Convex Optimization Forms

Lecturer: Pradeep Ravikumar
Co-instructor: Aarti Singh

Convex Optimization 10-725/36-725

Based on slides from Boyd, Vandenberghe, Tibshirani

Recall: standard form convex optimization problem

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

- f_0, f_1, . . . , f_m are convex; equality constraints are affine

often written as

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

Recall: equivalent transformations

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

Recall: equivalent transformations

some common transformations that preserve convexity:

- **eliminating equality constraints**

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

  is equivalent to

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

  where F and x_0 are such that

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

Recall: equivalent transformations

- **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

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

Recall: equivalent transformations

- **epigraph form**: standard form convex problem is equivalent to

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

Recall: equivalent transformations

- **minimizing over some variables**

$$
\begin{array}{ll}
\text{minimize} & f_0(x_1, x_2) \\
\text{subject to} & f_i(x_1) \le 0, \quad i = 1, \dots, m
\end{array}
$$

is equivalent to

$$
\begin{array}{ll}
\text{minimize} & \tilde{f}_0(x_1) \\
\text{subject to} & f_i(x_1) \le 0, \quad i = 1, \dots, m
\end{array}
$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

Recall: equivalent transformations

- **introducing slack variables for linear inequalities**

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

is equivalent to

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

Outline

- Canonical Optimization Problems
- Linear Programs (LP)
- Quadratic Programs (QP)
- Second Order Cone Programs (SOCP)
- Semi-definite Programs (SDP)
- Cone Programs

Outline

- Canonical Optimization Problems
- Linear Programs (LP)

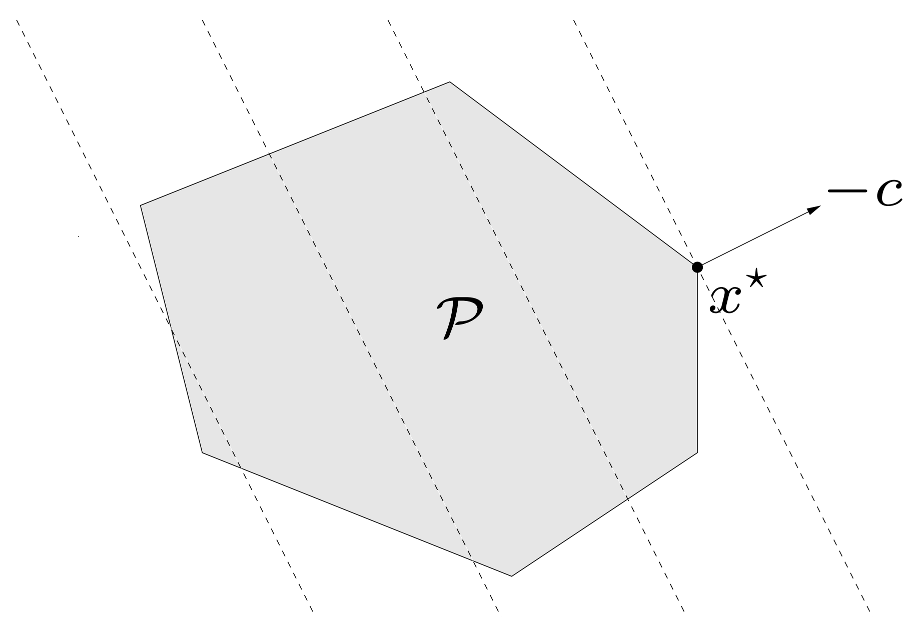
$$\text{LPs} \subseteq \text{QPs} \subseteq \text{SOCPs} \subseteq \text{SDPs} \subseteq \text{Conic programs}$$

- Semi-definite Programs (SDP)
- Cone Programs

Linear Program

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

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



Linear Program

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

- First introduced by **Kantorovich** in the late 1930s and **Dantzig** in the 1940s
- Dantzig's simplex algorithm gives a direct (non-iterative) solver for LPs (later in the course we'll see interior point methods)
- Fundamental problem in convex optimization. Many diverse applications, rich history

Example: Diets

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

Example: Diets

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

to find cheapest healthy diet,

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

Example: Transportation

Ship commodities from given sources to destinations at minimum cost

- s_i : supply at source i
- d_j : demand at destination j
- c_{ij} : per-unit shipping cost from i to j
- x_{ij} : units shipped from i to j

Example: Transportation

Ship commodities from given sources to destinations at minimum cost

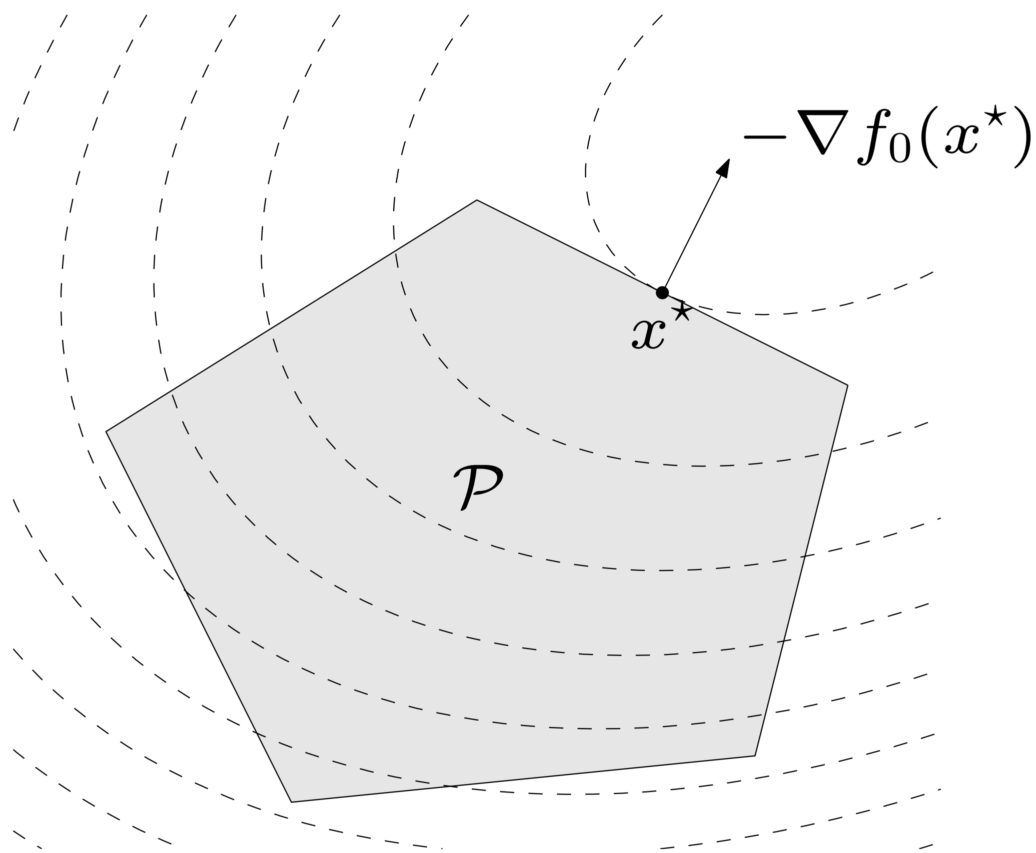
- s_i : supply at source i
- d_j : demand at destination j
- c_{ij} : per-unit shipping cost from i to j
- x_{ij} : units shipped from i to j

$$
\begin{array}{ll}
\displaystyle\min_x & \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\
\text{subject to} & \displaystyle\sum_{j=1}^{n} x_{ij} \le s_i,\ i = 1,\ldots,m \\
& \displaystyle\sum_{i=1}^{m} x_{ij} \ge d_j,\ j = 1,\ldots,n,\ \ x \ge 0
\end{array}
$$

Quadratic Programs

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

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



Example: 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, $e.g.$, $l \preceq x \preceq u$

Example: Linear Program with random cost

Consider:

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

Suppose:

- 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 $x^T \Sigma x$

Example: Linear Program with random cost

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

- 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)

Example: Support Vector Machines

Given $y \in \{-1,1\}^n$, $X \in \mathbb{R}^{n \times p}$ having rows $x_1, \ldots x_n$, recall the support vector machine or SVM problem:

$$
\begin{aligned}
\min_{\beta, \beta_0, \xi} \quad & \frac{1}{2}\|\beta\|_2^2 + C\sum_{i=1}^n \xi_i \\
\text{subject to} \quad & \xi_i \geq 0, \ i = 1, \ldots n \\
& y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i, \ i = 1, \ldots n
\end{aligned}
$$

This is a quadratic program

Second Order Cone Programming (SOCP)

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

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

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

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

Recall the definition of a second-order cone: $\{(x,t) : \|x\|_2 \le t\}$

Second Order Cone Programming (SOCP)

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

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

- 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

Example: Robust Linear Program

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

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

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 $a_i \in \mathcal{E}_i$

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

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

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

Example: Robust Linear Program — Deterministic

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

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

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

- robust LP

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

Example: Robust Linear Program — Deterministic

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

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

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

- robust LP

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

is equivalent to the SOCP

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

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

Example: Robust Linear Program — Stochastic

- 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

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

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

Example: Robust Linear Program — Stochastic

- 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

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

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

- robust LP

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

Example: Robust Linear Program — Stochastic

- 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

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

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

- robust LP

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

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

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

Generalized Inequalities

a convex cone $K \subseteq \mathbf{R}^n$ is a **proper cone** if

- K is closed (contains its boundary)
- K is solid (has nonempty interior)
- K is pointed (contains no line)

examples

- nonnegative orthant $K = \mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid x_i \geq 0,\ i = 1, \ldots, n\}$
- positive semidefinite cone $K = \mathbf{S}^n_+$

Generalized Inequalities

generalized inequality defined by a proper cone K:

$$x \preceq_K y \quad \Longleftrightarrow \quad y - x \in K, \qquad\qquad x \prec_K y \quad \Longleftrightarrow \quad y - x \in \mathbf{int}\, K$$

examples

- componentwise inequality ($K = \mathbf{R}^n_+$)

$$x \preceq_{\mathbf{R}^n_+} y \quad \Longleftrightarrow \quad x_i \leq y_i, \quad i = 1, \ldots, n$$

- matrix inequality ($K = \mathbf{S}^n_+$)

$$X \preceq_{\mathbf{S}^n_+} Y \quad \Longleftrightarrow \quad Y - X \text{ positive semidefinite}$$

these two types are so common that we drop the subscript in $\preceq_K$

properties: many properties of $\preceq_K$ are similar to $\leq$ on $\mathbf{R}$, *e.g.*,

$$x \preceq_K y, \quad u \preceq_K v \quad \Longrightarrow \quad x + u \preceq_K y + v$$

Generalized Inequality Programs

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 Programs

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 Programs

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

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

- inequality constraint is called linear matrix inequality (LMI)

Semidefinite Programs

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

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

- inequality constraint is called linear matrix inequality (LMI)

Both LPs and SDPs are special cases of conic programming. For LPs, $K = \mathbb{R}^n_+$; for SDPs, $K = \mathbb{S}^n_+$

Semidefinite Programs

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

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

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

Example: Eigenvalue Minimization

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

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

equivalent SDP

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

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

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

Example: 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 $A_i \in \mathbf{R}^{p \times q}$)

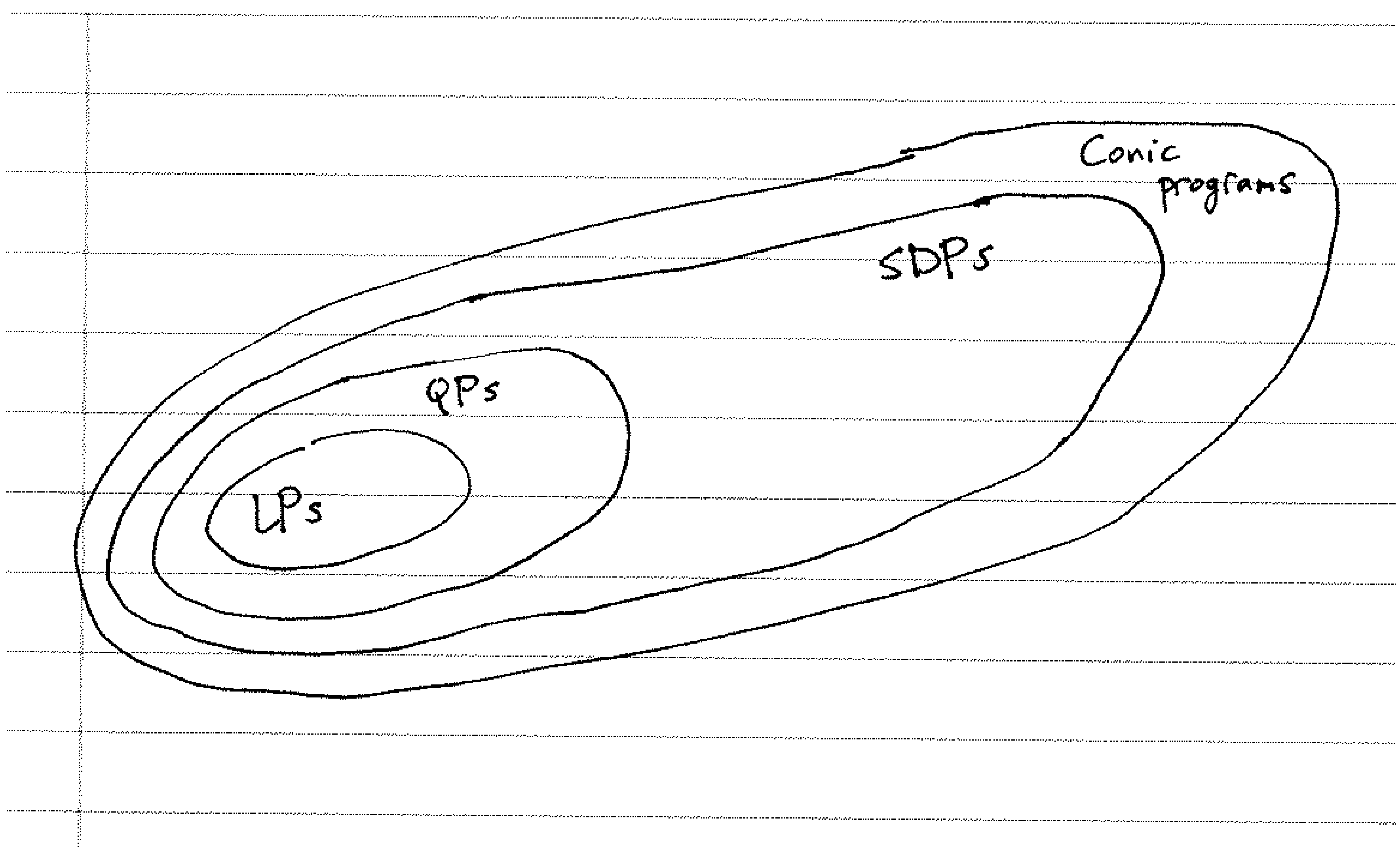
equivalent SDP

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

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

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

Hierarchy



LPs, QPs

LP

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to} \quad & Dx \leq d \\ & Ax = b \end{aligned}$$

QP

$$\begin{aligned} \min_x \quad & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} \quad & Dx \leq d \\ & Ax = b \end{aligned}$$

LPs $\subseteq$ QPs

QPs, SOCPs

Note that using tricks from equivalent transformations, we can rewrite QPs as:

$$\begin{aligned}
\min_{x,t} \quad & c^T x + t \\
\text{subject to} \quad & Dx \le d, \ \frac{1}{2} x^T Q x \le t \\
& Ax = b
\end{aligned}$$

Now write $\frac{1}{2} x^T Q x \le t \iff \|(\frac{1}{\sqrt{2}} Q^{1/2} x, \frac{1}{2}(1-t))\|_2 \le \frac{1}{2}(1+t)$

QPs $\subseteq$ SOCPs

SOCPs, SDPs

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

Hence we can write any SOCP constraint as an SDP constraint

SOCPs $\subseteq$ SDPs

SDPs, Conic Programs

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

- LPs, SOCPs, QPs, SDPs all naturally can be written as Conic Programs for appropriate cones, as noted earlier

Hierarchy

Take a breath (phew!). Thus we have established the hierachy

$$\text{LPs} \subseteq \text{QPs} \subseteq \text{SOCPs} \subseteq \text{SDPs} \subseteq \text{Conic programs}$$