Convex Optimization

(EE227A: UC Berkeley)

Lecture 6
(Conic optimization)

07 Feb, 2013

Suvrit Sra

Organizational Info

- Quiz coming up on 19th Feb.
- Project teams by 19th Feb
- Good if you can mix your research with class projects
- More info in a few days

Mini Challenge

Kummer's **confluent hypergeometric function**

$$M(a,c,x) := \sum_{j\geq 0} \frac{(a)_j}{(c)_j} \frac{x^j}{j!}, \qquad a,c,x \in \mathbb{R},$$

and $(a)_0 = 1$, $(a)_j = a(a+1)\cdots(a+j-1)$ is the **rising-factorial**.

Mini Challenge

Kummer's **confluent hypergeometric function**

$$M(a,c,x) := \sum_{j \geq 0} \frac{(a)_j}{(c)_j} \frac{x^j}{j!}, \qquad a, c, x \in \mathbb{R},$$

and $(a)_0 = 1$, $(a)_j = a(a+1)\cdots(a+j-1)$ is the **rising-factorial**.

Claim: Let $c > a > 0$ and $x \geq 0$. Then the function

$$h_{a,c}(\mu; x) := \mu \mapsto \frac{\Gamma(a+\mu)}{\Gamma(c+\mu)} M(a+\mu, c+\mu, x)$$

is strictly log-convex on $[0, \infty)$ (**note that h is a function of μ**).

Recall: $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt$ is the **Gamma function** (which is known to be log-convex for $x \geq 1$; see also Exercise 3.52 of BV).

LP formulation

Write $\min \quad \|Ax - b\|_1$ as a linear program.

$$\min \qquad \|Ax - b\|_1 \qquad x \in \mathbb{R}^n$$

$$\min \qquad \sum_i |a_i^T x - b_i|$$

$$\min_{x,t} \qquad \sum_i t_i, \qquad |a_i^T x - b_i| \leq t_i, \quad i = 1, \ldots, m.$$

$$\min_{x,t} \qquad \mathbf{1}^T t, \qquad -t_i \leq a_i^T x - b_i \leq t_i, \quad i = 1, \ldots, m.$$

LP formulation

Write $\min \quad \|Ax - b\|_1$ as a linear program.

$$
\begin{array}{lll}
\min & \|Ax-b\|_1 & x \in \mathbb{R}^n \\
\min & \sum_i |a_i^T x - b_i| & \\
\min\limits_{x,t} & \sum_i t_i, & |a_i^T x - b_i| \leq t_i, \quad i = 1, \ldots, m. \\
\min\limits_{x,t} & \mathbf{1}^T t, & -t_i \leq a_i^T x - b_i \leq t_i, \quad i = 1, \ldots, m.
\end{array}
$$

Exercise: Recast $\|Ax - b\|_2^2 + \lambda \|Bx\|_1$ as a QP.

Cone programs – overview

- Last time we briefly saw LP, QP, SOCP, SDP

Cone programs – overview

- Last time we briefly saw LP, QP, SOCP, SDP

LP (standard form)

$$\min \quad f^T x \quad \text{s.t. } Ax = b, \quad x \geq 0.$$

Feasible set $\mathcal{X} = \{x \mid Ax = b\} \cap \mathbb{R}^n_+$ (nonneg orthant)

Cone programs – overview

- Last time we briefly saw LP, QP, SOCP, SDP

LP (standard form)

$$\min \quad f^T x \quad \text{s.t. } Ax = b, \quad x \geq 0.$$

Feasible set $\mathcal{X} = \{x \mid Ax = b\} \cap \mathbb{R}^n_+$ (nonneg orthant)

Input data: (A, b, c)
Structural constraints: $x \geq 0$.

Cone programs – overview

- Last time we briefly saw LP, QP, SOCP, SDP

LP (standard form)

$$\min \quad f^T x \quad \text{s.t. } Ax = b, \quad x \geq 0.$$

Feasible set $\mathcal{X} = \{x \mid Ax = b\} \cap \mathbb{R}^n_+$ (nonneg orthant)

Input data: (A, b, c)
Structural constraints: $x \geq 0$.

How should we generalize this model?

Cone programs – overview

- Replace linear map $x \mapsto Ax$ by a nonlinear map?

Cone programs – overview

- Replace linear map $x \mapsto Ax$ by a nonlinear map?
- Quickly becomes nonconvex, potentially intractable

Cone programs – overview

- Replace linear map $x \mapsto Ax$ by a nonlinear map?
- Quickly becomes nonconvex, potentially intractable

 Generalize structural constraint $\mathbb{R}^n_+$

Cone programs – overview

- Replace linear map $x \mapsto Ax$ by a nonlinear map?
- Quickly becomes nonconvex, potentially intractable

 Generalize structural constraint $\mathbb{R}^n_+$

♣ Replace nonneg orthant by a convex cone $\mathcal{K}$;

Cone programs – overview

- Replace linear map $x \mapsto Ax$ by a nonlinear map?
- Quickly becomes nonconvex, potentially intractable

Generalize structural constraint $\mathbb{R}^n_+$

♣ Replace nonneg orthant by a convex cone $\mathcal{K}$;

♣ Replace $\geq$ by conic inequality $\succeq$

Cone programs – overview

- Replace linear map $x \mapsto Ax$ by a nonlinear map?
- Quickly becomes nonconvex, potentially intractable

Generalize structural constraint $\mathbb{R}^n_+$

- ♣ Replace nonneg orthant by a convex cone $\mathcal{K}$;
- ♣ Replace $\geq$ by conic inequality $\succeq$
- ♣ Nesterov and Nemirovski developed nice theory in late 80s
- ♣ Rich class of cones for which cone programs are tractable

Conic inequalities

- We are looking for “good” vector inequalities $\succeq$ on $\mathbb{R}^n$

Conic inequalities

- We are looking for "good" vector inequalities $\succeq$ on $\mathbb{R}^n$
- Characterized by the set

$$\mathcal{K} := \{x \in \mathbb{R}^n \mid x \succeq 0\}$$

of vector **nonneg w.r.t.** $\succeq$

$$x \succeq y \quad \Leftrightarrow \quad x - y \succeq 0 \quad \Leftrightarrow \quad x - y \in \mathcal{K}.$$

Conic inequalities

- We are looking for "good" vector inequalities $\succeq$ on $\mathbb{R}^n$
- Characterized by the set

$$\mathcal{K} := \{x \in \mathbb{R}^n \mid x \succeq 0\}$$

  of vector **nonneg w.r.t.** $\succeq$

$$x \succeq y \quad \Leftrightarrow \quad x - y \succeq 0 \quad \Leftrightarrow \quad x - y \in \mathcal{K}.$$

- **Necessary and sufficient** condition for a set $\mathcal{K} \subset \mathbb{R}^n$ to define a useful vector inequality $\succeq$ is: it should be a **nonempty, pointed cone**.

Cone programs – inequalities

- $\mathcal{K}$ is nonempty: $\mathcal{K} \neq \emptyset$
- $\mathcal{K}$ is closed wrt addition: $x, y \in \mathcal{K} \implies x + y \in \mathcal{K}$
- $\mathcal{K}$ closed wrt noneg scaling: $x \in \mathcal{K}$, $\alpha \geq 0 \implies \alpha x \in \mathcal{K}$
- $\mathcal{K}$ is pointed: $x, -x \in \mathcal{K} \implies x = 0$

Cone programs – inequalities

- $\mathcal{K}$ is nonempty: $\mathcal{K} \neq \emptyset$
- $\mathcal{K}$ is closed wrt addition: $x, y \in \mathcal{K} \implies x + y \in \mathcal{K}$
- $\mathcal{K}$ closed wrt noneg scaling: $x \in \mathcal{K}$, $\alpha \geq 0 \implies \alpha x \in \mathcal{K}$
- $\mathcal{K}$ is pointed: $x, -x \in \mathcal{K} \implies x = 0$

Cone inequality

$$x \succeq_{\mathcal{K}} y \iff x - y \in \mathcal{K}$$
$$x \succ_{\mathcal{K}} y \iff x - y \in \text{int}(\mathcal{K}).$$

Conic inequalities

- Cone underlying standard coordinatewise vector inequalities:

$$x \geq y \quad \Leftrightarrow \quad x_i \geq y_i \quad \Leftrightarrow \quad x_i - y_i \geq 0,$$

is the **nonegative orthant** $\mathbb{R}^n_+$.

Conic inequalities

- Cone underlying standard coordinatewise vector inequalities:

$$x \geq y \quad \Leftrightarrow \quad x_i \geq y_i \quad \Leftrightarrow \quad x_i - y_i \geq 0,$$

  is the **nonegative orthant** $\mathbb{R}^n_+$.
- Two more important properties that $\mathbb{R}^n_+$ has as a cone:
  - It is **closed** $\left\{x^i \in \mathbb{R}^n_+\right\} \to x \implies x \in \mathbb{R}^n_+$
  - It has nonempty interior (contains Euclidean ball of **positive** radius)

Conic inequalities

- Cone underlying standard coordinatewise vector inequalities:

$$x \geq y \quad \Leftrightarrow \quad x_i \geq y_i \quad \Leftrightarrow \quad x_i - y_i \geq 0,$$

  is the **nonegative orthant** $\mathbb{R}^n_+$.
- Two more important properties that $\mathbb{R}^n_+$ has as a cone:
  - It is **closed** $\{x^i \in \mathbb{R}^n_+\} \to x \implies x \in \mathbb{R}^n_+$
  - It has nonempty interior (contains Euclidean ball of **positive** radius)
- We'll require our cones to also satisfy these two properties.

Conic optimization problems

Standard form cone program

$$\min \quad f^T x \quad \text{s.t. } Ax = b,\ x \in \mathcal{K}$$

$$\min \quad f^T x \quad \text{s.t. } Ax \preceq_{\mathcal{K}} b.$$

Conic optimization problems

Standard form cone program

$$\min \quad f^T x \quad \text{s.t. } Ax = b,\ x \in \mathcal{K}$$

$$\min \quad f^T x \quad \text{s.t. } Ax \preceq_{\mathcal{K}} b.$$

- ♣ The nonnegative orthant $\mathbb{R}^n_+$
- ♣ The second order cone $\mathcal{Q}^n := \{(x, t) \in \mathbb{R}^n \mid \|x\|_2 \leq t\}$
- ♣ The semidefinite cone: $\mathcal{S}^n_+ := \left\{X = X^T \succeq 0\right\}$.

Conic optimization problems

Standard form cone program

$$\min \quad f^T x \quad \text{s.t. } Ax = b,\ x \in \mathcal{K}$$

$$\min \quad f^T x \quad \text{s.t. } Ax \preceq_{\mathcal{K}} b.$$

- ♣ The nonnegative orthant $\mathbb{R}^n_+$
- ♣ The second order cone $\mathcal{Q}^n := \{(x, t) \in \mathbb{R}^n \mid \|x\|_2 \leq t\}$
- ♣ The semidefinite cone: $\mathcal{S}^n_+ := \left\{X = X^T \succeq 0\right\}$.
- ♣ Other cones $\mathcal{K}$ given by Cartesian products of these

Conic optimization problems

Standard form cone program

$$\min \quad f^T x \quad \text{s.t. } Ax = b,\ x \in \mathcal{K}$$
$$\min \quad f^T x \quad \text{s.t. } Ax \preceq_{\mathcal{K}} b.$$

- ♣ The nonnegative orthant $\mathbb{R}^n_+$
- ♣ The second order cone $\mathcal{Q}^n := \{(x,t) \in \mathbb{R}^n \mid \|x\|_2 \leq t\}$
- ♣ The semidefinite cone: $\mathcal{S}^n_+ := \left\{X = X^T \succeq 0\right\}$.
- ♣ Other cones $\mathcal{K}$ given by Cartesian products of these
- ♣ These cones are "nice":
- ♣ LP, QP, SOCP, SDP: all are **cone programs**

Conic optimization problems

Standard form cone program

$$\min \quad f^T x \quad \text{s.t. } Ax = b,\ x \in \mathcal{K}$$
$$\min \quad f^T x \quad \text{s.t. } Ax \preceq_{\mathcal{K}} b.$$

- ♣ The nonnegative orthant $\mathbb{R}^n_+$
- ♣ The second order cone $\mathcal{Q}^n := \{(x,t) \in \mathbb{R}^n \mid \|x\|_2 \leq t\}$
- ♣ The semidefinite cone: $\mathcal{S}^n_+ := \left\{X = X^T \succeq 0\right\}$.
- ♣ Other cones $\mathcal{K}$ given by Cartesian products of these
- ♣ These cones are "nice":
- ♣ LP, QP, SOCP, SDP: all are **cone programs**
- ♣ Can treat them theoretically in a uniform way (roughly)

Conic optimization problems

Standard form cone program

$$\min \quad f^T x \quad \text{s.t. } Ax = b,\ x \in \mathcal{K}$$

$$\min \quad f^T x \quad \text{s.t. } Ax \preceq_{\mathcal{K}} b.$$

- ♣ The nonnegative orthant $\mathbb{R}^n_+$
- ♣ The second order cone $\mathcal{Q}^n := \{(x,t) \in \mathbb{R}^n \mid \|x\|_2 \le t\}$
- ♣ The semidefinite cone: $\mathcal{S}^n_+ := \left\{X = X^T \succeq 0\right\}$.
- ♣ Other cones $\mathcal{K}$ given by Cartesian products of these
- ♣ These cones are "nice":
- ♣ LP, QP, SOCP, SDP: all are **cone programs**
- ♣ Can treat them theoretically in a uniform way (roughly)
- ♣ Not all cones are nice!

Cone programs – tough case

Copositive cone

Def. Let $CP_n := \left\{ A \in \mathbb{S}^{n \times n} \mid x^T A x \geq 0,\ \forall x \geq 0 \right\}$.

Exercise: Verify that CP_n is a convex cone.

Cone programs – tough case

Copositive cone

Def. Let $CP_n := \{A \in \mathbb{S}^{n\times n} \mid x^T A x \geq 0,\ \forall x \geq 0\}$.

Exercise: Verify that CP_n is a convex cone.

If someone told you convex is "easy" ... they lied!

Cone programs – tough case

Copositive cone

Def. Let $CP_n := \{A \in \mathbb{S}^{n\times n} \mid x^T A x \geq 0, \ \forall x \geq 0\}$.

Exercise: Verify that CP_n is a convex cone.

If someone told you convex is "easy" ... they lied!

- Testing membership in CP_n is co-NP complete.
  (Deciding whether given matrix is not copositive is NP-complete.)
- Copositive cone programming: NP-Hard

Cone programs – tough case

Copositive cone

Def. Let $CP_n := \left\{A \in \mathbb{S}^{n\times n} \mid x^T A x \geq 0,\ \forall x \geq 0\right\}$.

Exercise: Verify that CP_n is a convex cone.

If someone told you convex is "easy" ... they lied!

- Testing membership in CP_n is co-NP complete.
  (Deciding whether given matrix is not copositive is NP-complete.)
- Copositive cone programming: NP-Hard

Exercise: Verify that the following matrix is copositive:

$$A := \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix}.$$

SOCP in conic form

$$\min \quad f^T x \quad \text{s.t.} \quad \|A_i x + b_i\|_2 \le c_i^T x + d_i \qquad i = 1, \ldots, m$$

Let $A_i \in \mathbb{R}^{n_i \times n}$; so $A_i x + b_i \in \mathbb{R}^{n_i}$.

SOCP in conic form

$$\min \quad f^T x \quad \text{s.t.} \quad \|A_i x + b_i\|_2 \le c_i^T x + d_i \qquad i = 1, \ldots, m$$

Let $A_i \in \mathbb{R}^{n_i \times n}$; so $A_i x + b_i \in \mathbb{R}^{n_i}$.

$$\mathcal{K} = \mathcal{Q}^{n_1} \times \mathcal{Q}^{n_2} \times \cdots \times \mathcal{Q}^{n_m}, \quad A = \begin{bmatrix} \begin{bmatrix} -A_1 \\ -c_1^T \end{bmatrix} \\ \begin{bmatrix} -A_2 \\ -c_2^T \end{bmatrix} \\ \vdots \\ \begin{bmatrix} -A_m \\ -c_m^T \end{bmatrix} \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ d_1 \\ b_2 \\ d_2 \\ \vdots \\ b_m \\ d_m \end{bmatrix}.$$

SOCP in conic form

$$\min \quad f^T x \quad Ax \preceq_{\mathcal{K}} b$$

SOCP representation

Exercise: Let $0 \prec Q = LL^T$, then show that

$$x^T Q x + b^T x + c \le 0 \Leftrightarrow \|L^T x + L^{-1} b\|_2 \le \sqrt{b^T Q^{-1} b - c}$$

SOCP representation

Exercise: Let $0 \prec Q = LL^T$, then show that

$$x^T Q x + b^T x + c \leq 0 \Leftrightarrow \|L^T x + L^{-1} b\|_2 \leq \sqrt{b^T Q^{-1} b - c}$$

Rotated second-order cone

$$\mathcal{Q}_r^n := \left\{ (x, y, z) \in \mathbb{R}^{n+1} \mid \|x\|_2 \leq \sqrt{yz}, y \geq 0, z \geq 0 \right\}.$$

SOCP representation

Exercise: Let $0 \prec Q = LL^T$, then show that

$$x^TQx + b^Tx + c \le 0 \Leftrightarrow \|L^Tx + L^{-1}b\|_2 \le \sqrt{b^TQ^{-1}b - c}$$

Rotated second-order cone

$$\mathcal{Q}_r^n := \left\{(x, y, z) \in \mathbb{R}^{n+1} \mid \|x\|_2 \le \sqrt{yz}, y \ge 0, z \ge 0\right\}.$$

Convert into standard SOC (verify!)

$$\left\| \begin{bmatrix} 2x \\ y - z \end{bmatrix} \right\|_2 \le (y + z) \quad \Longleftrightarrow \|x\|_2 \le \sqrt{yz}.$$

Exercise: Rewrite the constraint $x^TQx \le t$, where **both x and t are variables** using the rotated second order cone.

Convex QP as SOCP

$$\min \quad x^T Q x + c^T x \quad \text{s.t. } Ax = b.$$

Convex QP as SOCP

$$\min \quad x^T Q x + c^T x \quad \text{s.t. } Ax = b.$$

$$\begin{aligned} \min_{x,t} \quad & c^T x + t \\ \text{s.t.} \quad & Ax = b, \quad x^T Q x \le t. \end{aligned}$$

Convex QP as SOCP

$$\min \quad x^T Q x + c^T x \quad \text{s.t. } Ax = b.$$

$$\begin{aligned} \min_{x,t} \quad & c^T x + t \\ \text{s.t.} \quad & Ax = b, \quad x^T Q x \le t. \end{aligned}$$

$$\begin{aligned} \min_{x,t} \quad & c^T x + t \\ \text{s.t.} \quad & Ax = b, \quad (2L^T x, t, 1) \in \mathcal{Q}_r^n. \end{aligned}$$

Since, $x^T Q x = x^T L L^T x = \|L^T x\|_2^2$

Convex QCQPs as SOCP

Quadratically Constrained QP

$$\min \quad q_0(x) \quad \text{s.t. } q_i(x) \leq 0, \quad i = 1, \ldots, m$$

where each $q_i(x) = x^T P_i x + b_i^T x + c_i$ is a convex quadratic.

Convex QCQPs as SOCP

Quadratically Constrained QP

$$\min \quad q_0(x) \quad \text{s.t. } q_i(x) \leq 0, \quad i = 1, \ldots, m$$

where each $q_i(x) = x^T P_i x + b_i^T x + c_i$ is a convex quadratic.

Exercise: Show how QCQPs can be cast at SOCPs using $\mathcal{Q}_r^n$
Hint: See Lecture 5!

Convex QCQPs as SOCP

Quadratically Constrained QP

$$\min \quad q_0(x) \quad \text{s.t. } q_i(x) \leq 0, \quad i = 1, \ldots, m$$

where each $q_i(x) = x^T P_i x + b_i^T x + c_i$ is a convex quadratic.

Exercise: Show how QCQPs can be cast at SOCPs using $\mathcal{Q}_r^n$
Hint: See Lecture 5!

Exercise: Explain why we cannot cast SOCPs as QCQPs. That is, why cannot we simply use the equivalence

$$\|Ax + b\|_2 \leq c^T x + d \Leftrightarrow \|Ax + b\|_2^2 \leq (c^T x + d)^2, \; c^T x + d \geq 0.$$

Hint: Look carefully at the inequality!

Robust LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i \\ \text{where} \quad & \mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}. \end{aligned}$$

Robust LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \le b_i \quad \forall a_i \in \mathcal{E}_i \\ \text{where} \quad & \mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}. \end{aligned}$$

Robust half-space constraint:

- Wish to ensure $a_i^T x \le b_i$ holds irrespective of which a_i we pick from the *uncertainty set* $\mathcal{E}_i$.

Robust LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \le b_i \quad \forall a_i \in \mathcal{E}_i \\ \text{where} \quad & \mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}. \end{aligned}$$

Robust half-space constraint:

- Wish to ensure $a_i^T x \le b_i$ holds irrespective of which a_i we pick from the *uncertainty set* $\mathcal{E}_i$. This happens, if $b_i \ge \sup_{a_i \in \mathcal{E}_i} a_i^T x$.

Robust LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i \\ \text{where} \quad & \mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}. \end{aligned}$$

Robust half-space constraint:

- Wish to ensure $a_i^T x \leq b_i$ holds irrespective of which a_i we pick from the *uncertainty set* $\mathcal{E}_i$. This happens, if $b_i \geq \sup_{a_i \in \mathcal{E}_i} a_i^T x$.

$$\sup_{\|u\|_2 \leq 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2.$$

Robust LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \le b_i \quad \forall a_i \in \mathcal{E}_i \\ \text{where} \quad & \mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}. \end{aligned}$$

Robust half-space constraint:

- Wish to ensure $a_i^T x \le b_i$ holds irrespective of which a_i we pick from the *uncertainty set* $\mathcal{E}_i$. This happens, if $b_i \ge \sup_{a_i \in \mathcal{E}_i} a_i^T x$.

$$\sup_{\|u\|_2 \le 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2.$$

- We used the fact that $\sup_{\|u\|_2 \le 1} u^T v = \|v\|_2$ (recall dual-norms)

Robust LP

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \le b_i \quad \forall a_i \in \mathcal{E}_i \\ \text{where} \quad & \mathcal{E}_i := \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\}. \end{aligned}$$

Robust half-space constraint:

- Wish to ensure $a_i^T x \le b_i$ holds irrespective of which a_i we pick from the *uncertainty set* $\mathcal{E}_i$. This happens, if $b_i \ge \sup_{a_i \in \mathcal{E}_i} a_i^T x$.

$$\sup_{\|u\|_2 \le 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2.$$

- We used the fact that $\sup_{\|u\|_2 \le 1} u^T v = \|v\|_2$ (recall dual-norms)

SOCP formulation

$$\min \quad c^T x, \qquad \text{s.t.} \quad \bar{a}_i^T x + \|P_i^T x\|_2 \le b_i, \quad i = 1, \ldots, m.$$

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Standard form

- Think of x as a matrix variable X

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Standard form

- Think of x as a matrix variable X
- Wlog we may assume $\mathcal{K} = \mathcal{S}_+^n$ (Why?)

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Standard form

- Think of x as a matrix variable X
- Wlog we may assume $\mathcal{K} = \mathcal{S}_+^n$ (Why?)
- Say $\mathcal{K} = \mathcal{S}_+^{n_1} \times \mathcal{S}_+^{n_2}$

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Standard form

- Think of x as a matrix variable X
- Wlog we may assume $\mathcal{K} = \mathcal{S}_+^n$ (Why?)
- Say $\mathcal{K} = \mathcal{S}_+^{n_1} \times \mathcal{S}_+^{n_2}$
- The condition $(X_1, X_2) \in \mathcal{K} \Leftrightarrow X :=$ Diag$(X_1, X_2) \in \mathcal{S}_+^{n_1+n_2}$

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Standard form

- Think of x as a matrix variable X
- Wlog we may assume $\mathcal{K} = \mathcal{S}_+^n$ (Why?)
- Say $\mathcal{K} = \mathcal{S}_+^{n_1} \times \mathcal{S}_+^{n_2}$
- The condition $(X_1, X_2) \in \mathcal{K} \Leftrightarrow X := \text{Diag}(X_1, X_2) \in \mathcal{S}_+^{n_1+n_2}$
- Thus, by imposing non diagonals blocks to be zero, we reduce to where $\mathcal{K}$ is the semidefinite cone itself (of suitable dimension).

Semidefinite Program (SDP)

Cone program (semidefinite)

$$\min \quad c^T x \quad \text{s.t. } Ax = b, \quad x \in \mathcal{K},$$

where $\mathcal{K}$ is a product of semidefinite cones.

Standard form

- Think of x as a matrix variable X
- Wlog we may assume $\mathcal{K} = \mathcal{S}_+^n$ (Why?)
- Say $\mathcal{K} = \mathcal{S}_+^{n_1} \times \mathcal{S}_+^{n_2}$
- The condition $(X_1, X_2) \in \mathcal{K} \Leftrightarrow X := \text{Diag}(X_1, X_2) \in \mathcal{S}_+^{n_1+n_2}$
- Thus, by imposing non diagonals blocks to be zero, we reduce to where $\mathcal{K}$ is the semidefinite cone itself (of suitable dimension).
- So, in matrix notation:
  - $c^T x \to \text{Tr}(CX)$;
  - $a_i^T x = b_i \to \text{Tr}(A_i X) = b_i$; and
  - $x \in \mathcal{K}$ as $X \succeq 0$.

SDP (conic form)

$$\min_{y \in \mathbb{R}^n} \quad c^T y$$
$$\text{s.t.} \quad A(y) := A_0 + y_1 A_1 + y_2 A_2 + \ldots + y_n A_n \succeq 0.$$

SDP

SDP (conic form)

$$\min_{y\in\mathbb{R}^n} \quad c^T y$$
$$\text{s.t.} \quad A(y) := A_0 + y_1 A_1 + y_2 A_2 + \ldots + y_n A_n \succeq 0.$$

Standard form SDP

$$\begin{aligned} \min \quad & \mathrm{Tr}(CX) \\ \text{s.t.} \quad & \mathrm{Tr}(A_i X) = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0. \end{aligned}$$

SDP

SDP (conic form)

$$\min_{y \in \mathbb{R}^n} \quad c^T y$$

$$\text{s.t.} \quad A(y) := A_0 + y_1 A_1 + y_2 A_2 + \ldots + y_n A_n \succeq 0.$$

Standard form SDP

$$\begin{aligned} \min \quad & \mathrm{Tr}(CX) \\ \text{s.t.} \quad & \mathrm{Tr}(A_i X) = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0. \end{aligned}$$

One can be converted into another

SDP – CVX form

```
cvx_begin
    variables X(n,n) symmetric;
    minimize( trace(C*X) )
    subject to
        for i = 1:m,
            trace(A{i}*X) == b(i);
        end
        X == semidefinite(n);
cvx_end
```

Note: remember **symmetric** and **semidefinite**

SDP representation – LP

LP as SDP

$$\min \quad f^T x \qquad \text{s.t. } Ax \leq b.$$

SDP representation – LP

LP as SDP

$$\min \quad f^T x \qquad \text{s.t. } Ax \le b.$$

SDP formulation

$$\begin{aligned} \min \quad & f^T x \\ \text{s.t.} \quad & A(x) := \text{diag}(b_1 - a_1^T x, \ldots, b_m - a_m^T x) \succeq 0. \end{aligned}$$

SDP representation – SOCP

SOCP as SDP

$$\min \quad f^T x \qquad \text{s.t. } \|A_i^T x + b_i\| \le c_i^T x + d_i, \quad i = 1, \dots, m.$$

SDP representation – SOCP

SOCP as SDP

$$\min \quad f^T x \qquad \text{s.t. } \|A_i^T x + b_i\| \leq c_i^T x + d_i, \quad i = 1, \ldots, m.$$

SDP formulation

$$\|x\|_2 \leq t \iff \begin{bmatrix} t & x^T \\ x & tI \end{bmatrix} \succeq 0$$

Schur-complements: $\begin{bmatrix} A & B^T \\ B & C \end{bmatrix} \succeq 0 \iff A - B^T C^{-1} B \succeq 0.$

SDP representation – SOCP

SOCP as SDP

$$\min \quad f^T x \qquad \text{s.t. } \|A_i^T x + b_i\| \leq c_i^T x + d_i, \quad i = 1, \ldots, m.$$

SDP formulation

$$\|x\|_2 \leq t \iff \begin{bmatrix} t & x^T \\ x & tI \end{bmatrix} \succeq 0$$

Schur-complements: $\begin{bmatrix} A & B^T \\ B & C \end{bmatrix} \succeq 0 \iff A - B^T C^{-1} B \succeq 0.$

$$\|A_i^T x + b_i\| \leq c_i^T x + d_i \iff \begin{bmatrix} c_i^T x + d_i & (A_i^T x + b_i)^T \\ A_i^T x + b_i & (c_i^T x + d_i) \end{bmatrix} \succeq 0.$$

SDP / LMI representation

Def. A set $S \subset \mathbb{R}^n$ is called **linear matrix inequality** (LMI) representable if there exist symmetric matrices $A_0, \ldots, A_n$ such that

$$S = \{x \in \mathbb{R}^n \mid A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}.$$

S is called **SDP representable** if it equals the projection of some higher dimensional LMI representable set.

SDP / LMI representation

Def. A set $S \subset \mathbb{R}^n$ is called **linear matrix inequality** (LMI) representable if there exist symmetric matrices $A_0, \ldots, A_n$ such that

$$S = \{x \in \mathbb{R}^n \mid A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}.$$

S is called **SDP representable** if it equals the projection of some higher dimensional LMI representable set.

♠ **Linear inequalities:** $Ax \leq b$ iff

$$\begin{bmatrix} b_1 - a_1^T x & & \\ & \ddots & \\ & & b_m - a_m^T x \end{bmatrix} \succeq 0.$$

SDP / LMI representation

♠ **Convex quadratics:** $x^T L L^T x + b^T x \leq c$ iff

$$\begin{bmatrix} I & L^T x \\ x^T L & c - b^T x \end{bmatrix} \succeq 0$$

SDP / LMI representation

♠ **Convex quadratics:** $x^T L L^T x + b^T x \leq c$ iff

$$\begin{bmatrix} I & L^T x \\ x^T L & c - b^T x \end{bmatrix} \succeq 0$$

♠ **Eigenvalue inequalities:**

$$\lambda_{\max}(X) \quad \leq \quad t, \quad \text{iff} \quad tI - X \succeq 0$$

SDP / LMI representation

♠ **Convex quadratics:** $x^T L L^T x + b^T x \leq c$ iff

$$\begin{bmatrix} I & L^T x \\ x^T L & c - b^T x \end{bmatrix} \succeq 0$$

♠ **Eigenvalue inequalities:**

$$\begin{aligned} \lambda_{\max}(X) &\leq t, \quad \text{iff} \quad tI - X \succeq 0 \\ \lambda_{\min}(X) &\geq t \quad \text{iff} X - tI \succeq 0 \\ \lambda_{\max} \text{ cvx} & \qquad \lambda_{\min} \text{concave.} \end{aligned}$$

SDP / LMI representation

♠ **Convex quadratics:** $x^T L L^T x + b^T x \leq c$ iff

$$\begin{bmatrix} I & L^T x \\ x^T L & c - b^T x \end{bmatrix} \succeq 0$$

♠ **Eigenvalue inequalities:**

$$\begin{array}{lcll} \lambda_{\max}(X) & \leq & t, & \text{iff} \quad tI - X \succeq 0 \\ \lambda_{\min}(X) & \geq & t & \text{iff} X - tI \succeq 0 \\ \lambda_{\max} \text{ cvx} & & & \lambda_{\min} \text{concave.} \end{array}$$

♠ **Matrix norm:** $X \in \mathbb{R}^{m \times n}$, $\|X\|_2 \leq t$ (i.e., $\sigma_{\max}(X) \leq t$) iff

$$\begin{bmatrix} tI_m & X \\ X^T & tI_n \end{bmatrix} \succeq 0.$$

Proof. $t^2 I \succeq X X^T \implies t^2 \geq \lambda_{\max}(X X^T) = \sigma^2_{\max}(X)$.

SDP / LMI representation

♠ **Sum of top eigenvalues:** For $X \in \mathbb{S}^n$, $\sum_{i=1}^{k} \lambda_i(X) \leq t$ iff

$$\begin{aligned} t - ks - \mathrm{Tr}(Z) &\geq 0 \\ Z &\succeq 0 \\ Z - X + sI &\succeq 0. \end{aligned}$$

SDP / LMI representation

♠ **Sum of top eigenvalues:** For $X \in \mathbb{S}^n$, $\sum_{i=1}^k \lambda_i(X) \le t$ iff

$$\begin{aligned} t - ks - \mathrm{Tr}(Z) &\ge 0 \\ Z &\succeq 0 \\ Z - X + sI &\succeq 0. \end{aligned}$$

Proof:
Suppose $\sum_{i=1}^k \lambda_i(X) \le t$.

SDP / LMI representation

♠ **Sum of top eigenvalues:** For $X \in \mathbb{S}^n$, $\sum_{i=1}^k \lambda_i(X) \leq t$ iff

$$\begin{aligned} t - ks - \mathrm{Tr}(Z) &\geq 0 \\ Z &\succeq 0 \\ Z - X + sI &\succeq 0. \end{aligned}$$

Proof:

Suppose $\sum_{i=1}^k \lambda_i(X) \leq t$. Then, choosing $s = \lambda_k$ and $Z = \mathrm{Diag}(\lambda_1 - s, \ldots, \lambda_k - s, 0, \ldots, 0)$, above LMIs hold.

SDP / LMI representation

♠ **Sum of top eigenvalues:** For $X \in \mathbb{S}^n$, $\sum_{i=1}^k \lambda_i(X) \leq t$ iff

$$\begin{aligned} t - ks - \mathrm{Tr}(Z) &\geq 0 \\ Z &\succeq 0 \\ Z - X + sI &\succeq 0. \end{aligned}$$

Proof:
Suppose $\sum_{i=1}^k \lambda_i(X) \leq t$. Then, choosing $s = \lambda_k$ and $Z = \mathrm{Diag}(\lambda_1 - s, \ldots, \lambda_k - s, 0, \ldots, 0)$, above LMIs hold.
Conversely, if above LMI holds, then, (since $Z \succeq 0$)

$$\begin{aligned} X \preceq Z + sI \qquad \implies \sum\nolimits_{i=1}^k \lambda_i(X) &\leq \sum\nolimits_{i=1}^k (\lambda_i(Z) + s) \\ &\leq \sum\nolimits_{i=1}^n \lambda_i(Z) + ks \\ &\leq t \qquad \text{(from first ineq.)}. \end{aligned}$$

SDP / LMI Representation

♠ **Nuclear norm:** $X \in \mathbb{R}^{m\times n}$; $\|X\|_{\mathsf{tr}} := \sum_{i=1}^{n} \sigma_i(X) \leq t$ iff

SDP / LMI Representation

♠ **Nuclear norm:** $X \in \mathbb{R}^{m\times n}$; $\|X\|_{\mathsf{tr}} := \sum_{i=1}^{n} \sigma_i(X) \le t$ iff

$$\begin{aligned} t - ns - \mathrm{Tr}(Z) &\ge 0 \\ Z &\succeq 0 \\ Z - \begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix} + sI_{m+n} &\succeq 0. \end{aligned}$$

SDP / LMI Representation

♠ **Nuclear norm:** $X \in \mathbb{R}^{m \times n}$; $\|X\|_{\text{tr}} := \sum_{i=1}^{n} \sigma_i(X) \leq t$ iff

$$\begin{aligned} t - ns - \operatorname{Tr}(Z) &\geq 0 \\ Z &\succeq 0 \\ Z - \begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix} + sI_{m+n} &\succeq 0. \end{aligned}$$

Follows from: $\lambda\left(\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}\right) = (\pm\sigma(X), 0, \ldots, 0)$.

SDP / LMI Representation

♠ **Nuclear norm:** $X \in \mathbb{R}^{m\times n}$; $\|X\|_{\mathrm{tr}} := \sum_{i=1}^{n} \sigma_i(X) \le t$ iff

$$\begin{aligned} t - ns - \mathrm{Tr}(Z) &\ge 0 \\ Z &\succeq 0 \\ Z - \begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix} + sI_{m+n} &\succeq 0. \end{aligned}$$

Follows from: $\lambda\left(\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}\right) = (\pm\sigma(X), 0, \ldots, 0)$.

Alternatively, we may SDP-represent nuclear norm as

$$\|X\|_{\mathrm{tr}} \le t \quad \Leftrightarrow \quad \exists U, V : \begin{bmatrix} U & X \\ X^T & V \end{bmatrix} \succeq 0, \quad \mathrm{Tr}(U+V) \le 2t.$$

Proof is slightly more involved (see lecture notes).

SDP example

Logarithmic Chebyshev approximation

$$\min \quad \max_{1 \le i \le m} |\log(a_i^T x) - \log b_i|$$

SDP example

Logarithmic Chebyshev approximation

$$\min \quad \max_{1 \le i \le m} |\log(a_i^T x) - \log b_i|$$

$$|\log(a_i^T x) - \log b_i| = \log \max(a_i^T x / b_i, b_i / a_i^T x)$$

SDP example

Logarithmic Chebyshev approximation

$$\min \quad \max_{1 \le i \le m} |\log(a_i^T x) - \log b_i|$$

$$|\log(a_i^T x) - \log b_i| = \log \max(a_i^T x / b_i, b_i / a_i^T x)$$

Reformulation

$$\min_{x,t} \quad t \qquad \text{s.t. } 1/t \le a_i^T x / b_i \le t, \quad i = 1, \ldots, m.$$

SDP example

Logarithmic Chebyshev approximation

$$\min \quad \max_{1 \le i \le m} |\log(a_i^T x) - \log b_i|$$

$$|\log(a_i^T x) - \log b_i| = \log \max(a_i^T x / b_i, b_i / a_i^T x)$$

Reformulation

$$\min_{x,t} \quad t \qquad \text{s.t. } 1/t \le a_i^T x / b_i \le t, \quad i = 1, \ldots, m.$$

$$\begin{bmatrix} a_i^T x / b_i & 1 \\ 1 & t \end{bmatrix} \succeq 0, \qquad i = 1, \ldots, m.$$

Least-squares SDP

$$\min \quad \|X - Y\|_2^2 \qquad \text{s.t. } X \succeq 0.$$

Exercise 1: Try solving using CVX (assume $Y^T = Y$); note $\|\cdot\|_2$ above is the operator 2-norm; not the Frobenius norm.

Exercise 2: Recast as SDP. *Hint:* Begin with $\min_{X,t} t \qquad \text{s.t. } \ldots$
Exercise 3: Solve the two questions also with $\|X - Y\|_{\mathsf{F}}^2$

Exercise 4: Verify against analytic solution: $X = U\Lambda^+ U^T$, where $Y = U\Lambda U^T$, and $\Lambda^+ = \text{Diag}(\max(0, \lambda_1), \ldots, \max(0, \lambda_n))$.

SDP relaxation

Binary Least-squares

$$\min \quad \|Ax - b\|^2$$
$$x_i \in \{-1, +1\} \quad i = 1, \ldots, n.$$

- Fundamental problem (engineering, computer science)
- Nonconvex; $x_i \in \{-1, +1\}$ – 2^n possible solutions
- Very hard in general (even to approximate)

SDP relaxation

Binary Least-squares

$$\min \quad \|Ax - b\|^2$$
$$x_i \in \{-1, +1\} \quad i = 1, \ldots, n.$$

- Fundamental problem (engineering, computer science)
- Nonconvex; $x_i \in \{-1, +1\}$ – 2^n possible solutions
- Very hard in general (even to approximate)

$$\min \qquad x^T A^T A x - 2x^T A^T b + b^T b \qquad x_i^2 = 1$$
$$\min \qquad \mathrm{Tr}(A^T A x x^T) - 2b^T A x \qquad x_i^2 = 1$$

SDP relaxation

Binary Least-squares

$$\min \quad \|Ax - b\|^2$$
$$x_i \in \{-1, +1\} \quad i = 1, \dots, n.$$

- Fundamental problem (engineering, computer science)
- Nonconvex; $x_i \in \{-1, +1\}$ – 2^n possible solutions
- Very hard in general (even to approximate)

$$\min \qquad x^T A^T A x - 2x^T A^T b + b^T b \qquad x_i^2 = 1$$
$$\min \qquad \mathrm{Tr}(A^T A x x^T) - 2b^T A x \qquad x_i^2 = 1$$
$$\min \qquad \mathrm{Tr}(A^T A Y) - 2b^T A x \qquad \text{s.t. } Y = xx^T,\ \mathrm{diag}(Y) = 1.$$

SDP relaxation

Binary Least-squares

$$\min \quad \|Ax - b\|^2$$
$$x_i \in \{-1, +1\} \quad i = 1, \ldots, n.$$

- Fundamental problem (engineering, computer science)
- Nonconvex; $x_i \in \{-1, +1\}$ – 2^n possible solutions
- Very hard in general (even to approximate)

$$\min \quad x^T A^T A x - 2x^T A^T b + b^T b \qquad x_i^2 = 1$$
$$\min \quad \mathrm{Tr}(A^T A x x^T) - 2b^T A x \qquad x_i^2 = 1$$
$$\min \quad \mathrm{Tr}(A^T A Y) - 2b^T A x \qquad \text{s.t. } Y = xx^T,\ \mathrm{diag}(Y) = 1.$$

- Still hard: $Y = xx^T$ is a nonconvex constraint.

SDP relaxation

Replace $Y = xx^T$ by $Y \succeq xx^T$. Thus, we obtain

$$\begin{aligned} \min \quad & \mathrm{Tr}(A^T A Y) - 2b^T A x \\ & Y \succeq xx^T, \quad \mathsf{diag}(Y) = 1. \end{aligned}$$

SDP relaxation

Replace $Y = xx^T$ by $Y \succeq xx^T$. Thus, we obtain

$$\begin{aligned} \min \quad & \mathrm{Tr}(A^T A Y) - 2b^T A x \\ & Y \succeq xx^T, \quad \mathrm{diag}(Y) = 1. \end{aligned}$$

This is an SDP, since

$$Y \succeq xx^T \quad \Leftrightarrow \quad \begin{bmatrix} Y & x \\ x^T & 1 \end{bmatrix} \succeq 0$$

(using Schur complements).

SDP relaxation

Replace $Y = xx^T$ by $Y \succeq xx^T$. Thus, we obtain

$$\begin{aligned} \min \quad & \mathrm{Tr}(A^T AY) - 2b^T Ax \\ & Y \succeq xx^T, \quad \mathsf{diag}(Y) = 1. \end{aligned}$$

This is an SDP, since

$$Y \succeq xx^T \quad \Leftrightarrow \quad \begin{bmatrix} Y & x \\ x^T & 1 \end{bmatrix} \succeq 0$$

(using Schur complements).

- Optimal value gives lower bound on binary LS
- Recover binary x by *randomized rounding*

Exercise: Try the above problem in CVX.

Nonconvex quadratic optimization

$$\begin{aligned} \min \quad & x^T A x + b^T x \\ & x^T P_i x + b_i^T x + c \leq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Nonconvex quadratic optimization

$$\begin{aligned} \min \quad & x^T A x + b^T x \\ & x^T P_i x + b_i^T x + c \leq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Exercise: Show that $x^T Q x = \mathrm{Tr}(Q x x^T)$ (where Q is symmetric).

Nonconvex quadratic optimization

$$\begin{aligned} \min \quad & x^T A x + b^T x \\ & x^T P_i x + b_i^T x + c \le 0, \quad i = 1, \ldots, m. \end{aligned}$$

Exercise: Show that $x^T Q x = \mathrm{Tr}(Q x x^T)$ (where Q is symmetric).

$$\begin{aligned} \min_{X,x} \quad & \mathrm{Tr}(AX) + b^T x \\ & \mathrm{Tr}(P_i X) + b_i^T x + c \le 0, \quad i = 1, \ldots, m \\ & X \succeq 0, \qquad \textbf{rank}(X) = 1. \end{aligned}$$

Nonconvex quadratic optimization

$$\begin{aligned} \min \quad & x^T A x + b^T x \\ & x^T P_i x + b_i^T x + c \leq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Exercise: Show that $x^T Q x = \mathrm{Tr}(Q x x^T)$ (where Q is symmetric).

$$\begin{aligned} \min_{X,x} \quad & \mathrm{Tr}(AX) + b^T x \\ & \mathrm{Tr}(P_i X) + b_i^T x + c \leq 0, \quad i = 1, \ldots, m \\ & X \succeq 0, \qquad \textbf{rank}(X) = 1. \end{aligned}$$

- Relax nonconvex $\textbf{rank}(X) = 1$ to $X \succeq x x^T$.
- Can be quite bad, but sometimes also quite tight.

References

1. L. Vandenberghe. MLSS 2012 Lecture slides; EE236B slides
2. A. Nemirovski. Lecture slides on modern convex optimization.