatrix} 1 \\ (y_i)_{2:n+1} \end{pmatrix}.
\end{aligned}$$

So $(y_i)_{2:n+1}$ is an optimal solution.

Randomized approximation algorithm for max-cut

- QCQP model

$$
\begin{array}{rl}
Z_{MC} = \max & \frac{1}{4} \sum_{i,j=1}^{n} w_{ij}(1 - x_i x_j) \\
s.t. & x_i^2 = 1, i = 1, 2, \cdots, n.
\end{array}
$$

- SDP relaxation model

$$
\begin{array}{rl}
Z_{SDP} = \max & \frac{1}{4} \sum_{i,j=1}^{n} w_{ij}(1 - x_{ij}) \\
s.t. & X = (x_{ij})_{n \times n} \succeq 0 \\
& x_{ii} = 1, i = 1, 2, \cdots, n.
\end{array}
$$

Randomized approximation algorithm for max-cut

For an optimal solution $X \in \mathcal{S}_+^n$, there exists full rank matrix $B \in \mathcal{M}(m,n)$ such that $X = B^T B$. Let $B = (v^1, v^2, \ldots, v^n)$. Then $X = B^T B = \big((v^i)^T v^j\big)$, $(v^i)^T v^j = x_{ij}$ and $(v^i)^T v^i = x_{ii} = 1$.

Step 0 Solve the SDP relaxation model and get one optimal X with $(v^1, v^2, \ldots, v^n), v^i \in \mathbb{R}^m, i = 1, 2 \ldots, n, m = \mathrm{rank}(X)$;

Step 1 Choose a randomized a over the surface of $\{x \in \mathbb{R}^m \mid \|x\| = 1\}$;

Step 2 For $i = 1, 2, \ldots, n$, if $a^T v^i \geq 0$, then $\eta_i = 1$, otherwise $\eta_i = -1$.

SDP relaxation of max-cut—Analytic results

- $\Pr(\mathrm{sign}(a^T v^i) \neq \mathrm{sign}(a^T v^j)) = \frac{\arccos(v^i, v^j)}{\pi}$.

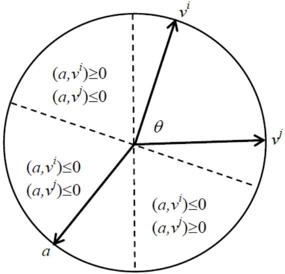

Figure: $\Pr(\mathrm{sign}(a^T v^i) \neq \mathrm{sign}(a^T v^j))$

SDP relaxation of max-cut—Analytic results

- $\Pr(\mathrm{sign}(a^T v^i) \neq \mathrm{sign}(a^T v^j)) = \frac{\arccos(v^i, v^j)}{\pi}$.
- Denote $\theta = \arccos(v^i, v^j)$ and

$$\alpha = \min_{0 \le \theta \le \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta},$$

  then $\alpha \approx 0.87856$.

-

$$\begin{array}{rlrl} v_{RA} = & E\left[\frac{1}{4}\sum_{i,j=1}^n w_{ij}(1 - \eta_i\eta_j)\right] & = \frac{1}{2}\sum_{i,j=1}^n w_{ij}\frac{\arccos(v^i, v^j)}{\pi} \\ \geq & \frac{\alpha}{4}\sum_{i,j=1}^n w_{ij}(1 - (v^i)^T v^j) & = \frac{\alpha}{4}\sum_{i,j=1}^n w_{ij}(1 - x_{ij}) \\ = & \alpha\, v_{SDP}. & \end{array}$$

- $v_{RA} \geq \alpha\, Z_{SDP} \geq \alpha\, Z_{MC}$

Uncertain dynamical linear system (ULS)

$$\frac{d}{dt}x(t) = A(t)x(t), \;\; x(0) = x^0,$$

where $A(t)$ is an $n \times n$ uncertainty matrix, $x(t)$ is an $n \times 1$ vector, x^0 is an initial point.

- Stable (ULS): $x(t) \to 0$ if $t \to +\infty$.
- Conditions of $A(t)$ and x^0 for a stable ULS?

For a dynamical system

$$\frac{d}{dt}x(t) = f(t, x(t)), \;\; x(0) = x^0,$$

where $f(t, 0) = 0$ and $f(x, t)$ is assumed smooth.

$$f(t, x(t)) = f(t, 0) + \int_0^1 \frac{\partial}{\partial s} f(t, sx)xds,$$

$$\frac{d}{dt}x(t) = A(x, t)x(t), \quad x(0) = x^0,$$

where $A(x, t) = \int_0^1 \frac{\partial}{\partial s} f(t, sx)ds$.

Theorem

If there exist an $\alpha > 0$ and a positive-definite matrix X for (ULS) such that $L(x) = x^T Xx$ and

$$\frac{d}{dt}L(x(t)) \leq -\alpha L(x(t)),$$

then (ULS) is stable.

- $L(x) = x^T Xx$ is called Lyapunov's quadratic function.

Theorem

Let $\mathcal{U}$ be the uncertain set of A in (ULS). If the optimal value of the following semi-definite programming problem is negative

$$\begin{array}{ll} \min & s \\ \text{s.t.} & sI_n - A^T X - XA \succeq 0, \ \forall A \in \mathcal{U} \\ & X \succeq I_n \\ & X \in \mathcal{S}^n_+, s \in \mathbb{R}, \end{array}$$

then the dynamic programming is stable.

An easy case

$$\mathcal{U} = \mathrm{conv}\{A_1, A_2, \cdots, A_K\},$$

where A_i is a fixed $n \times n$ matrix,

$$\begin{array}{ll} \min & s \\ \text{s.t.} & sI_n - A_i^T X - XA_i \succeq 0, \ i = 1, 2, \cdots, K \\ & X \succeq I_n \\ & X \in \mathcal{S}^n_+, s \in \mathbb{R}, \end{array}$$

Interior Point Methods

Content

- Interior Points and Primal-Dual Model
- Barrier Functions and Optimal Systems
- Central Path and Newton Methods
- Path Following Method

Interior Point Method

- Interior point method
  - Start from an interior point solution.
  - If the current solution is not good enough, then move to another interior point solution.
  - Stop at an interior point solution whose objective value is close to the optimum (within an ϵ gap).
- Advantages:
  - Polynomial time complexity (comparing with the simplex method for LP)
  - Excellent computational performance in practice (comparing with the ellipsoid method)
- Three types: primal; dual; primal-dual

Primal-dual Model

- Primal-dual type of LP

$$
\begin{array}{ll}
\min & s^T x \\
s.t. & Ax = b \\
 & A^T y + s = c \\
 & x \geq_{\mathbb{R}^n_+} 0, s \geq_{\mathbb{R}^n_+} 0
\end{array}
\qquad \text{(LPD)}
$$

- Primal-dual type of SDP

$$
\begin{array}{ll}
\min & S \bullet X \\
s.t. & \mathcal{A}X = b \\
 & \mathcal{A}^* y + S = C \\
 & X \succeq 0, S \succeq 0
\end{array}
\qquad \text{(SDPD)}
$$

- Note:

$\mathcal{A}X = [A_1 \bullet X, \cdots, A_m \bullet X]^T$
and $\mathcal{A}^* y = \sum_{i=1}^m y_i A_i$

Interior Points

$$\text{feas}^+(\text{LP}) = \{x | Ax = b, x >_{\mathbb{R}^n_+} 0\}$$
$$\text{feas}^+(\text{LD}) = \{(y,s) | A^T y + s = c, s >_{\mathbb{R}^n_+} 0\}$$
$$\text{feas}^+(\text{LPD}) = \text{feas}^+(\text{LP}) \times \text{feas}^+(\text{LD})$$

$$\text{feas}^+(\text{SDP}) = \{X | \mathcal{A}X = b, X \succ 0\}$$
$$\text{feas}^+(\text{SDD}) = \{(y,S) | \mathcal{A}^* y + S = C, S \succ 0\}$$
$$\text{feas}^+(\text{SDPD}) = \text{feas}^+(\text{SDP}) \times \text{feas}^+(\text{SDD})$$

- Assumptions:
  - $\text{feas}^+(\text{LP})$ and $\text{feas}^+(\text{LD})$ are not empty and the rows of A are linearly independent.
  - $\text{feas}^+(\text{SDP})$ and $\text{feas}^+(\text{SDD})$ are not empty and the vectors formed by A_i in $\mathcal{A}$ are linearly independent.

Barrier function

- Properties required:
  - Strictly convex (concave).
  - Goes to $+\infty$ $(-\infty)$ when the point is close to the boundary.
  - Sufficient continuous differentiability.

- Barrier functions:

$$
\begin{array}{llll}
\text{LP}: & -\sum_{i=1}^n \log x_i & \text{SDP}: & -\log\det(X) \\
\text{LD}: & \sum_{i=1}^n \log s_i & \text{SDD}: & \log\det(S) \\
\text{LPD}: & -\sum_{i=1}^n \log(x_i s_i) & \text{SDPD}: & -\log\det(XS)
\end{array}
$$

LP with Barrier

$$
\begin{array}{ll}
\min & c^T x - \mu \sum_{i=1}^{n} \log x_i \\
s.t. & Ax = b \\
 & x >_{\mathbb{R}^n_+} 0
\end{array}
\tag{LPB}
$$

$$
\begin{array}{ll}
\max & b^T y + \mu \sum_{i=1}^{n} \log s_i \\
s.t. & A^T y + s = c \\
 & s >_{\mathbb{R}^n_+} 0
\end{array}
\tag{LDB}
$$

$$
\begin{array}{ll}
\min & s^T x - \mu \sum_{i=1}^{n} \log(x_i s_i) \\
s.t. & Ax = b \\
 & A^T y + s = c \\
 & x >_{\mathbb{R}^n_+} 0, s >_{\mathbb{R}^n_+} 0
\end{array}
\tag{LPDB}
$$

Common Optimal System for LP with Barrier

$$
\begin{aligned}
&Ax = b \\
&A^T y + s = c \\
&\Lambda_x s = \mu e \\
&x >_{\mathbb{R}^n_+} 0, s >_{\mathbb{R}^n_+} 0,
\end{aligned}
$$

where $e = (1, \ldots, 1)^T$ and Λ_x is a diagonal matrix with $(\Lambda_x)_{ii} = x_i$, $i = 1, \ldots, n$.

Notice that

$$
\mu = \frac{x^T s}{n} = \frac{c^T x - b^T y}{n}
$$

When $\mu \to 0$, $s^T x \to 0$. Optimal!

SDP with Barrier

$$
\begin{aligned}
\min \quad & C \bullet X - \mu \log\det(X) \\
s.t. \quad & \mathcal{A}X = b \\
& X \succ 0
\end{aligned}
\qquad \text{(SDPB)}
$$

$$
\begin{aligned}
\min \quad & b^T y + \mu \log\det(S) \\
s.t. \quad & \mathcal{A}^* y + S = C \\
& S \succ 0
\end{aligned}
\qquad \text{(SDDB)}
$$

$$
\begin{aligned}
\min \quad & S \bullet X - \mu \log\det(XS) \\
s.t. \quad & \mathcal{A}X = b \\
& \mathcal{A}^* y + S = C \\
& X \succ 0, S \succ 0
\end{aligned}
\qquad \text{(SDPDB)}
$$

Common Optimal System for SDP with Barrier

$$
\begin{aligned}
&\mathcal{A}X = b \\
&\mathcal{A}^* y + S = C \\
&XS = \mu I \\
&X \succ 0, S \succ 0
\end{aligned}
$$

Notice that

$$
\mu = \frac{S \bullet X}{n} = \frac{C \bullet X - b^T y}{n}
$$

When $\mu \to 0$, $S \bullet X \to 0$. Optimal!

Central Path for LP and SDP

$$\mathcal{C}_{\text{LP}} = \{(x, y, s) \in \text{feas}^+(\text{LPD}) | \Lambda_x s = \mu e, 0 < \mu < +\infty\}$$

$$\mathcal{C}_{\text{SDP}} = \{(X, y, S) \in \text{feas}^+(\text{SDPD}) | XS = \mu I, 0 < \mu < +\infty\}$$

Under proper assumptions:

- For any $0 < \mu < +\infty$, there exists a unique point on central path.

  LP: $(x(\mu), y(\mu), s(\mu))$

  SDP: $(X(\mu), y(\mu), S(\mu))$

- Given $\bar{\mu} > 0$, the set $\{(x, y, s) \in \text{feas}^+(\text{LPD}) | \Lambda_x s = \mu e, 0 < \mu < \bar{\mu}\}$ is bounded.

  Given $\bar{\mu} > 0$, the set $\{(X, y, S) \in \text{feas}^+(\text{SDPD}) | XS = \mu I, 0 < \mu < \bar{\mu}\}$ is bounded.

Example: Central Path

$$
\begin{aligned}
Min \quad & x_1 + x_2 \\
s.t. \quad & x_1 + x_2 \leq 3 \\
& x_1 - x_2 \leq 1 \\
& x_2 \leq 2 \\
& x_1 \geq 0, x_2 \geq 0
\end{aligned}
$$

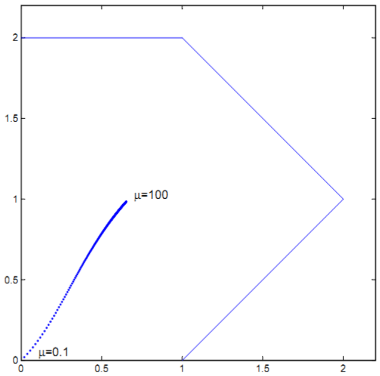

Figure: Projection of central path on (x_1, x_2)

Newton Method for LP

Given $(x^0, y^0, s^0) \in \text{feas}^+(\text{LPD})$ with $\mu^0 = \frac{(s^0)^T x^0}{n}$ and $0 \le \gamma \le 1$, find (d_x, d_y, d_s) satisfying

$$
\begin{array}{l}
A(x^0 + d_x) = b \\
A^T(y^0 + d_y) + (s^0 + d_s) = c \\
\Lambda_{x^0 + d_x}(s^0 + d_s) = \gamma \mu^0 e \\
x^0 + d_x >_{\mathbb{R}^n_+} 0, s^0 + d_s >_{\mathbb{R}^n_+} 0,
\end{array}
$$

After linearization

$$
\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ \Lambda_{s^0} & 0 & \Lambda_{x^0} \end{bmatrix} \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \gamma \mu^0 e - \Lambda_{x^0} \Lambda_{s^0} e \end{bmatrix}
$$
$$
x^0 + d_x >_{\mathbb{R}^n_+} 0, \ s^0 + d_s >_{\mathbb{R}^n_+} 0,
$$

Directly solve the equation is not easy.

Newton Method for LP

Linear scaling: Given a positive diagonal matrix $D \in \mathbb{R}^{n\times n}$,

$$\bar{A} = AD, \bar{x}^0 = D^{-1}x^0, \bar{s}^0 = Ds^0, \bar{c} = Dc$$

$$\begin{bmatrix} \bar{A} & 0 & 0 \\ 0 & \bar{A}^T & I \\ \Lambda_{\bar{s}^0} & 0 & \Lambda_{\bar{x}^0} \end{bmatrix} \begin{bmatrix} \bar{d}_x \\ \bar{d}_y \\ \bar{d}_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \gamma\mu^0 e - \Lambda_{\bar{x}^0}\Lambda_{\bar{s}^0} e \end{bmatrix}$$

$$\bar{x}^0 + \bar{d}_x >_{\mathbb{R}^n_+} 0,\ \bar{s}^0 + \bar{d}_s >_{\mathbb{R}^n_+} 0,$$

- $D = \Lambda_{x^0}$: $\quad \bar{x}^0 = e \quad \Rightarrow \quad \bar{x}^0 + \bar{d}_x >_{\mathbb{R}^n_+} 0, \forall \|\bar{d}_x\|_2 < 1$ (Primal)
- $D = \Lambda_{s^0}^{-1}$: $\quad \bar{s}^0 = e \quad \Rightarrow \quad \bar{s}^0 + \bar{d}_s >_{\mathbb{R}^n_+} 0, \forall \|\bar{d}_s\|_2 < 1$ (Dual)
- $D = \Lambda_{x^0}^{1/2}\Lambda_{s^0}^{-1/2}$: $\quad v^0 = \bar{x}^0 = \bar{s}^0 = \Lambda_{x^0}^{1/2}\Lambda_{s^0}^{1/2} e$ (Primal-dual)

Primal-Dual Interior-Point Method for LP

$D = \Lambda_{x^0}^{1/2}\Lambda_{s^0}^{-1/2}$:

$$\begin{bmatrix} \bar{A} & 0 & 0 \\ 0 & \bar{A}^T & I \\ I & 0 & I \end{bmatrix} \begin{bmatrix} \bar{d}_x \\ \bar{d}_y \\ \bar{d}_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \gamma\mu^0\Lambda_{v^0}^{-1}e - v^0 \end{bmatrix}$$
$$\bar{x}^0 + \bar{d}_x >_{\mathbb{R}_+^n} 0,\ \bar{s}^0 + \bar{d}_s >_{\mathbb{R}_+^n} 0,$$

One can solve

$$\bar{A}\bar{A}^T\bar{d}_y = -\bar{A}(\gamma\mu^0\Lambda_{v^0}^{-1}e - v^0)$$

And then solve $\bar{d}_s$ and $\bar{d}_x$:

$$\begin{aligned} \bar{d}_s &= -\bar{A}^T\bar{d}_y \\ \bar{d}_x &= -\bar{d}_s + \gamma\mu^0\Lambda_{v^0}^{-1}e - v^0 \end{aligned}$$

Newton Method for SDP

Given $(X^0, y^0, S^0) \in \text{feas}^+(\text{SDPD})$ with $\mu^0 = \frac{S^0 \bullet X^0}{n}$ and $0 \le \gamma \le 1$, find $(\triangle X, d_y, \triangle S)$ satisfying

$$
\begin{array}{l}
\mathcal{A}(X^0 + \triangle X) = b \\
\mathcal{A}^*(y^0 + d_y) + (S^0 + \triangle S) = C \\
(X^0 + \triangle X)(S^0 + \triangle S) = \gamma \mu^0 I \\
X^0 + \triangle X \succ 0, S^0 + \triangle S \succ 0
\end{array}
$$

After linearization

$$
\begin{array}{lccccc}
\mathcal{A}\triangle X & & & & = & 0 \\
 & \mathcal{A}^* dy & + & \triangle S & = & 0 \\
\triangle X S^0 & & + & X^0 \triangle S & = & \gamma \mu^0 I - X^0 S^0
\end{array}
$$

$$X^0 + \triangle X \succ 0, S^0 + \triangle S \succ 0.$$

Directly solve the equation is not easy.

Newton Method for SDP

Linear transformation: Given an invertible matrix $L \in \mathbb{R}^{n\times n}$, let
$\bar{\mathcal{A}} = (\bar{A}_1, \ldots, \bar{A}_m)$, $\bar{A}_i = L^T A_i L$ for $i = 1, \ldots, m$.
$\bar{X}^0 = L^{-1} X^0 L^{-T}$, $\bar{S}^0 = L^T S^0 L$, $\bar{C} = L^T C L$.

$$
\begin{array}{rcrcl}
\bar{\mathcal{A}} \triangle \bar{X} & & & = & 0 \\
& \bar{\mathcal{A}}^* \bar{d}_y & + \quad \triangle \bar{S} & = & 0 \\
\triangle \bar{X} \bar{S}^0 & & + \quad \bar{X}^0 \triangle \bar{S} & = & \gamma \mu^0 I - \bar{X}^0 \bar{S}^0 \\
\multicolumn{5}{l}{\bar{X}^0 + \triangle \bar{X} \succ 0, \bar{S}^0 + \triangle \bar{S} \succ 0}
\end{array}
$$

- $L = (X^0)^{1/2}$: $\quad \bar{X}^0 = I \quad \Rightarrow \quad \bar{X}^0 + \triangle \bar{X} \succ 0, \forall \|\triangle \bar{X}\|_F < 1$ (Primal)
- $L = (S^0)^{-1/2}$: $\quad \bar{S}^0 = I \quad \Rightarrow \quad \bar{S}^0 + \triangle \bar{S} \succ 0, \forall \|\triangle \bar{S}\|_F < 1$ (Dual)
- $LL^T = (S^0)^{-1/2}[(S^0)^{1/2} X^0 (S^0)^{1/2}]^{1/2} (S^0)^{-1/2}$:
  $V^0 = \bar{X}^0 = \bar{S}^0$ (Primal-dual)

Primal-Dual Interior-Point Method for SDP

$LL^T = (S^0)^{-\frac{1}{2}}[(S^0)^{\frac{1}{2}}X^0(S^0)^{\frac{1}{2}}]^{\frac{1}{2}}(S^0)^{-\frac{1}{2}}$:

$$\begin{bmatrix} \bar{\mathcal{A}} & 0 & 0 \\ 0 & \bar{\mathcal{A}}^* & I \\ I & 0 & I \end{bmatrix} \begin{bmatrix} \triangle\bar{X} \\ \bar{d}_y \\ \triangle\bar{S} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \gamma\mu^0(V^0)^{-1} - V^0 \end{bmatrix}$$

$$\bar{X}^0 + \triangle\bar{X} \succ 0, \bar{S}^0 + \triangle\bar{S} \succ 0$$

One can solve

$$\bar{\mathcal{A}}\bar{\mathcal{A}}^*\bar{d}_y = -\bar{\mathcal{A}}(\gamma\mu^0(V^0)^{-1} - V^0)$$

And then solve $\triangle\bar{S}$ and $\triangle\bar{X}$:

$$\begin{aligned} \triangle\bar{S} &= -\bar{\mathcal{A}}^*\bar{d}_y \\ \triangle\bar{X} &= -\triangle\bar{S} + \gamma\mu^0(V^0)^{-1} - V^0 \end{aligned}$$

Neighborhood of Central Path for LP

Notice that $\bar{x}^0 = \bar{s}^0 = v^0$

- Distance to central path: $u >_{\mathbb{R}^n_+} 0$

$$\delta(u) = \|e - \frac{n}{u^T u} \Lambda_u u\|_2$$

- Neighborhood of the central path

$$\mathcal{N}_2(\beta) = \{u | u >_{\mathbb{R}^n_+} 0, \delta(u) \leq \beta\}$$

$$\mathcal{N}_{-\infty}(\beta) = \{u | u >_{\mathbb{R}^n_+} 0, \Lambda_u u \geq_{\mathbb{R}^n_+} (1 - \beta) \frac{u^T u}{n} e\}$$

Examples: $\mathcal{N}_2(\frac{1}{2})$ and $\mathcal{N}_{-\infty}(\frac{1}{2})$

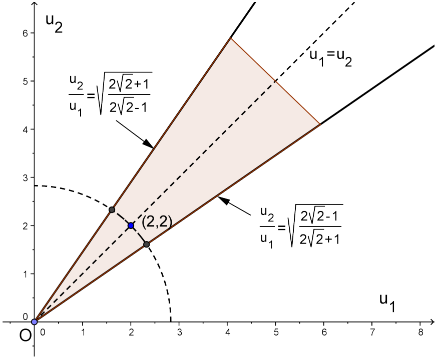

Figure: Neighborhood $\mathcal{N}_2(\frac{1}{2})$

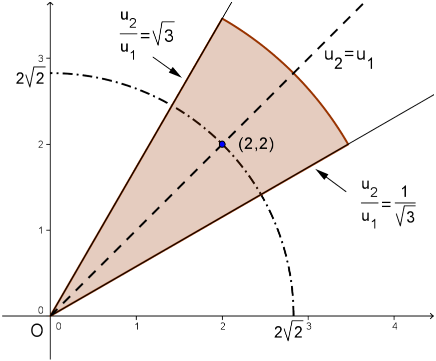

Figure: Neighborhood $\mathcal{N}_{-\infty}(\frac{1}{2})$

Finding Step Length for LP

$$\begin{matrix} \bar{x}^0 + \alpha \bar{d}_x \\ \bar{s}^0 + \alpha \bar{d}_s \end{matrix} \xrightarrow{\text{scaling back}} \begin{bmatrix} x^1 \\ s^1 \end{bmatrix} \xrightarrow{\text{new scaling}} v^1 = \bar{x}^1 = \bar{s}^1$$

Lemma

For any $0 \leq \alpha \leq 1$,

$$\mu^1 = \frac{\|v^1\|_2^2}{n} = \frac{(\bar{x}^0 + \alpha \bar{d}_x)^T (\bar{s}^0 + \alpha \bar{d}_s)}{n} = (1 - \alpha + \gamma\alpha)\mu^0$$

Finding Step Length for LP

Lemma

If $\delta(v^0) < 1$ and α satisfies $\bar{x}^0 + \alpha \bar{d}_x >_{\mathbb{R}^n_+} 0$ and $\bar{s}^0 + \alpha \bar{d}_s >_{\mathbb{R}^n_+} 0$, then

$$(1 - \alpha + \gamma\alpha)\delta(v^1) \leq (1-\alpha)\delta(v^0) + \frac{\alpha^2}{2}\left(\frac{\gamma^2\delta(v^0)^2}{1-\delta(v^0)} + n(1-\gamma)^2\right)$$

Proof:

$$\begin{aligned}
\mu^1\delta(v^1) &= \mu^1\|e - \frac{1}{\mu^1}\Lambda_{v^1}v^1\|_2 \\
&= \|(1-\alpha+\gamma\alpha)\mu^0 e - \Lambda_{(v^0+\alpha\bar{d_x})}(v^0 + \alpha\bar{d_s})\|_2 \\
&\leq \|(1-\alpha)\mu^0(e - \frac{1}{\mu^0}\Lambda_{v^0}v^0)\|_2 + \|\alpha^2\Lambda_{\bar{d_x}}\bar{d_s}\|_2 \\
&\leq (1-\alpha)\mu^0\delta(v^0) + \frac{\alpha^2}{2}\|\bar{d_x} + \bar{d_s}\|_2^2 \\
&= (1-\alpha)\mu^0\delta(v^0) + \frac{\alpha^2}{2}(\gamma^2\|\mu^0\Lambda_{v^0}^{-1}e - v^0\|_2^2 + (1-\gamma)^2 n\mu^0) \\
&\leq (1-\alpha)\mu^0\delta(v^0) + \frac{\alpha^2}{2}(\frac{\mu^0\gamma^2\delta(v^0)^2}{1-\delta(v^0)} + (1-\gamma)^2 n\mu^0)
\end{aligned}$$

Finding Step Length for LP

Lemma

If $v^0 \in \mathcal{N}_2(\beta)$ with $\beta = \frac{1}{2}$, $\gamma = \frac{1}{1+1/\sqrt{2n}}$ and $\alpha = 1$, then

(i) $v^1 \in \mathcal{N}_2(\beta)$

(ii) $x^1 \bullet s^1 = \bar{x}^1 \bullet \bar{s}^1 = \|v^1\|_2^2 = \gamma\mu^0$

Path Following Algorithm for LP

Step 1: (Initialization)
$\epsilon > 0$, (x^0, y^0, s^0) with $v^0 \in \mathcal{N}(\beta)$, where $\beta = \frac{1}{2}$.
Set $k = 0$, $\gamma = \frac{1}{1+1/\sqrt{2n}}$, and $\alpha = 1$.

Step 2: Solve the Newton system introduced above and get (d_x, d_y, d_s).
Set

$$\begin{cases} x^{k+1} = x^k + \alpha d_x \\ y^{k+1} = y^k + \alpha d_y \\ s^{k+1} = s^k + \alpha d_s \end{cases}$$

with $v^{k+1} = \Lambda_{x^{k+1}}^{1/2} \Lambda_{s^{k+1}}^{1/2} e$.
Set $k = k + 1$.

Step 3: If $x^k \bullet s^k < \epsilon$, stop. Otherwise, go to Step 2.

Complexity for LP

Theorem

Given the above settings, we have

(i) $v^k \in \mathcal{N}_2(\beta)$, $k = 0, 1, 2, \ldots$.

(ii) The algorithms stops in

$$O(\sqrt{n} \log \frac{x^0 \bullet s^0}{\epsilon})$$

steps and output a primal-dual solution satisfying

$$x^k \bullet s^k < \epsilon$$

Neighborhood of Central Path for SDP

Notice that $\bar{X}^0 = \bar{S}^0 = V^0$

- Distance to central path: $U \in \mathcal{S}_+^n$ and $U \succ 0$

$$\delta(U) = \|I - \frac{n}{I \bullet U^2} U^2\|_F, \text{ with } U^2 = UU$$

- Neighborhood of the central path

$$\mathcal{N}_2(\beta) = \{U | U \succ 0, \delta(U) \le \beta\}$$

$$\mathcal{N}_{-\infty}(\beta) = \{U | U \succ 0, U^2 \succeq (1-\beta) \frac{I \bullet U^2}{n} I\}$$

Finding Step Length for SDP

$$\begin{matrix} \bar{X}^0 + \alpha \triangle \bar{X} \\ \bar{S}^0 + \alpha \triangle \bar{S} \end{matrix} \xrightarrow{\text{scaling back}} \begin{bmatrix} X^1 \\ S^1 \end{bmatrix} \xrightarrow{\text{new scaling}} V^1 = \bar{X}^1 = \bar{S}^1$$

Lemma

For any $0 \leq \alpha \leq 1$,

$$\mu^1 = \frac{\|V^1\|_F^2}{n} = \frac{\text{tr}[(\bar{X}^0 + \alpha \triangle \bar{X})(\bar{S}^0 + \alpha \triangle \bar{S})]}{n} = (1 - \alpha + \gamma\alpha)\mu^0.$$

Finding Step Length for SDP

Lemma

For any square matrix U, we have

$$\mathrm{tr}(U^2) = \|\frac{U+U^T}{2}\|_F^2 - \|\frac{U-U^T}{2}\|_F^2 \leq \|\frac{U+U^T}{2}\|_F^2$$

Lemma

Suppose $\delta(V^0) < 1$ and $\alpha \geq 0$ satisfies $\bar{X}^0 + \alpha\triangle\bar{X} \succ 0$ and $\bar{S}^0 + \alpha\triangle\bar{S} \succ 0$. Let

$$W = \frac{(\bar{X}^0 + \alpha\triangle\bar{X})(\bar{S}^0 + \alpha\triangle\bar{S}) + ((\bar{X}^0 + \alpha\triangle\bar{X})(\bar{S}^0 + \alpha\triangle\bar{S}))^T}{2}$$

then

$$W = (1-\alpha)(V^0)^2 + \alpha\gamma\mu^0 I + \alpha^2\frac{\triangle\bar{X}\triangle\bar{S} + \triangle\bar{S}\triangle\bar{X}}{2}$$

and

$$\delta(V^1)^2 \leq \|I - \frac{1}{\mu^1}W\|_F^2$$

Finding Step Length for SDP

Lemma

Suppose $\delta(V^0) < 1$ and $\alpha \geq 0$ satisfies $\bar{X}^0 + \alpha\triangle\bar{X} \succ 0$ and $\bar{S}^0 + \alpha\triangle\bar{S} \succ 0$. Then

$$(1-\alpha+\gamma\alpha)\delta(V^1) \leq (1-\alpha)\delta(V^0) + \frac{\alpha^2}{2}\left(\frac{\gamma^2\delta(V^0)^2}{1-\delta(V^0)} + n(1-\gamma)^2\right)$$

Proof

$$\begin{aligned}
\mu^1\delta(V^1) &\leq (1-\alpha)\mu^0\delta(V^0) + \alpha^2\|\tfrac{\triangle\bar{X}\triangle\bar{S}+\triangle\bar{S}\triangle\bar{X}}{2}\|_F \\
&\leq (1-\alpha)\mu^0\delta(V^0) + \tfrac{\alpha^2}{2}\|\triangle\bar{X} + \triangle\bar{S}\|_F^2 \\
&= (1-\alpha)\mu^0\delta(V^0) + \tfrac{\alpha^2}{2}(\gamma^2\|\mu^0(V^0)^{-1} - V^0\|_F^2 + (1-\gamma)^2 n\mu^0) \\
&\leq (1-\alpha)\mu^0\delta(V^0) + \tfrac{\alpha^2\mu^0}{2}\left(\tfrac{\gamma^2\delta(V^0)^2}{1-\delta(V^0)} + n(1-\gamma)^2\right)
\end{aligned}$$

Finding Step Length for SDP

Lemma

If $V^0 \in \mathcal{N}_2(\beta)$ with $\beta = \frac{1}{2}$, $\gamma = \frac{1}{1+1/\sqrt{2n}}$ and $\alpha = 1$, then

(i) $V^1 \in \mathcal{N}_2(\beta)$

(ii) $X^1 \bullet S^1 = \bar{X}^1 \bullet \bar{S}^1 = \|V^1\|_F^2 = \gamma\mu^0$

Path Following Algorithm for SDP

Step 1: (Initialization)
$\epsilon > 0$, (X^0, y^0, S^0) with $V^0 \in \mathcal{N}(\beta)$, where $\beta = \frac{1}{2}$.
Set $k = 0$, $\gamma = \frac{1}{1+1/\sqrt{2n}}$, and $\alpha = 1$.

Step 2: Solve the equation system introduced above and get $(\triangle X, d_y, \triangle S)$.
Set

$$\begin{cases} X^{k+1} = X^k + \alpha \triangle X \\ y^{k+1} = y^k + \alpha d_y \\ S^{k+1} = X^k + \alpha \triangle S \end{cases}$$

with $V^{k+1} = \bar{X}^{k+1} = \bar{S}^{k+1}$.
Set $k = k + 1$.

Step 3: If $X^k \bullet S^k < \epsilon$, stop. Otherwise, go to Step 2.

Complexity

Theorem

Given the above settings, we have

(i) $V^k \in \mathcal{N}_2(\beta)$, $k = 0, 1, 2, \ldots$.

(ii) The algorithms stops in

$$O(\sqrt{n} \log \frac{X^0 \bullet S^0}{\epsilon})$$

steps and output a primal-dual solution satisfying

$$X^k \bullet S^k < \epsilon$$

Example: Path Following Algorithm

$$
\begin{aligned}
\min \quad & x_1 + x_2 \\
s.t. \quad & x_1 + x_2 \leq 3 \\
& x_1 - x_2 \leq 1 \\
& x_2 \leq 2 \\
& x_1 \geq 0, x_2 \geq 0
\end{aligned}
$$

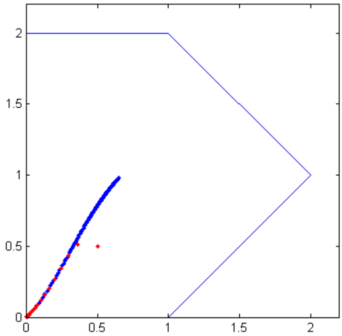

Figure: Path following algorithm with $\beta = 1/2$

Initialization and Improve the Performance

Initialization

- Big-M Method
- Two-Phase Method
- Self-Dual Embedding Method

Different Path-Following Methods

- Short Step Algorithm
- Long Step Algorithm
- Predictor-Corrector Algorithm
- Largest Step Algorithm

Reference: *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, edited by Wolkowicz H., Saigal R. and Vandenberghe L., Kluwer Academic Publisher: Norwell, MA USA 2000