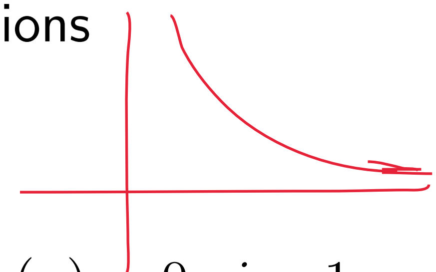
4. Convex optimization problems

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

Optimization problem in standard form

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

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



optimal value: infimum ← minimum

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

- $p^\star = \infty$ 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 $x \in \mathbf{dom}\, f_0$ and it satisfies the constraints

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

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

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

local

optimal

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

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

$f_0'(x) = 3x^2 - 3 = 0$

Implicit constraints

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

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

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

$-\log x$: convex.

example:

convex

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

$b_i - a_i^T x > 0$,

$(i = 1, 2, \ldots, k)$

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

convex set

Feasibility problem

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

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

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

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

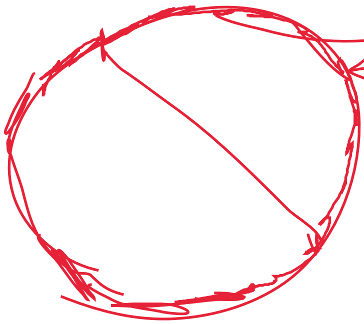
$x \in \mathrm{dom} f$

Convex optimization problem

$C_\alpha = \{x \mid f(x) \leq \alpha\}$

(α-sublevel)

standard form convex optimization problem



convex set

$\|x\| = 0$

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

$h_i(x) = 0$

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

often written as

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

x

$A \in \mathbf{R}^{p \times n}$

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

example

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\quad f_0(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\|x\|_2^2 = x^T P x$

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

- 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

  $\theta \in [0,1]$

  proof: suppose x is locally optimal, but there exists a feasible y with $f_0(y) < f_0(x)$

  $z = \theta y + (1-\theta)x$ y global

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

$$z \text{ feasible}, \quad \|z - x\|_2 \le R \quad \Longrightarrow \quad f_0(z) \ge f_0(x)$$

$\boxed{f(y) \le f(x)}$

local

  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

$f(z) \le \theta f(y) + (1-\theta) f(x)$

$= \theta \underbrace{(f(y) - f(x))}_{\le 0} + f(x)$

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

$\Longrightarrow f(z) \le f(x)$

  which contradicts our assumption that x is locally optimal

  contradict

$\nabla f_0(x)^T (y-x) \geq 0, \ \forall y$

$\nabla f_0(x)^T z \geq 0, \ \forall z$

$\Downarrow$

$\nabla f_0(x) = 0$

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

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

$\min_x f_0(x), \ \forall x$

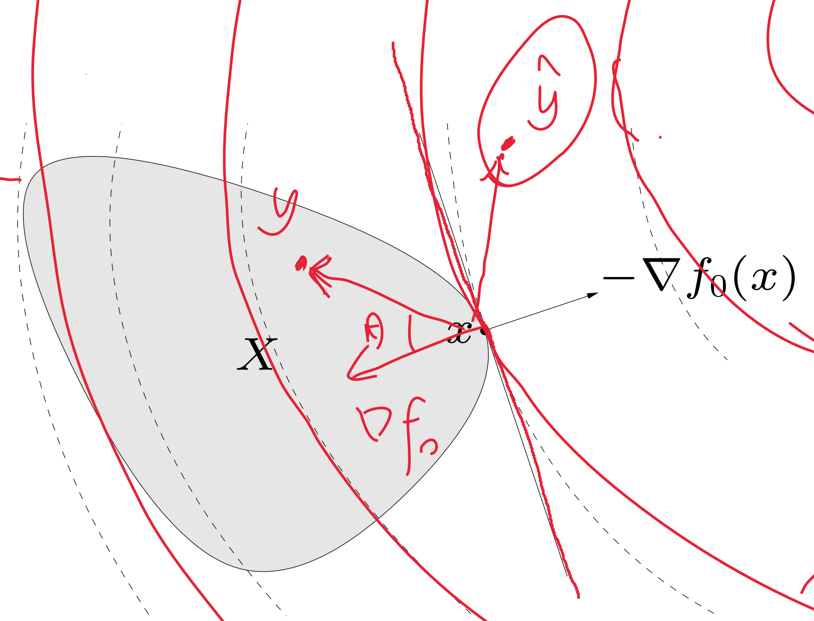
s.t. $f_i(x) \leq 0, \ \forall i$

$Ax = b$

$f(x)$: convex. i.i.f.

$f(y) \geq f(x) + \underbrace{\nabla f(x)^T (y-x)}_{\geq 0}$

$\Rightarrow f(y) \geq f(x), \ \forall y$



$f_0(x) = \text{const}$

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

- **unconstrained problem**: x is optimal if and only if

$$x \in \mathbf{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

- **equality constrained problem**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad Ax = b$$

x is optimal if and only if there exists a ν such that

$$x \in \mathbf{dom}\, f_0, \qquad Ax = b, \qquad \nabla f_0(x) + A^T \nu = 0$$

- **minimization over nonnegative orthant**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad x \succeq 0$$

x is optimal if and only if

$\nabla f(x)^T x = 0$

$$x \in \mathbf{dom}\, f_0, \qquad x \succeq 0, \qquad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

Equivalent convex problems

two problems are (informally) **equivalent** if the solution of one is readily obtained from the solution of the other, and vice-versa

some common transformations that preserve convexity:

- **eliminating equality constraints**

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

$f(x)$: convex

$f(Ax+b)$: convex

is equivalent to

$$\begin{array}{ll} \text{minimize (over } z) & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{array}$$

where F and x_0 are such that $(R(F) = N(A), \ Ax_0 = b)$

$$Ax = b \quad \Longleftrightarrow \quad x = Fz + x_0 \text{ for some } z$$

$A(Fz + x_0) = \underline{AF} \cdot z + Ax_0 = 0 + b = b$

- **introducing equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(A_0x + b_0) \\ \text{subject to} & f_i(A_ix + b_i) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x, \ y_i) & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, \quad i = 1, \ldots, m \\ & y_i = A_ix + b_i, \quad i = 0, 1, \ldots, m \end{array}$$

- **introducing slack variables for linear inequalities**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^Tx \leq b_i, \quad i = 1, \ldots, m \end{array}$$

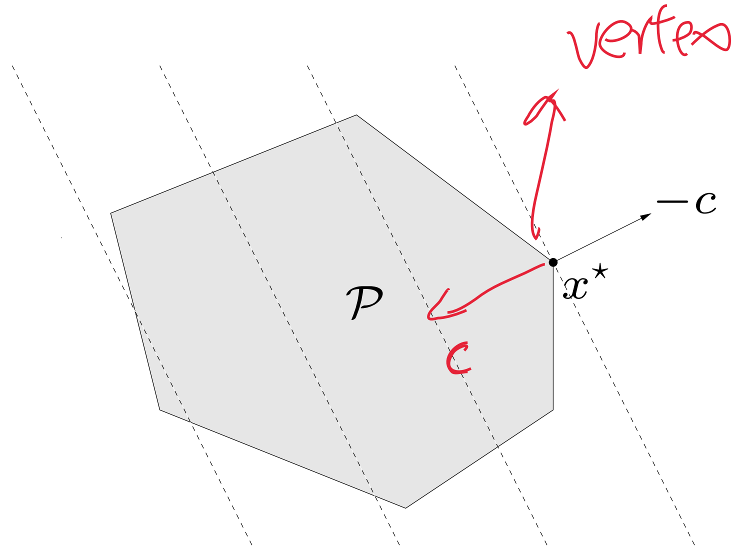
is equivalent to

$$\begin{array}{ll} \text{minimize (over } x, \ s) & f_0(x) \\ \text{subject to} & a_i^Tx + s_i = b_i, \quad i = 1, \ldots, m \\ & s_i \geq 0, \quad i = 1, \ldots m \end{array}$$

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



$c^T x + d = \text{const}_1$

$c^T x + d = \text{const}_2$ $(\text{const}_2 > \text{const}_1)$

Examples

diet problem: choose quantities $x_1, \ldots, x_n$ of n foods

- one unit of food j costs c_j, contains amount a_{ij} of nutrient i
- healthy diet requires nutrient i in quantity at least b_i

$a_i^T x \geq b_i$

$(i = 1, 2, \ldots, m)$

$\downarrow$

$Ax \geq b$

$(\mathbb{R}^n_+)$

to find cheapest healthy diet,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \succeq b, \quad x \succeq 0 \end{array}$$

piecewise-linear minimization

$$\text{minimize} \quad \max_{i=1,\ldots,m}(a_i^T x + b_i)$$

equivalent to an LP

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

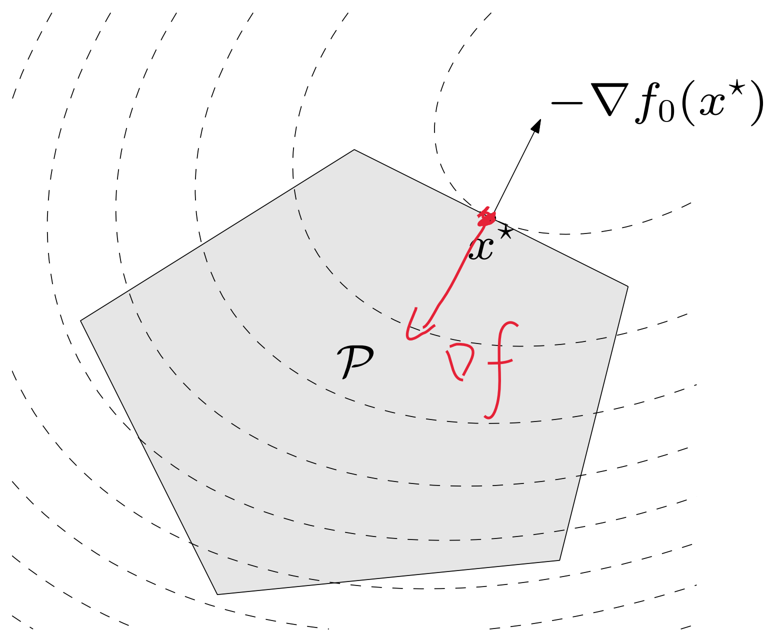
Quadratic program (QP)

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

$\nabla^2 f_0 = \underline{P \succeq 0}$

$(K = S^n_+)$

- $P \in \mathbf{S}^n_+$, 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 $x^\star = A^\dagger b$ ($A^\dagger$ is pseudo-inverse)
- can add linear constraints, *e.g.*, $l \preceq x \preceq u$

$$A = U\Sigma V^T$$

$$U^TU = UU^T = I$$

$$V^TV = VV^T = I$$

$$x^TA^TAx - 2x^TA^Tb + b^Tb$$

$$Ax = b$$

$$U\Sigma V^T x = b$$

$$\Sigma V^T x = U^T b$$

$$V^T x = \Sigma^{-1} U^T b$$

$$x = V\Sigma^{-1}U^T b$$

$$A^\dagger$$

$$A^\dagger = \lim_{\varepsilon \to 0} (A^TA + \varepsilon I)^{-1} A^T$$

$$= \lim_{\varepsilon \to 0} A^T (AA^T + \varepsilon I)^{-1}$$

$$\boxed{\begin{array}{ll} \min & \|Ax - b\|_2^2 \\ \text{s.t.} & \|x\|_2^2 \le t \end{array}}$$

Quadratically constrained quadratic program (QCQP)

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

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

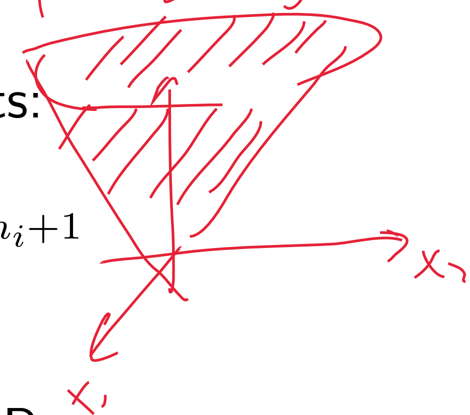
$$
\begin{tikzcd}
\mathrm{QCQP} \arrow[r, "P_i=0"] \arrow[rr, bend right, "P_i=P_0=0"'] & \mathrm{QP} \arrow[r, "P_0=0"] & \mathrm{LP}
\end{tikzcd}
$$

Second-order cone programming

$$
\begin{array}{ll}
\text{minimize} & f^T x \\
\text{subject to} & \|A_i x + b_i\|_2^2 \le (c_i^T x + d_i)^2, \quad i = 1, \ldots, m \\
& F x = g
\end{array}
$$

second-order cone

$\{(x,t) \mid \|x\|_2 \le t\}$



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

- 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

Geometric programming

monomial function

$$f(x) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$

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

posynomial function: sum of monomials

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

geometric program (GP)

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

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

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

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

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

- geometric program transforms to convex problem

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

Generalized inequality constraints

convex problem with generalized inequality constraints

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

- $f_0 : \mathbf{R}^n \to \mathbf{R}$ convex; $f_i : \mathbf{R}^n \to \mathbf{R}^{k_i}$ 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 program (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,

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

is equivalent to single LMI

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

(LMI.)

LP and SOCP as SDP

$\mathbf{diag}(Ax) \preceq \mathbf{diag}(b)$

LP and equivalent SDP

$Ax = [\;] \le [\;] = b$

LP:	minimize	$c^T x$	SDP:	minimize	$c^T x$
	subject to	$Ax \preceq b$		subject to	$\mathbf{diag}(Ax - b) \preceq 0$

(LMI)

(note different interpretation of generalized inequality $\preceq$)

$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \succeq 0$

$\Longleftrightarrow A \succeq 0,\quad D - C \cdot A^{-1} B \succeq 0$

complement

SOCP and equivalent SDP

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

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

$\Downarrow \; (c_i^T x + d_i) I \succeq 0, \quad (c_i^T x + d_i) - (A_i x + b_i)^T \frac{1}{c_i^T x + d_i} (A_i x + b_i) \succeq 0$

Eigenvalue minimization

SDP

LP, QP, QCQP

SOCP

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

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

equivalent SDP

LMI

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

SDP

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

$$\lambda_{\max}(A) \le t \quad \Longleftrightarrow \quad A \preceq tI$$

Matrix norm minimization

$$\text{minimize} \quad \|A(x)\|_2 = \left(\lambda_{\max}(A(x)^T A(x))\right)^{1/2}$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbf{R}^{p\times q}$)

equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} tI & A(x) \\ A(x)^T & tI \end{array}\right] \succeq 0 \end{array}$$

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

$$\begin{aligned} \|A\|_2 \leq t \quad &\Longleftrightarrow \quad A^T A \preceq t^2 I, \quad t \geq 0 \\ &\Longleftrightarrow \quad \left[\begin{array}{cc} tI & A \\ A^T & tI \end{array}\right] \succeq 0 \end{aligned}$$