FOUNDATIONS OF OPTIMIZATION: IE6001

Convex Optimization Problems

Napat Rujeerapaiboon

Semester I, AY2022/2023

Optimization Problems in Standard Form

$$
\begin{array}{lll}
\mathsf{P}: & \text{minimize} & f_0(x) \\[1em]
& \text{subject to} & f_i(x) \le 0 \quad \forall i = 1, \ldots, m \\
& & h_i(x) = 0 \quad \forall i = 1, \ldots, p
\end{array}
$$

$x \in \mathbb{R}^n$	optimization or decision variable
$f_0 : \mathbb{R}^n \to (-\infty, +\infty]$	objective or cost function
$f_i : \mathbb{R}^n \to (-\infty, +\infty]$	inequality constraint functions
$h_i : \mathbb{R}^n \to \mathbb{R}$	equality constraint functions

$$
\begin{array}{lll}
\inf \mathsf{P} = +\infty & \Longleftrightarrow & \mathsf{P} \text{ is infeasible} \\
\inf \mathsf{P} = -\infty & \Longleftrightarrow & \mathsf{P} \text{ is unbounded}
\end{array}
$$

Transforming Problems to the Standard Form

- "$\geq$"-inequalites:

$$\hat{f}(x) \geq 0 \iff f(x) = -\hat{f}(x) \leq 0$$

- nonzero right hand sides:

$$\hat{f}(x) \leq b \iff f(x) = \hat{f}(x) - b \leq 0$$

- maximization problems:

$$\sup_{x \in \mathcal{X}} \hat{f}_0(x) = -\inf_{x \in \mathcal{X}} f_0(x) \quad \text{where} \quad f_0(x) = -\hat{f}_0(x)$$

Note: In principle, all equality constraints can be transformed to inequality constraints:

$$f(x) = 0 \iff f(x) \leq 0 \text{ and } -f(x) \leq 0$$

Convex Optimization Problems

$$
\begin{array}{lll}
\mathrm{P}: & \text{minimize} & f_0(x) \\
& \text{subject to} & f_i(x) \leq 0 \quad \forall i = 1, \ldots, m \\
& & a_i^\top x = b_i \quad \forall i = 1, \ldots, p
\end{array}
$$

The objective and inequality constraint functions are *convex*, while all equality constraint functions are *linear* (affine). The equality constraints can be written compactly as $Ax = b$.

Proposition: The *feasible set* of a convex optimization problem is *convex*.

Proof: The feasible set of P is given by

$$\mathrm{dom}(f_0) \cap_{i=1}^{m} \{x : f_i(x) \leq 0\} \cap_{i=1}^{p} \{x : a_i^\top x = b\}.$$

It is convex as the domains and sublevel sets of convex functions are convex, while hyperplanes are also convex. Intersections of convex sets are convex.

Local Optima are Global Optima

Theorem: Every locally optimal solution of a convex optimization problem is also globally optimal.

Proof: Suppose that x is *not* a global minimizer, and let y be a global minimizer. By convexity, the decision $\theta x + (1-\theta)y$ is feasible for every $\theta \in [0,1]$, and its objective value satisfies

$$\begin{aligned} f_0(\theta x + (1-\theta)y) &\leq \theta f_0(x) + (1-\theta) f_0(y) \\ &< \theta f_0(x) + (1-\theta) f_0(x) \\ &= f_0(x) \end{aligned}$$

for all $\theta \in (0,1)$. Therefore, x cannot be a local minimizer. By contraposition, if x is a local minimizer, then it must be a global minimizer.

Optimality Criterion (Differentiable Objective)

Theorem: For P convex, $x^\star$ is optimal iff it is feasible and

$$\nabla f_0(x^\star)^\top (x - x^\star) \geq 0 \quad \text{for all feasible } x. \qquad (*)$$

Proof: Assume that $(*)$ holds for some feasible $x^\star$ and let x be any feasible point. The 1st-order conditions for f_0 imply

$$f_0(x) \geq f_0(x^\star) + \nabla f_0(x^\star)^\top (x - x^\star) \quad \forall x \in \mathbb{R}^n.$$

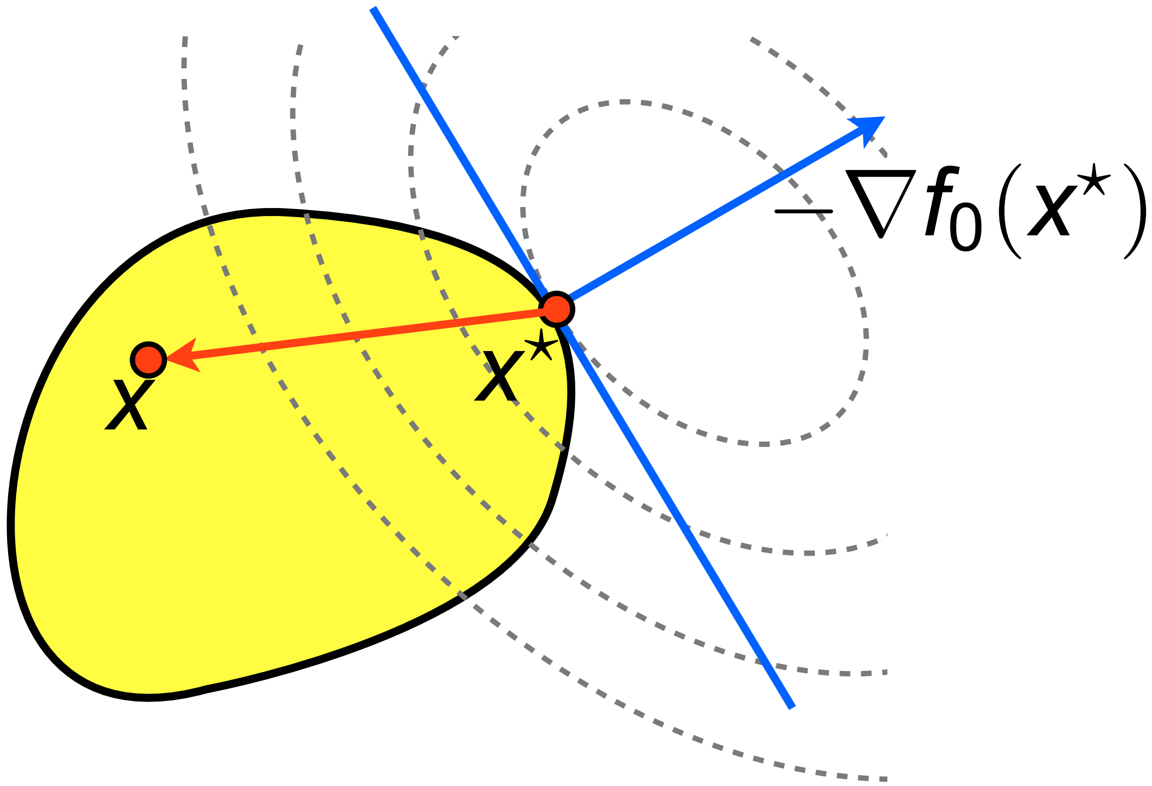
Using $(*)$, we conclude that $f_0(x) \geq f_0(x^\star)$. Thus, $x^\star$ is optimal. Conversely, assume that $x^\star$ is optimal and let x be any feasible point. By convexity, the line segment $[x^\star, x]$ is contained in the feasible set. By optimality of $x^\star$, the function

$$\psi(t) = f_0((1-t)x^\star + tx)$$

must satisfy $\psi(t) \geq \psi(0)\ \forall t \in [0, 1]$. This is only possible if

$$\psi'(0) = \nabla f_0(x^\star)^\top (x - x^\star) \geq 0.$$

Optimality Criterion (Differentiable Objective)



$x^\star$ is optimal if $-\nabla f_0(x^\star)$ has an angle of more than 90° with $(x - x^\star)$ for every feasible x.

Special case:

Unconstrained problem:

$$x^\star \text{ optimal } \iff x^\star \in \mathrm{dom}(f_0),\ \nabla f_0(x^\star) = 0$$

Equivalent Optimization Problems

Two problems P and P' are *equivalent* (written as P $\iff$ P') if the solution of P' is obtained from that of P via "elementary transformations" and vice versa.

Epigraphical reformulation: The standard form convex optimization problem is equivalent to

$$
\begin{array}{lll}
\text{P'}: & \text{minimize (over } x, t) & t \\
& \text{subject to} & f_0(x) - t \leq 0 \\
& & f_i(x) \leq 0 \quad \forall i = 1, \ldots, m \\
& & Ax = b
\end{array}
$$

(i.e., one can always assume that the objective is linear.)

Equivalent Optimization Problems

Partial minimization: The problem

$$\begin{array}{lll} \text{P:} & \text{minimize (over } x_1, x_2) & f_0(x_1, x_2) \\ & \text{subject to} & f_i(x_1) \leq 0 \quad \forall i = 1, \ldots, m \\ & & g_i(x_2) \leq 0 \quad \forall i = 1, \ldots, q \end{array}$$

is equivalent to

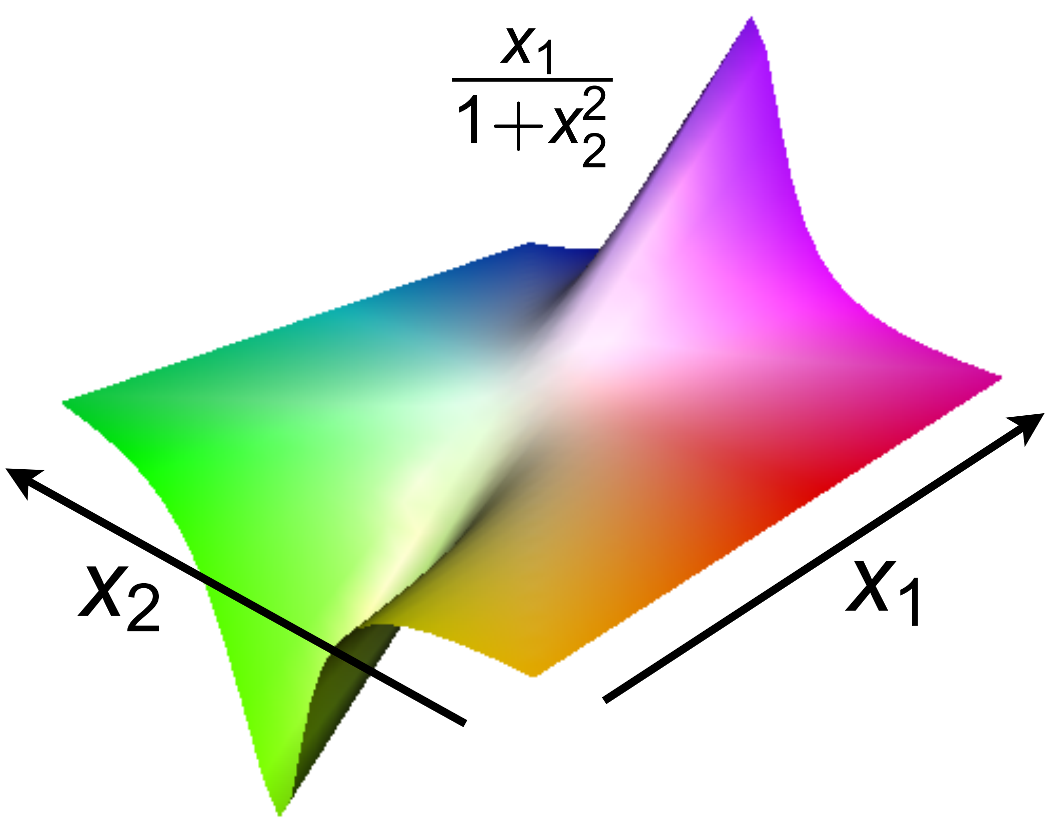
$$\begin{array}{lll} \text{P':} & \text{minimize (over } x_1) & \hat{f}_0(x_1) \\ & \text{subject to} & f_i(x_1) \leq 0 \quad \forall i = 1, \ldots, m \end{array}$$

where

$$\hat{f}_0(x_1) = \inf_{x_2} \{ f_0(x_1, x_2) : g_i(x_2) \leq 0 \ \forall i = 1, \ldots, q \}.$$

Equivalent Optimization Problems

Example: The following problems are equivalent.



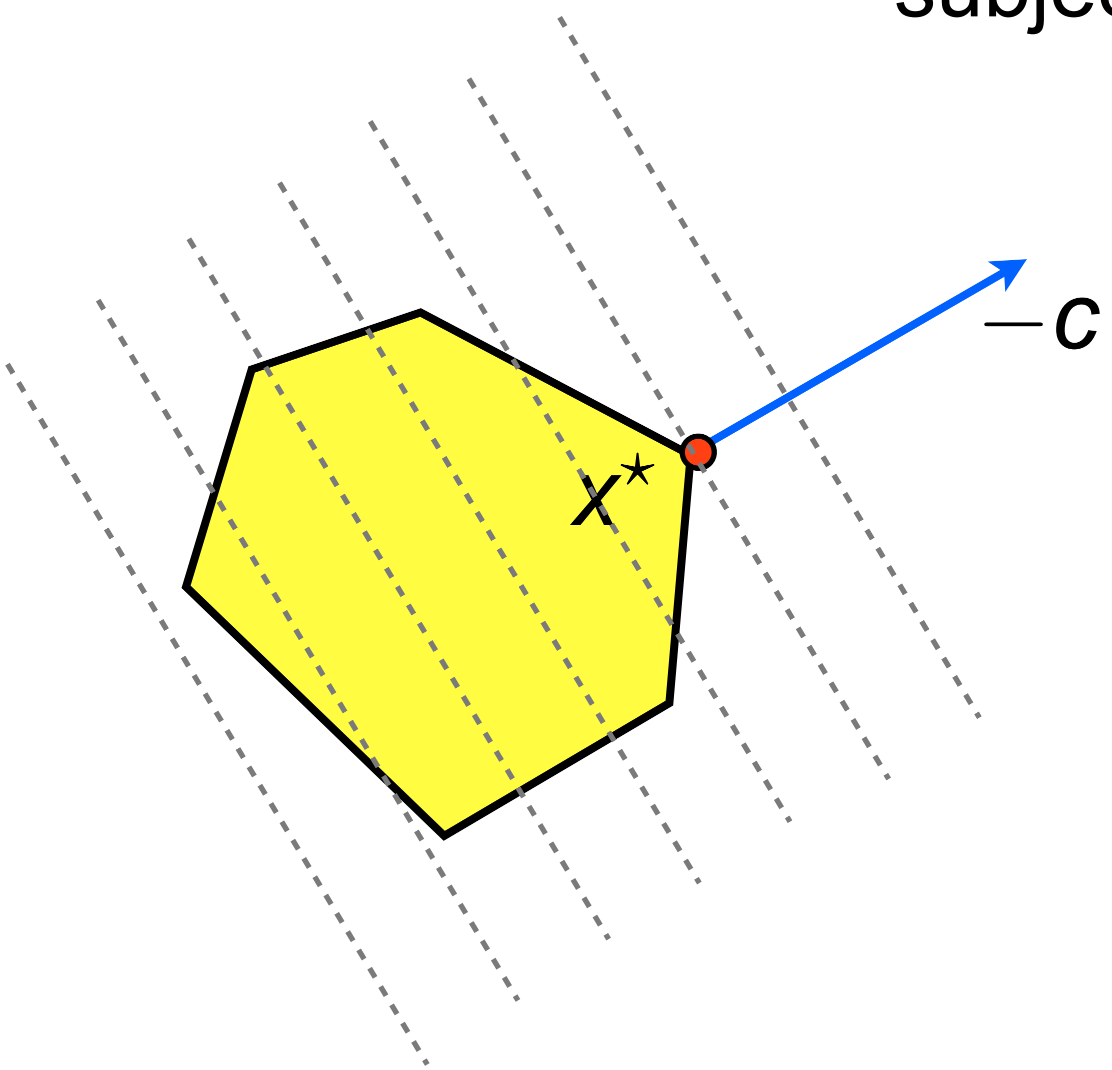
$$\text{P:} \quad \begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1/(1+x_2^2) \le 0 \\ & (x_1+x_2)^2 = 0 \end{array}$$

$$\text{P':} \quad \begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \le 0 \\ & x_1+x_2 = 0 \end{array}$$

However, P' is a convex optimization problem according to our definition, but P is not. Indeed, P involves a nonconvex inequality constraint function and a nonlinear equality constraint function.

Linear Program (LP)

$$\begin{array}{ll} \text{minimize} & c^\top x + d \\ \text{subject to} & Ax = b \\ & Cx \leq g \end{array}$$

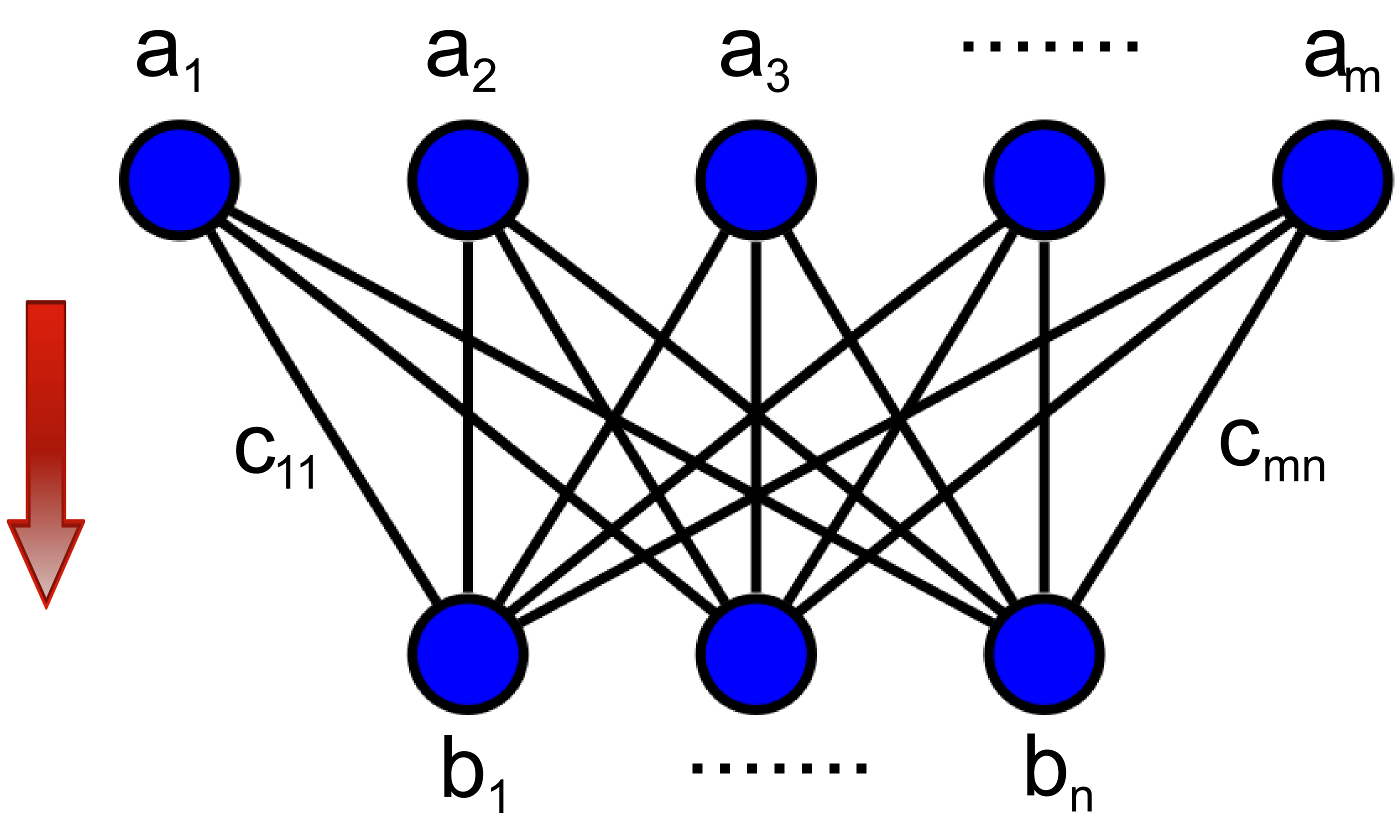


Affine objective and constraints, polyhedral feasible set.

Problems with n variables and encoded in L input bits can be solved in $\mathcal{O}(n^{3.5}L)$ arithmetic operations via interior point methods.

Transportation Problem

minimize	$\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$		transportation costs
subject to	$\sum_{j=1}^{n} x_{ij} = a_i$	$\forall i$	total shipment from source i
	$\sum_{i=1}^{m} x_{ij} = b_j$	$\forall j$	total shipment to destination j
	$x_{ij} \geq 0$	$\forall i, j$	nonnegative shipments only



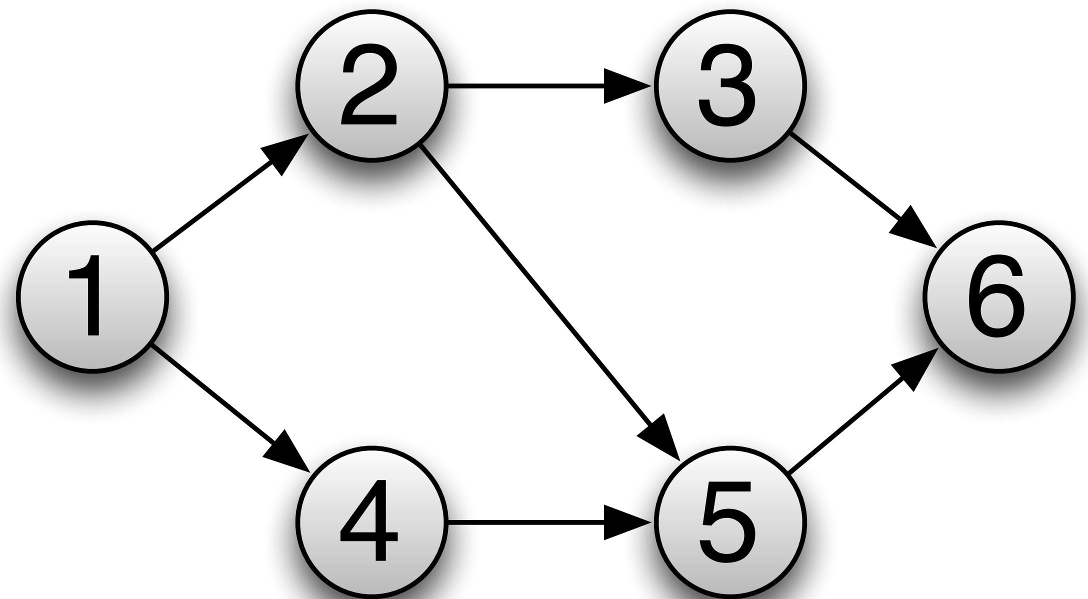
Problem feasible if total supply = total demand:

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

Project Scheduling

Project: directed acyclic graph $G = (V, E)$

- *tasks* $V = \{1, \ldots, n\}$
- *precedences* $E \subseteq V \times V$
- *task durations* $\zeta_v, \; v \in V$

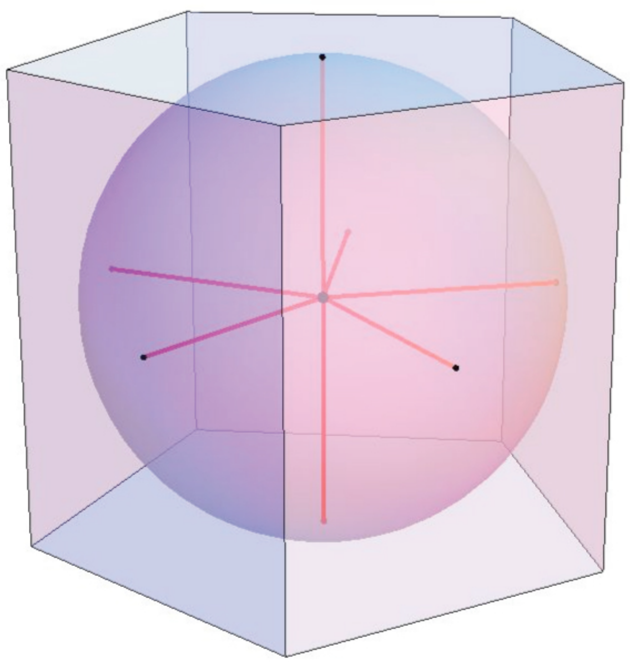


Early Start Policy:

- find task start times *via an LP*

$$\begin{aligned} \min_{s \in \mathbb{R}^n_+} \quad & s_n + \zeta_n \\ \text{s.t.} \quad & s_v \geq s_u + \zeta_u \quad \forall (u, v) \in E \end{aligned}$$

Chebyshev Center of a Polyhedron



The Chebyshev center of a polyhedron

$$\mathcal{P} = \{x : a_i^\top x \leq b_i,\ i = 1, \ldots, m\}$$

is the center of the largest ball inside $\mathcal{P}$

$$\begin{array}{ll} \underset{x_c, r}{\text{maximize}} & r \\ \text{subject to} & \underbrace{a_i^\top x \leq b_i \quad \forall x \in \mathcal{B}(x_c, r)}, \ i = 1 \ldots, m \end{array}$$

$$\begin{aligned} \iff \quad & a_i^\top (x_c + u) \leq b_i \qquad \forall u : \|u\|_2 \leq r \\ \iff \quad & \max_{\|u\|_2 \leq r} a_i^\top (x_c + u) \leq b_i \\ \iff \quad & a_i^\top x_c + r\|a_i\|_2 \leq b_i \end{aligned}$$

Quadratic Program (QP)

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}x^\top P x + q^\top x + r \\
\text{subject to} & Ax = b \\
& Cx \leq g
\end{array}
$$

Convex quadratic objective ($P \in \mathbb{S}^n_+$) and affine constraints, i.e., a polyhedral feasible set.

Example: Least-squares problem:

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

If A has full column rank, we have

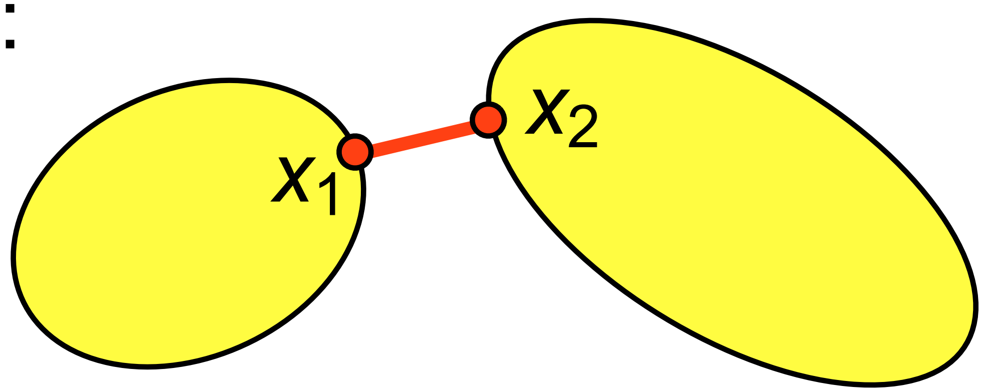
$$
\begin{array}{c}
0 = \nabla \|Ax - b\|_2^2 = 2A^\top Ax - 2A^\top b \\
\implies \quad x = (A^\top A)^{-1} A^\top b.
\end{array}
$$

Quadratically Constrained QP (QCQP)

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

Convex quadratic objective and constraints ($P_i \in \mathbb{S}^n_+$), i.e., the feasible set is an intersection of ellipsoids (for $P_i \succ 0$) and halfspaces (for $P_i = 0$). Every LP is a QCQP with $P_i = 0$.

Example: Distance between ellipsoids:



$$\begin{array}{ll} \text{minimize} & \|x_1 - x_2\|_2^2 \\ \text{subject to} & (x_1 - \mu_1)^\top \Sigma_1^{-1}(x_1 - \mu_1) \leq 1 \\ & (x_2 - \mu_2)^\top \Sigma_2^{-1}(x_2 - \mu_2) \leq 1 \end{array}$$

Second-Order Cone Program (SOCP)

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

$$(f \in \mathbb{R}^n,\ A_i \in \mathbb{R}^{n_i \times n},\ b_i \in \mathbb{R}^{n_i},\ c_i \in \mathbb{R}^n,\ d_i \in \mathbb{R},\ F \in \mathbb{R}^{p \times n},\ g \in \mathbb{R}^p)$$

The inequalities are called *second-order cone constraints* as the $(n_i + 1)$-dimensional vector $(A_i x + b_i, c_i^\top x + d_i)$ belongs to the second-order cone in $\mathbb{R}^{n_i+1}$.

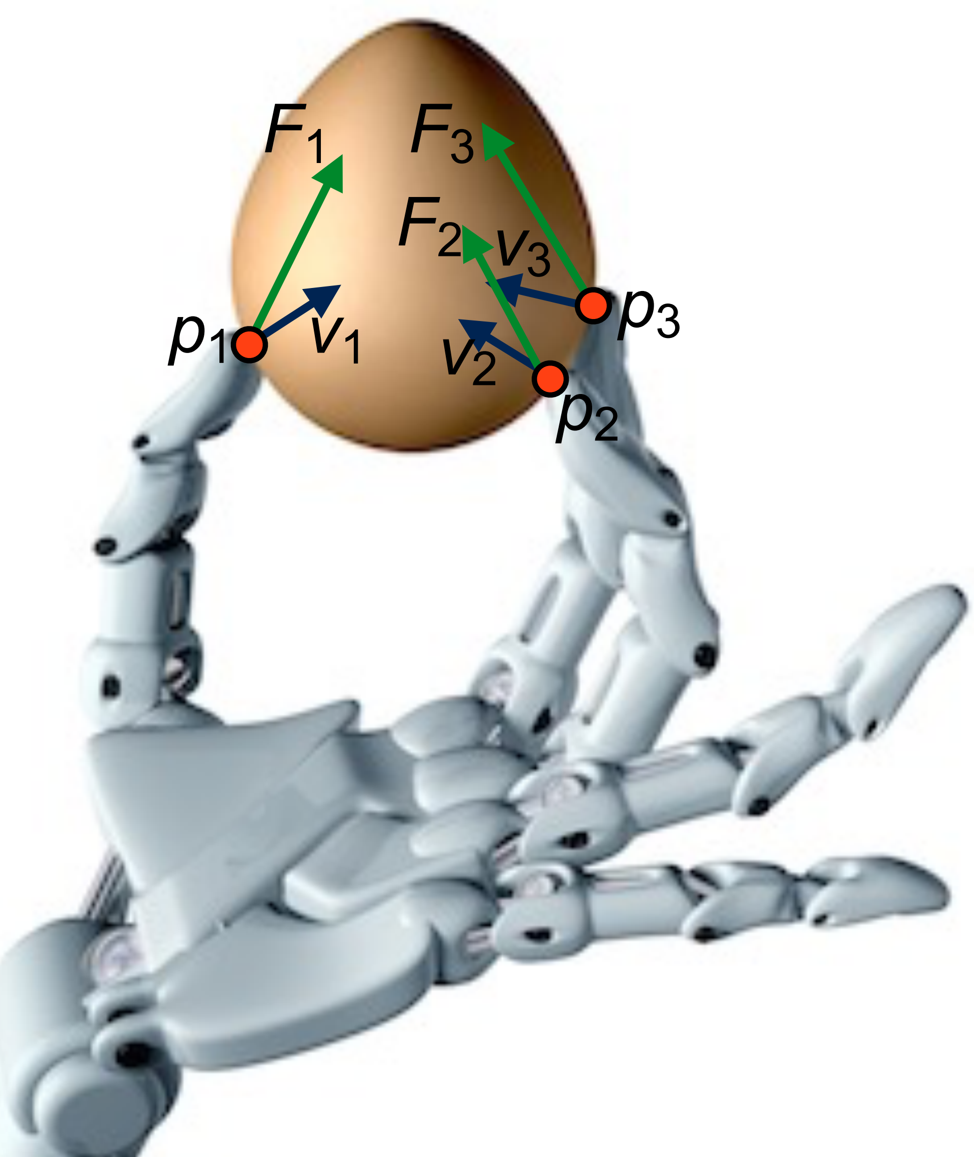
Every LP is an SOCP with $A_i = 0,\ b_i = 0$. Every SOCP with $c_i = 0$ is a QCQP, and every QCQP is an SOCP.

An ε-optimal solution can be found in $\mathcal{O}((\sum_{i=1}^m n_i)^{3.5} \log(\varepsilon^{-1}))$ arithmetic operations via interior point methods.

Grasping Force Optimization

Consider a *rigid body* held by N *robotic fingers*.

- Center of gravity at 0
- Contact points $p_1, \ldots, p_N \in \mathbb{R}^3$
- Contact forces $F_1, \ldots, F_N \in \mathbb{R}^3$
- inward pointing unit normal vectors $v_1, \ldots, v_N \in \mathbb{R}^3$



Friction-cone constraints:

$$\underbrace{\left\| (I - v_i v_i^\top) F_i \right\|_2}_{\text{tangential force}} \leq \underbrace{\mu}_{\text{friction coefficient}} \cdot \underbrace{v_i^\top F_i}_{\text{normal force}}$$

Static equilibrium constraints:

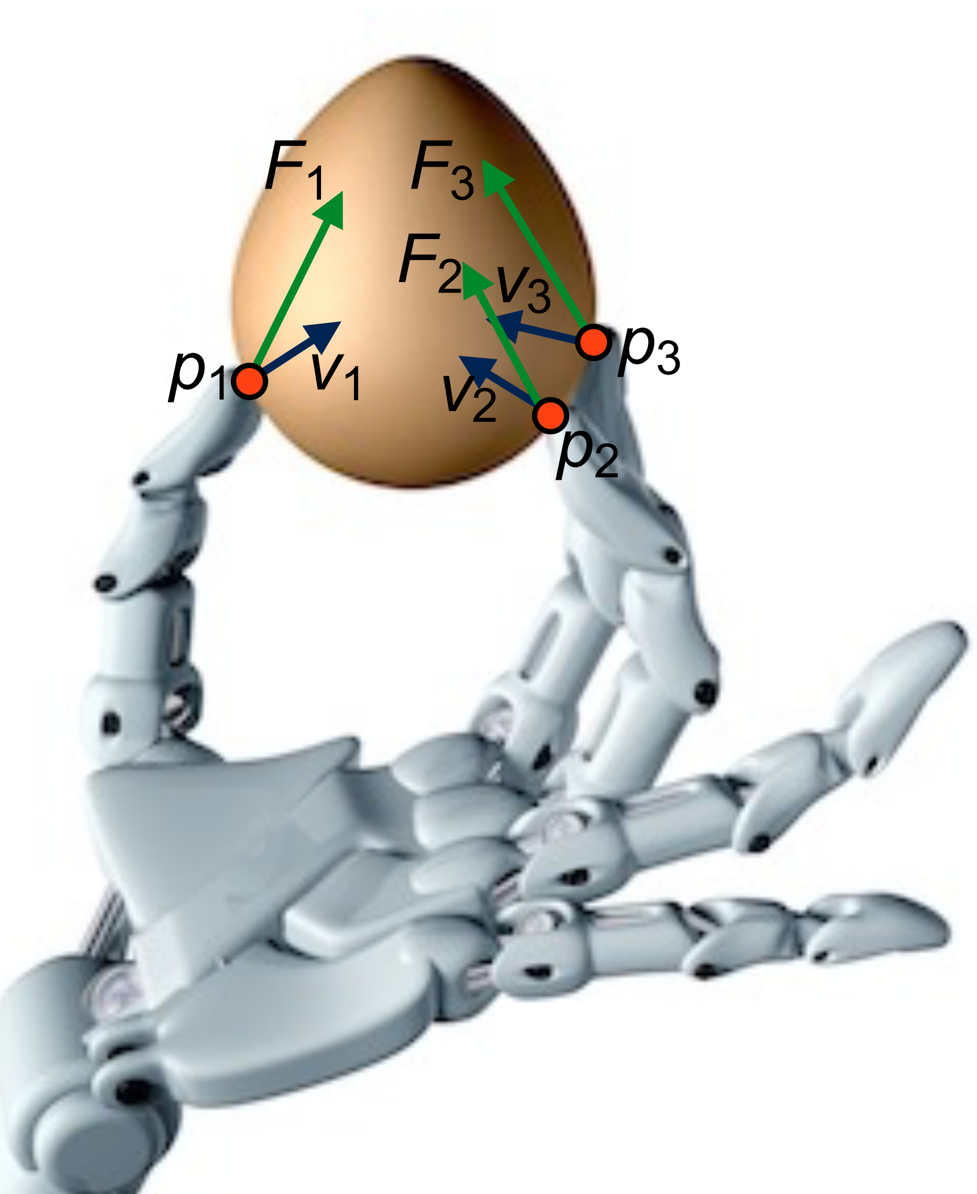
$$\textstyle\sum_{i=1}^N F_i = -F_{\text{ext}} \longleftarrow \text{external force}$$

$$\textstyle\sum_{i=1}^N p_i \times F_i = -T_{\text{ext}} \longleftarrow \text{external torque}$$

Grasping Force Optimization

Consider a *rigid body* held by N *robotic fingers*.

- Center of gravity at 0
- Contact points $p_1, \ldots, p_N \in \mathbb{R}^3$
- Contact forces $F_1, \ldots, F_N \in \mathbb{R}^3$
- inward pointing unit normal vectors $v_1, \ldots, v_N \in \mathbb{R}^3$



Minimize maximal normal force:

$$
\begin{array}{ll}
\underset{t, F_i}{\text{minimize}} & t \\
\text{subject to} & v_i^\top F_i \leq t \ \ \forall i \\
& \left\| (I - v_i v_i^\top) F_i \right\|_2 \leq \mu \cdot v_i^\top F_i \ \ \forall i \\
& \sum_{i=1}^N F_i = -F_{\text{ext}} \\
& \sum_{i=1}^N p_i \times F_i = -T_{\text{ext}}
\end{array}
$$

“Hidden” SOCP Constraints

Quadratic constraints:

$$\|x\|_2^2 \leq t \quad \Longleftrightarrow \quad \left\| \begin{pmatrix} 2x \\ t-1 \end{pmatrix} \right\|_2 \leq t+1$$

Hyperbolic constraints:

$$\|x\|_2^2 \leq st, \quad s,t \geq 0 \quad \Longleftrightarrow \quad \left\| \begin{pmatrix} 2x \\ s-t \end{pmatrix} \right\|_2 \leq s+t, \quad s,t \geq 0$$

Problems with Generalized Inequalities

Standard form convex optimization problem with *generalized inequalities*:

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

Assumptions:

- $f_0 : \mathbb{R}^n \to \mathbb{R}$ convex, $f_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ K_i-convex
- $K_i \subseteq \mathbb{R}^{k_i}$ is a proper convex cone

Observations:

- P has a convex feasible set
- All local optima of P are global optima

Conic Form Problems

Special case where *objective* and *constraints* are *affine* in x.

$$\begin{array}{lll} \mathrm{P}: & \text{minimize} & c^\top x \\ & \text{subject to} & Fx + g \preceq_K 0 \\ & & Ax = b \end{array}$$

The conic form problem *reduces to an LP* if $K = \mathbb{R}^m_+$.

Note also that the SOCP

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

is equivalent to

$$\begin{array}{lll} \text{minimize} & f^\top x & \\ \text{subject to} & -(A_i x + b_i, c_i^\top x + d_i) \preceq_{K_i} 0 & i = 1, \ldots, m \\ & Fx = g & \end{array}$$

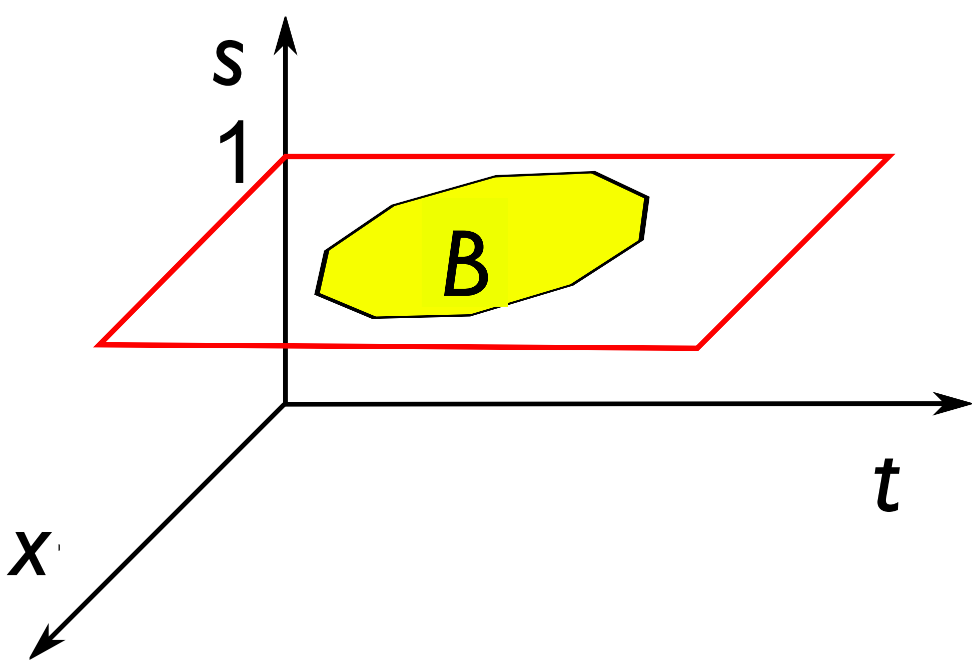
where $K_i = \{(y, t) \in \mathbb{R}^{n_i+1} : \|y\|_2 \leq t\}$.

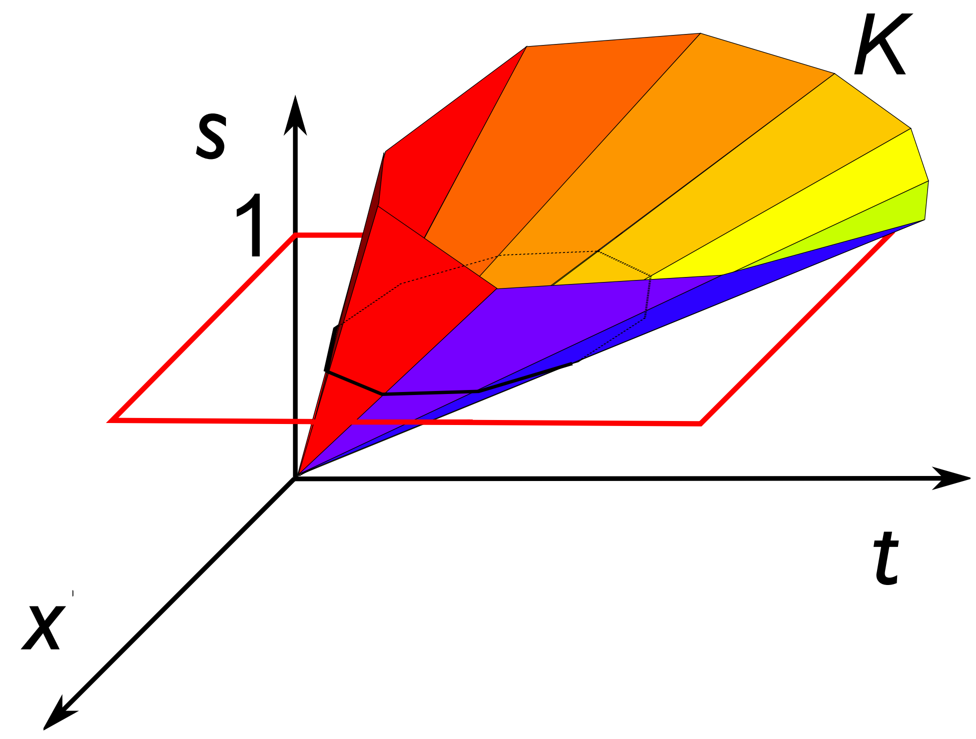
Every Convex Problem is Conic

Theorem: Every convex optimization problem can be reformulated as a conic form problem.

Proof idea: Consider an arbitrary convex problem $\inf_{x \in C} f_0(x)$, where C is a convex feasible set. This problem is equivalent to

$$
\begin{array}{lll}
 & \inf\{t : f_0(x) \le t,\ x \in C\} & \\
\iff & \inf\{t : (t,x) \in B\} & B = \{(t,x) : f_0(x) \le t,\ x \in C\} \\
\iff & \inf\{t : s = 1,\ (s,t,x) \succeq_K 0\} & K = \{(s,t,x) : (t,x)/s \in B,\ s \ge 0\}.
\end{array}
$$





Note: K is a convex cone (why?)

Semidefinite Program (SDP)

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

$$(c \in \mathbb{R}^n,\ F_1, \ldots, F_n, G \in \mathbb{S}^k,\ A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m)$$

The semidefinite constraint *linear matrix inequality* (LMI). Note that several LMIs can be combined to a single LMI:

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

is equivalent to

$$\begin{pmatrix} \tilde{F}_1 & 0 \\ 0 & \hat{F}_1 \end{pmatrix} x_1 + \cdots + \begin{pmatrix} \tilde{F}_n & 0 \\ 0 & \hat{F}_n \end{pmatrix} x_n \preceq \begin{pmatrix} \tilde{G} & 0 \\ 0 & \hat{G} \end{pmatrix}.$$

An ε-optimal solution can be found in $\mathcal{O}(n^2 k^{2.5} \log(\varepsilon^{-1}))$ arithmetic operations via interior point methods.

Every SOCP is an SDP

The SOCP

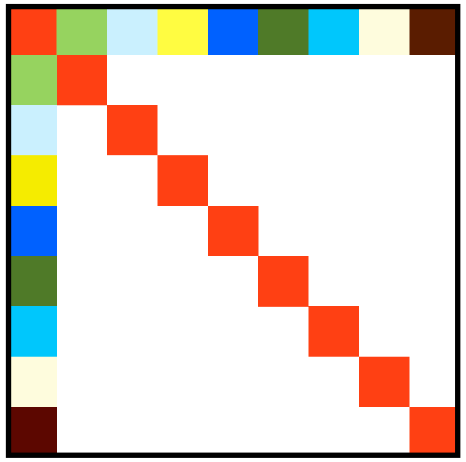
$$\begin{array}{ll} \text{minimize} & f^\top x \\ \text{subject to} & \|A_i x + b_i\|_2 \le c_i^\top x + d_i \quad i = 1, \dots, m \\ & Fx = g \end{array}$$

is equivalent to an SDP as by Schur's lemma

$$\|A_i x + b_i\|_2 \le c_i^\top x + d_i$$

$$\iff \quad \begin{pmatrix} c_i^\top x + d_i & (A_i x + b_i)^\top \\ A_i x + b_i & (c_i^\top x + d_i) I \end{pmatrix} \succeq 0$$

Arrow matrix:



Every LP is an SDP

We have already seen that every LP is an SOCP and every SOCP is an SDP. Thus, every LP is an SDP.

Direct argument: The LP

$$\begin{array}{ll} \text{minimize} & c^\top x + d \\ \text{subject to} & Ax = b \\ & Cx \leq g \end{array}$$

is equivalent to the SDP

$$\begin{array}{ll} \text{minimize} & c^\top x + d \\ \text{subject to} & \begin{pmatrix} \operatorname{diag}(b - Ax) & & \\ & \operatorname{diag}(Ax - b) & \\ & & \operatorname{diag}(g - Cx) \end{pmatrix} \succeq 0. \end{array}$$

Eigenvalue Minimization

Minimize the largest eigenvalue of a matrix, i.e., solve

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

where $A(x) = A_0 + A_1 x_1 + \cdots + A_n x_n,\ A_i \in \mathbb{S}^k\ \forall i = 0, \ldots, n$.

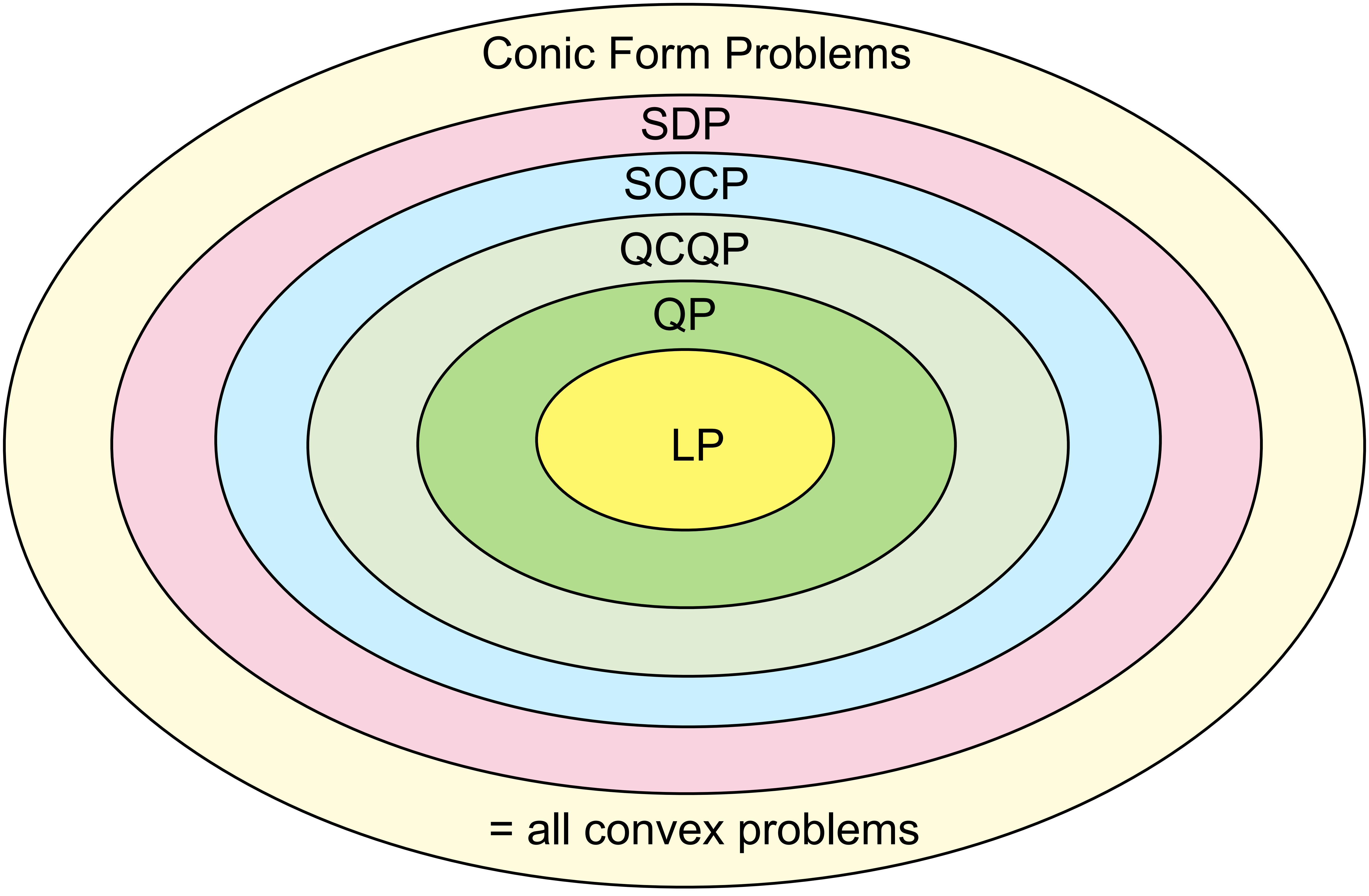
This problem is equivalent to the SDP

$$\underset{x,t}{\text{minimize}}\ \{t : A(x) \preceq tI\}.$$

Consider $X = R\Lambda R^\top \in \mathbb{S}^k$ with $\Lambda = \operatorname{diag}(\lambda_1(X), \ldots, \lambda_k(X))$ and R orthogonal ($R^\top = R^{-1}$).

$$\begin{aligned}
\lambda_{\max}(X) \le t \quad &\iff \quad \lambda_i(X) \le t \quad \forall i = 1, \ldots, k \\
&\iff \quad \Lambda \preceq tI \\
&\iff \quad X = R\Lambda R^\top \preceq tRIR^\top = tI
\end{aligned}$$

Hierarchy of Convex Optimization Problems



Popular Solvers

CPLEX (https://www.ibm.com/analytics/cplex-optimizer):

- caters for LP, QP, SOCP and MILP (mixed-integer LP); free for academic use

Gurobi (www.gurobi.com)

- caters for LP, QP, SOCP and MISOCP (mixed-integer SOCP); can be deployed on the cloud; free for academic use

MOSEK (http://www.mosek.com)

- caters for LP, QP, SOCP, SDP and MISDP (mixed-integer SDP); ideal for large-scale sparse problems; free for academic use

SDPT3 (http://www.math.nus.edu.sg/~mattohkc/sdpt3.html):

- caters for SDP; free, runs on Matlab

IPOPT (https://projects.coin-or.org/Ipopt)

- caters for general NLPs (nonlinear programs); free

YALMIP (https://yalmip.github.io) is a modelling language for optimization problems implemented as a free toolbox for Matlab.

A comprehensive list of solvers that can be used with Yalmip is available from https://yalmip.github.io/allsolvers/.

Main Take-Away Points

- **Convex optimization problems:** convex objective and inequality constraints, linear equality constraints; feasible set is always convex; every local minimizer is a global minimizer
- **Optimality criterion:** $x^\star$ is optimal iff the gradient of the objective makes an acute angle with every feasible direction
- **Problem classes:** definitions of LP, QP, QCQP, SOCP, SDP and conic form problems; (strict!) subset relations between problem classes