Lecture 10
Semidefinite programming

- semidefinite programming
- applications

Semidefinite programming (SDP)

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & F(x) = F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0 \\
& Ax = b
\end{array}
$$

where $F_i = F_i^T \in \mathbf{R}^{p \times p}$

- SDP is cvx opt problem in standard form with generalized (matrix) inequality
- LMI $F(x) \preceq 0$ is equivalent to a set of polynomial inequalities (nonnegative diagonal minors of $-F$)
- multiple LMIs can be combined into one (block diagonal) LMI
- many nonlinear cvx problems can be cast as SDPs

Examples

LP as SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

can be expressed as SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{diag}\,(Ax - b) \preceq 0 \end{array}$$

since $Ax - b \preceq 0 \iff \mathbf{diag}\,(Ax - b) \preceq 0$
(that's tricky notation!)

maximum eigenvalue minimization

$$\text{minimize}_x \ \lambda_{\max}(A(x))$$

$A(x) = A_0 + x_1 A_1 + \cdots + x_m A_m$, $A_i = A_i^T$
SDP with variables $x \in \mathbf{R}^m$ and $t \in \mathbf{R}$:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A(x) - tI \preceq 0 \end{array}$$

Schur complements

$$X = X^T = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

– $S = C - B^T A^{-1} B$ is the **Schur complement** of A in X (provided $\det A \neq 0$)

– useful to represent nonlinear convex constraints as LMIs

facts: (exercise)

– $X \succ 0$ if and only if $A \succ 0$ and $S \succ 0$

– if $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$

example. (convex) quadratic inequality

$$(Ax+b)^T(Ax+b) - c^T x - d \leq 0$$

is equivalent to the LMI

$$\begin{bmatrix} I & Ax+b \\ (Ax+b)^T & c^T x + d \end{bmatrix} \succeq 0$$

QCQP as SDP

the quadratically constrained quadratic program

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, L \end{array}$$

where $f_i(x) \triangleq (A_i x + b)^T (A_i x + b) - c_i^T x - d_i$

can be expressed as SDP (in x and t)

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} I & A_0 x + b_0 \\ (A_0 x + b_0)^T & c_0^T x + d_0 + t \end{array} \right] \succeq 0, \\ & \left[\begin{array}{cc} I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{array} \right] \succeq 0, \quad i = 1, \ldots, L \end{array}$$

extends to problems over second-order cone:

$$\|Ax + b\| \leq e^T x + d$$

is equivalent to LMI

$$\left[\begin{array}{cc} (e^T x + d) I & Ax + b \\ (Ax + b)^T & e^T x + d \end{array} \right] \succeq 0$$

Simple nonlinear example

$$\begin{array}{ll} \text{minimize} & \dfrac{(c^T x)^2}{d^T x} \\ \text{subject to} & Ax \preceq b \end{array}$$

(assume $d^T x > 0$ whenever $Ax \preceq b$)

1. equivalent problem with linear objective:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & Ax \preceq b \\ & t - \dfrac{(c^T x)^2}{d^T x} \geq 0 \end{array}$$

2. SDP (in x, t) using Schur complement:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} \mathbf{diag}\,(b - Ax) & 0 & 0 \\ 0 & t & c^T x \\ 0 & c^T x & d^T x \end{bmatrix} \succeq 0 \end{array}$$

Maximum eigenvalue optimization

$$\text{minimize } \lambda_{\max}(A(x))$$

where

$$A(x) = A_0 + x_1 A_1 + \cdots + x_m A_m, \quad A_i \in \mathbf{R}^{n \times n}, \quad A_i = A_i^T$$

and $\lambda_{\max}(A)$ is largest eigenvalue of (symmetric) matrix A

can cast as SDP:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & tI - A(x) \succeq 0 \end{array}$$

Matrix norm minimization

$$\text{minimize } \|A(x)\|$$

where

$$A(x) = A_0 + x_1 A_1 + \cdots + x_m A_m, \quad A_i \in \mathbf{R}^{p \times q}$$

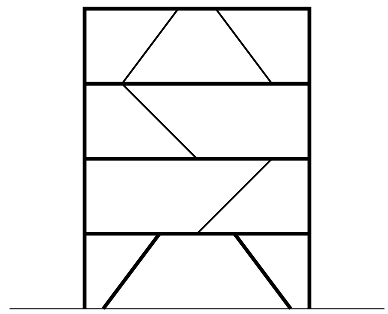
and $\|A\| = \sigma_1(A) = \left(\lambda_{\max}(A^T A)\right)^{1/2}$

can cast as SDP:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0 \end{array}$$

Optimizing structural dynamics

linear elastic structure



dynamics (ignoring damping): $M\ddot{d} + Kd = 0$

- $d(t) \in \mathbf{R}^k$: vector of displacements
- $M = M^T \succ 0$: mass matrix
- $K = K^T \succ 0$: stiffness matrix

solutions have form $d_i(t) = \Sigma_{j=1}^k \alpha_{ij} \cos(\omega_j t - \phi_j)$

- modal frequencies: $\omega \geq 0$ s.t. $\det(K - M\omega^2) = 0$
  $\omega_1 \leq \omega_2 \leq \cdots \leq \omega_k$
- fundamental frequency: $\omega_1 = \lambda_{\min}^{1/2}(M, K)$
  (structure behaves like mass below ω_1)

lower bound on fundamental frequency

$$\omega_1 \geq \Omega \iff M\Omega^2 - K \preceq 0$$

- design variables: x_i, cross-sectional area of structural member i (geometry of structure fixed)
- $M(x) = M_0 + \sum_i x_i M_i$
- $K(x) = K_0 + \sum_i x_i K_i$
- structure weight $w = w_0 + \sum_i x_i w_i$

problem:
minimize weight s.t. fundamental frequency $\geq \Omega$, limits on cross-sectional areas

as SDP:

$$\begin{array}{ll} \text{minimize} & w_0 + \sum_i x_i w_i \\ \text{subject to} & M(x)\Omega^2 - K(x) \preceq 0 \\ & l_i \leq x_i \leq u_i \end{array}$$

Measurements with unknown sensor noise variance

random vectors $y = x + v \in \mathbf{R}^k$

- x: random vector of interest,
  $\mathbf{E}x = \bar{x}$, $\mathbf{E}(x - \bar{x})(x - \bar{x})^T = \Sigma$
- v: measurement noise, independent of x,
  $\mathbf{E}v = 0$, $\mathbf{E}vv^T = F$, diagonal but otherwise unknown
- y: measured data, $\mathbf{E}y = \bar{x}$,
  $\mathbf{E}(y - \bar{x})(y - \bar{x})^T = \widehat{\Sigma} = \Sigma + F$

take **many** samples of $y \Rightarrow \bar{x}$, $\widehat{\Sigma}$ known

covariance Σ is unknown, but lies in (convex) set

$$\mathbf{S} = \{\widehat{\Sigma} - D \mid D \succeq 0 \text{ diagonal}, \widehat{\Sigma} - D \succeq 0\}$$

can bound linear function of Σ by solving SDP over $\mathbf{S}$

example. can bound variance of $c^T x$ by solving SDP:

upper bound:

$$\mathbf{E}(c^T x - c^T \bar{x})^2 = c^T \Sigma c \leq \sup_{\Sigma \in \mathbf{S}} c^T \Sigma c = c^T \widehat{\Sigma} c$$

lower bound:

$$\mathbf{E}(c^T x - c^T \bar{x})^2 = c^T \Sigma c \geq \inf_{\Sigma \in \mathbf{S}} c^T \Sigma c$$

i.e., solve SDP in D:

$$\begin{array}{ll} \text{minimize} & c^T \widehat{\Sigma} c - c^T D c \\ \text{subject to} & D \text{ diag.}, \ D \succeq 0 \\ & \widehat{\Sigma} - D \succeq 0 \end{array}$$

special case. 'educational testing problem' $(c = \mathbf{1})$

- x: 'ability' of a random student on k tests
- y: score of a random student on k tests
- v: testing error of k tests
- $\mathbf{1}^T x$: total ability on tests
- $\mathbf{1}^T y$: total test score
- $\mathbf{1}^T \Sigma \mathbf{1}$: variance in total ability
- $\mathbf{1}^T \widehat{\Sigma} \mathbf{1}$: variance in total score
- reliability of the test:

$$\frac{\mathbf{1}^T \Sigma \mathbf{1}}{\mathbf{1}^T \widehat{\Sigma} \mathbf{1}} = 1 - \frac{\mathbf{Tr}\, F}{\mathbf{1}^T \widehat{\Sigma} \mathbf{1}}$$

can lower bound reliability by solving SDP:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}\, D \\ \text{subject to} & D \text{ diagonal}, \;\; D \succeq 0 \\ & \widehat{\Sigma} - D \succeq 0 \end{array}$$

Covariance matrix reconstruction

Let W be the second-moment matrix of a random variable v

$$W = \mathbf{E}(vv^T)$$

assume we know W only partially: there is a subset of indices $\mathcal{I} \times \mathcal{J}$ such that

$$w_{ij}^{\mathrm{low}} \le W_{ij} \le w_{ij}^{\mathrm{up}}, \quad (i,j) \in \mathcal{I} \times \mathcal{J} \qquad (*)$$

where the numbers $w_{ij}^{\mathrm{low}}, w_{ij}^{\mathrm{up}}$ are given

reconstruction problem: find a matrix W that is consistent with observation, *i.e.*:

$$W \succ 0 \text{ and } (*)$$

This is an SDP (feasibility problem)

Reconstruction from moments

Given $m_0, \ldots, m_{2n}$, find if there exists a random variable such that m_i is the i-th moment of X for all i

Fact: the sequence $m_0, \ldots, m_{2n}$ is a sequence of moments iff

$$H(m_0, \ldots, m_{2n}) := \begin{bmatrix} m_0 & m_1 & \ldots & m_n \\ m_1 & m_2 & \ldots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \ldots & m_{2n} \end{bmatrix} \succeq 0$$

From there: can find maximum variance among all random variables subject to constraints on their moments:

$$\begin{gathered} \max \ y \text{ subject to } l_i \leq m_i \leq u_i, \quad i = 0, \ldots, 2n, \\ H(m_0, \ldots, m_{2n}) \succeq 0, \\ \begin{bmatrix} x_2 - y & x_1 \\ x_1 & 1 \end{bmatrix} \succeq 0 \end{gathered}$$