

4. Convex optimization problems

- optimization problem in standard form
- convex optimization problems
- linear optimization
- quadratic optimization
- semidefinite programming
- vector optimization

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:

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

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

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

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

- 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

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

Local and global optima

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

proof: suppose x is locally optimal, but there exists a feasible y 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

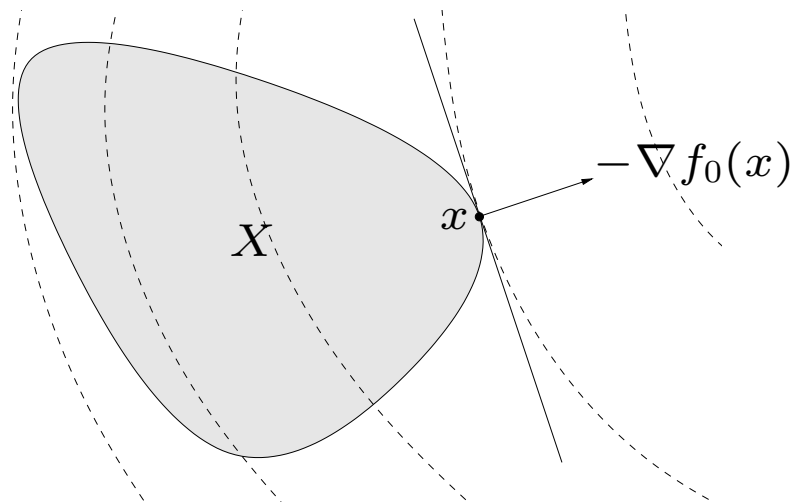
$$f_0(z) \leq \theta f_0(y) + (1 - \theta)f_0(x) < f_0(x)$$

which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

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



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

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

- **equality constrained problem**

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

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

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

- **minimization over nonnegative orthant**

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

x is optimal if and only if

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

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**

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

is equivalent to

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

where F and x_0 are such that

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

- **introducing equality constraints**

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

is equivalent to

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

- **introducing slack variables for linear inequalities**

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

is equivalent to

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

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

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

- **minimizing over some variables**

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

is equivalent to

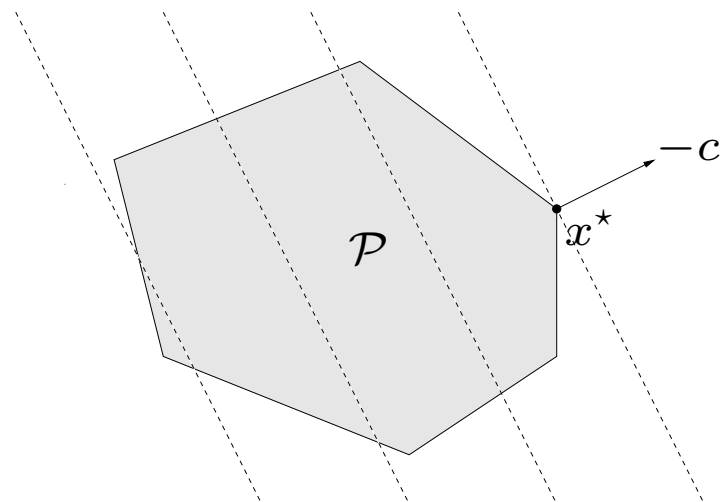
$$\begin{array}{ll}
 \text{minimize} & \tilde{f}_0(x_1) \\
 \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \dots, m
 \end{array}$$

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

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



Examples

diet problem: choose quantities x_1, \dots, 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,

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

piecewise-linear minimization

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

equivalent to an LP

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

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

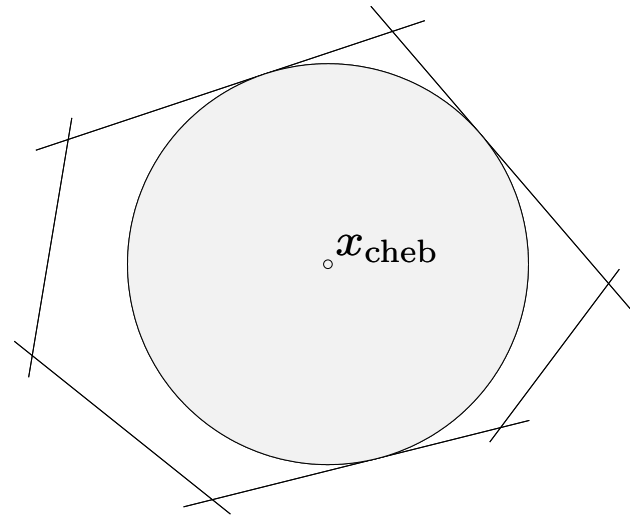
$$\mathcal{B} = \{x_c + u \mid \|u\|_2 \leq r\}$$

- $a_i^T x \leq b_i$ 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

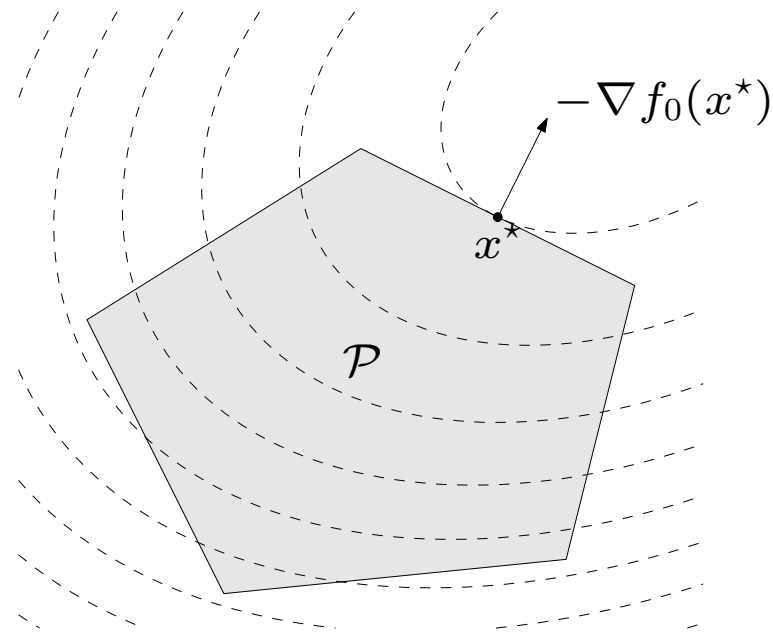
$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \leq b_i, \quad i = 1, \dots, m \end{array}$$



Quadratic program (QP)

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

- $P \in \mathbf{S}_+^n$, 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)
- can add linear constraints, *e.g.*, $l \preceq x \preceq u$

linear program with random cost

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

- 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)

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

- $P_i \in \mathbf{S}_+^n$; 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

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

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

- 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

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

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

is equivalent to the SOCP

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

(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

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

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

- robust LP

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

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

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

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,

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

is equivalent to single LMI

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

LP and SOCP as SDP

LP and equivalent SDP

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

(note different interpretation of generalized inequality \preceq)

SOCP and equivalent SDP

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

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

Eigenvalue minimization

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

where $A(x) = A_0 + x_1A_1 + \cdots + x_nA_n$ (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

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

Matrix norm minimization

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

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

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

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

$$\begin{aligned} \|A\|_2 \leq t & \iff A^T A \preceq t^2 I, \quad t \geq 0 \\ & \iff \begin{bmatrix} tI & A \\ A^T & tI \end{bmatrix} \succeq 0 \end{aligned}$$