

EE5138 OPTIMIZATION FOR ELECTRICAL ENGINEERING/ EE6138 OPTIMIZATION FOR ELECTRICAL ENGINEERING (ADVANCED)

Lecture 3: Convex Optimization Problems

Outline

- optimization problem in standard form
- convex optimization problems
- quasiconvex optimization
- linear optimization
- quadratic optimization
- geometric programming
- generalized inequality constraints
- semidefinite programming

Required reading: textbook chapter 4 (4.1-4.6)

Optimization problem in standard form

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

- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \rightarrow \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \rightarrow \mathbf{R}$, $i = 1, \dots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \rightarrow \mathbf{R}$ are the equality constraint functions

optimal value:

$$p^\star = \inf\{f_0(x) \mid f_i(x) \leq 0, \quad i = 1, \dots, m, \quad h_i(x) = 0, \quad i = 1, \dots, p\}$$

- $p^\star = \infty$ if problem is infeasible (no x satisfies the constraints)
- $p^\star = -\infty$ if problem is unbounded below

Optimal and locally optimal points

x is **feasible** if $x \in \text{dom } f_0$ and it satisfies the constraints

a feasible x is **optimal** if $f_0(x) = p^*$; X_{opt} is the set of optimal points

x is **locally optimal** if there is an $R > 0$ such that x is optimal for

$$\begin{array}{ll} \text{minimize (over } z) & f_0(z) \\ \text{subject to} & f_i(z) \leq 0, \quad i = 1, \dots, m, \quad h_i(z) = 0, \quad i = 1, \dots, p \\ & \|z - x\|_2 \leq R \end{array}$$

examples (with $n = 1$, $m = p = 0$)

- $f_0(x) = 1/x$, $\text{dom } f_0 = \mathbf{R}_{++}$: $p^* = 0$, no optimal point
- $f_0(x) = -\log x$, $\text{dom } f_0 = \mathbf{R}_{++}$: $p^* = -\infty$
- $f_0(x) = x \log x$, $\text{dom } f_0 = \mathbf{R}_{++}$: $p^* = -1/e$, $x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$, $p^* = -\infty$, local optimum at $x = 1$

Implicit constraints

the standard form optimization problem has an **implicit constraint**

$$x \in \mathcal{D} = \bigcap_{i=0}^m \text{dom } f_i \cap \bigcap_{i=1}^p \text{dom } h_i,$$

- we call \mathcal{D} the **domain** of the problem
- the constraints $f_i(x) \leq 0$, $h_i(x) = 0$ are the explicit constraints
- a problem is **unconstrained** if it has no explicit constraints ($m = p = 0$)

example:

$$\text{minimize } f_0(x) = -\sum_{i=1}^k \log(b_i - a_i^T x)$$

is an unconstrained problem with implicit constraints $a_i^T x < b_i$

Feasibility problem

$$\begin{array}{ll}\text{find} & x \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p\end{array}$$

can be considered a special case of the general problem with $f_0(x) = 0$:

$$\begin{array}{ll}\text{minimize} & 0 \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p\end{array}$$

- $p^* = 0$ if constraints are feasible; any feasible x is optimal
- $p^* = \infty$ if constraints are infeasible

Convex optimization problem

standard form convex optimization problem

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & a_i^T x = b_i, \quad i = 1, \dots, p\end{array}$$

- f_0, f_1, \dots, f_m are convex; equality constraints are affine
- problem is *quasiconvex* if f_0 is quasiconvex (and f_1, \dots, f_m convex)

often written as

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

important property: feasible set of a convex optimization problem is convex

example

$$\begin{array}{ll}\text{minimize} & f_0(x) = x_1^2 + x_2^2 \\ \text{subject to} & f_1(x) = x_1/(1 + x_2^2) \leq 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0\end{array}$$

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (according to our definition): f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

$$\begin{array}{ll}\text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \leq 0 \\ & x_1 + x_2 = 0\end{array}$$

Local and global optima

any locally optimal point of a convex problem is (globally) optimal

proof: suppose x is locally optimal and y is optimal with $f_0(y) < f_0(x)$

x locally optimal means there is an $R > 0$ such that

$$z \text{ feasible, } \|z - x\|_2 \leq R \implies f_0(z) \geq f_0(x)$$

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and

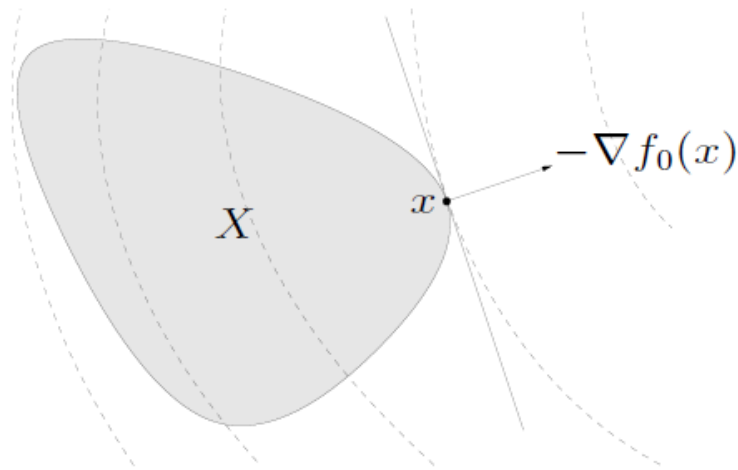
$$f_0(z) \leq \theta f_0(x) + (1 - \theta)f_0(y) < f_0(x)$$

which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable function

x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T (y - x) \geq 0 \quad \text{for all feasible } y$$



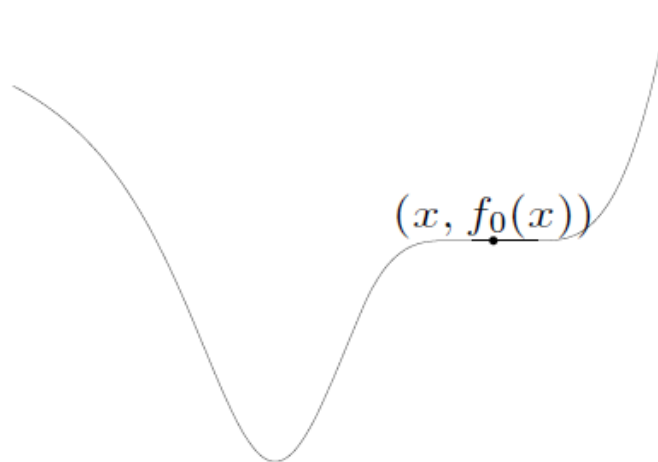
if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

Quasiconvex optimization

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

with $f_0 : \mathbf{R}^n \rightarrow \mathbf{R}$ quasiconvex, f_1, \dots, f_m convex

can have locally optimal points that are not (globally) optimal



convex representation of sublevel sets of f_0

if f_0 is quasiconvex, there exists a family of functions ϕ_t such that:

- $\phi_t(x)$ is convex in x for fixed t
- t -sublevel set of f_0 is 0-sublevel set of ϕ_t , i.e.,

$$f_0(x) \leq t \iff \phi_t(x) \leq 0$$

example

$$f_0(x) = \frac{p(x)}{q(x)}$$

with p convex, q concave, and $p(x) \geq 0$, $q(x) > 0$ on $\text{dom } f_0$

can take $\phi_t(x) = p(x) - tq(x)$:

- for $t \geq 0$, ϕ_t convex in x
- $p(x)/q(x) \leq t$ if and only if $\phi_t(x) \leq 0$

quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \leq 0, \quad f_i(x) \leq 0, \quad i = 1, \dots, m, \quad Ax = b \quad (1)$$

- for fixed t , a convex feasibility problem in x
- if feasible, we can conclude that $t \geq p^*$; if infeasible, $t \leq p^*$

Bisection method for quasiconvex optimization

given $l \leq p^*$, $u \geq p^*$, tolerance $\epsilon > 0$.

repeat

1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (1).
3. **if** (1) is feasible, $u := t$; **else** $l := t$.

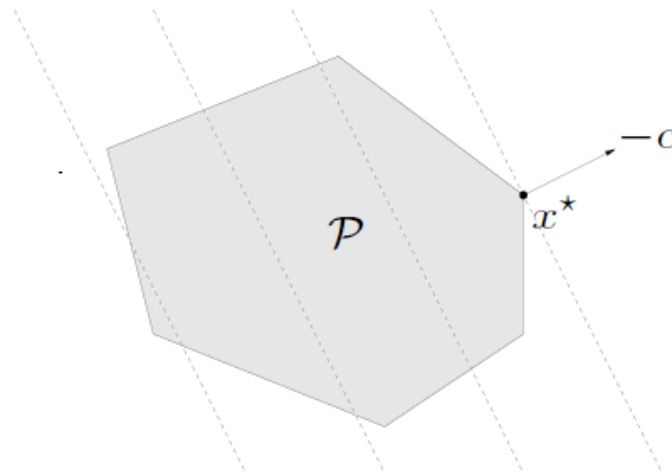
until $u - l \leq \epsilon$.

requires exactly $\lceil \log_2((u - l)/\epsilon) \rceil$ iterations (where u, l are initial values)

Linear program (LP)

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

- convex problem with affine objective and constraint functions
- feasible set is a polyhedron



Example

Chebyshev center of a polyhedron

Chebyshev center of

$$\mathcal{P} = \{x \mid a_i^T x \leq b_i, \quad i = 1, \dots, m\}$$

is center of largest inscribed ball

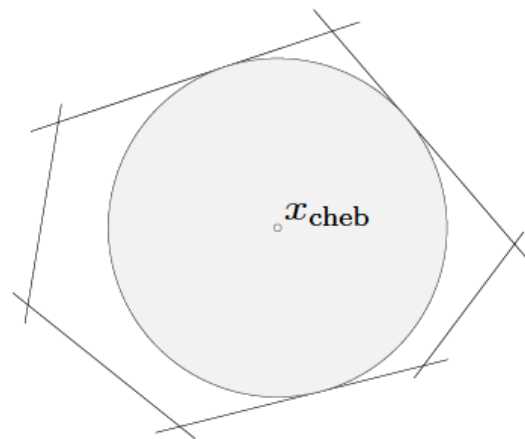
$$\mathcal{B} = \{x_c + u \mid \|u\|_2 \leq r\}$$

- $a_i^T x \leq b_i$ for all $x \in \mathcal{B}$ if and only if

$$\sup\{a_i^T (x_c + u) \mid \|u\|_2 \leq r\} = a_i^T x_c + r\|a_i\|_2 \leq b_i$$

- hence, x_c, r can be determined by solving the LP

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \leq b_i, \quad i = 1, \dots, m \end{array}$$



Linear-fractional program

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & Gx \preceq h \\ & Ax = b\end{array}$$

linear-fractional program

$$f_0(x) = \frac{c^T x + d}{e^T x + f}, \quad \text{dom } f_0(x) = \{x \mid e^T x + f > 0\}$$

- a quasiconvex optimization problem; can be solved by bisection

generalized linear-fractional program

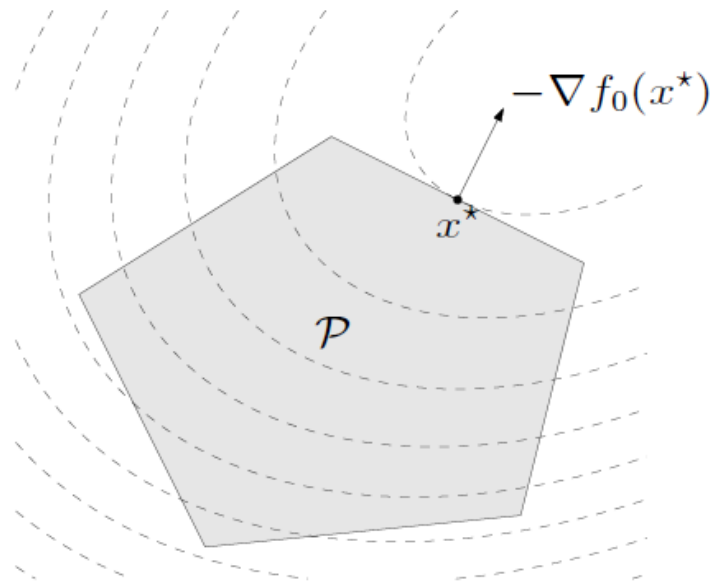
$$f_0(x) = \max_{i=1,\dots,r} \frac{c_i^T x + d_i}{e_i^T x + f_i}, \quad \text{dom } f_0(x) = \{x \mid e_i^T x + f_i > 0, i = 1, \dots, r\}$$

a quasiconvex optimization problem; can be solved by bisection

Quadratic program (QP)

$$\begin{array}{ll}\text{minimize} & (1/2)x^T Px + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b\end{array}$$

- $P \in \mathbf{S}_{+}^n$, so objective is convex quadratic
- minimize a convex quadratic function over a polyhedron



Examples

least-squares

$$\text{minimize} \quad \|Ax - b\|_2^2$$

- analytical solution $x^* = A^\dagger b$ (A^\dagger is pseudo-inverse) (textbook, A.5.4)
- can add linear constraints, *e.g.*, $l \preceq x \preceq u$

linear program with random cost

$$\begin{aligned} &\text{minimize} && \bar{c}^T x + \gamma x^T \Sigma x = \mathbf{E} c^T x + \gamma \mathbf{var}(c^T x) \\ &\text{subject to} && Gx \preceq h, \quad Ax = b \end{aligned}$$

- c is random vector with mean \bar{c} and covariance Σ
- hence, $c^T x$ is random variable with mean $\bar{c}^T x$ and variance $x^T \Sigma x$
- $\gamma > 0$ is risk aversion parameter; controls the trade-off between expected cost and variance (risk)

Quadratically constrained quadratic program (QCQP)

$$\begin{array}{ll}\text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m \\ & Ax = b\end{array}$$

- $P_i \in \mathbf{S}_{+}^n$; objective and constraints are convex quadratic
- if $P_1, \dots, P_m \in \mathbf{S}_{++}^n$, feasible region is intersection of m ellipsoids and an affine set

Second-order cone programming (SOCP)

$$\begin{array}{ll}\text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \dots, m \\ & Fx = g\end{array}$$

$$(A_i \in \mathbf{R}^{n_i \times n}, F \in \mathbf{R}^{p \times n})$$

- inequalities are called second-order cone (SOC) constraints:

$$(A_i x + b_i, c_i^T x + d_i) \in \text{second-order cone in } \mathbf{R}^{n_i+1}$$

- for $n_i = 0$, reduces to an LP; if $c_i = 0$, reduces to a QCQP
- more general than QCQP and LP

Robust linear programming (optional)

the parameters in optimization problems are often uncertain, *e.g.*, in an LP

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \dots, m,\end{array}$$

there can be uncertainty in c , a_i , b_i

two common approaches to handling uncertainty (in a_i , for simplicity)

- deterministic model: constraints must hold for all $a_i \in \mathcal{E}_i$

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{E}_i, \quad i = 1, \dots, m,\end{array}$$

- stochastic model: a_i is random variable; constraints must hold with probability η

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \dots, m\end{array}$$

deterministic approach via SOCP

- choose an ellipsoid as \mathcal{E}_i :

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\} \quad (\bar{a}_i \in \mathbf{R}^n, \quad P_i \in \mathbf{R}^{n \times n})$$

center is \bar{a}_i , semi-axes determined by singular values/vectors of P_i

- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \dots, m \end{array}$$

is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \|P_i^T x\|_2 \leq b_i, \quad i = 1, \dots, m \end{array}$$

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

stochastic approach via SOCP

- assume a_i is Gaussian with mean \bar{a}_i , covariance Σ_i ($a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$)
- $a_i^T x$ is Gaussian r.v. with mean $\bar{a}_i^T x$, variance $x^T \Sigma_i x$; hence

$$\text{prob}(a_i^T x \leq b_i) = \Phi \left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2} \right)$$

where $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^x e^{-t^2/2} dt$ is CDF of $\mathcal{N}(0, 1)$

- robust LP

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && \text{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \dots, m, \end{aligned}$$

with $\eta \geq 1/2$, is equivalent to the SOCP

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i, \quad i = 1, \dots, m \end{aligned}$$

Geometric programming (GP)

monomial function

$$f(x) = cx_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}, \quad \text{dom } f = \mathbf{R}_{++}^n$$

with $c > 0$; exponent a_i can be any real number

posynomial function: sum of monomials

$$f(x) = \sum_{k=1}^K c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \quad \text{dom } f = \mathbf{R}_{++}^n$$

geometric program (GP)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 1, \quad i = 1, \dots, m \\ & h_i(x) = 1, \quad i = 1, \dots, p \end{array}$$

with f_i posynomial, h_i monomial

Geometric program in convex form

change variables to $y_i = \log x_i$, and take logarithm of cost, constraints

- monomial $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ transforms to

$$\log f(e^{y_1}, \dots, e^{y_n}) = a^T y + b \quad (b = \log c)$$

- posynomial $f(x) = \sum_{k=1}^K c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

$$\log f(e^{y_1}, \dots, e^{y_n}) = \log \left(\sum_{k=1}^K e^{a_k^T y + b_k} \right) \quad (b_k = \log c_k)$$

- geometric program transforms to convex problem

$$\begin{aligned} & \text{minimize} && \log \left(\sum_{k=1}^K \exp(a_{0k}^T y + b_{0k}) \right) \\ & \text{subject to} && \log \left(\sum_{k=1}^K \exp(a_{ik}^T y + b_{ik}) \right) \leq 0, \quad i = 1, \dots, m \\ & && Gy + d = 0 \end{aligned}$$

Generalized inequality constraints (optional)

convex problem with generalized inequality constraints

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & Ax = b\end{array}$$

- $f_0 : \mathbf{R}^n \rightarrow \mathbf{R}$ convex; $f_i : \mathbf{R}^n \rightarrow \mathbf{R}^{k_i}$ K_i -convex w.r.t. proper cone K_i
- same properties as standard convex problem (convex feasible set, local optimum is global, etc.)

conic form problem: special case with affine objective and constraints

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & Fx + g \preceq_K 0 \\ & Ax = b\end{array}$$

extends linear programming ($K = \mathbf{R}_+^m$) to nonpolyhedral cones

Semidefinite programming (SDP)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b \end{array}$$

with $F_i, G \in \mathbf{S}^k$

- inequality constraint is called linear matrix inequality (LMI)
- includes problems with multiple LMI constraints: for example,

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \quad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0$$

is equivalent to single LMI

$$x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \tilde{G} \end{bmatrix} \preceq 0$$

Recast LP and SOCP as SDP

LP and equivalent SDP

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

(note different interpretation of generalized inequality \preceq)

SOCP and equivalent SDP

$$\begin{array}{ll} \text{SOCP:} & \begin{array}{l} \text{minimize } f^T x \\ \text{subject to } \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \dots, m \end{array} \end{array}$$

$$\begin{array}{ll} \text{SDP:} & \begin{array}{l} \text{minimize } f^T x \\ \text{subject to } \begin{bmatrix} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0, \quad i = 1, \dots, m \end{array} \end{array}$$

(Applying Schur complement, see textbook A5.5)

Eigenvalue minimization

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

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbf{S}^k$)

equivalent SDP

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

- variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$
- follows from

$$\lambda_{\max}(A) \leq t \iff A \preceq tI$$