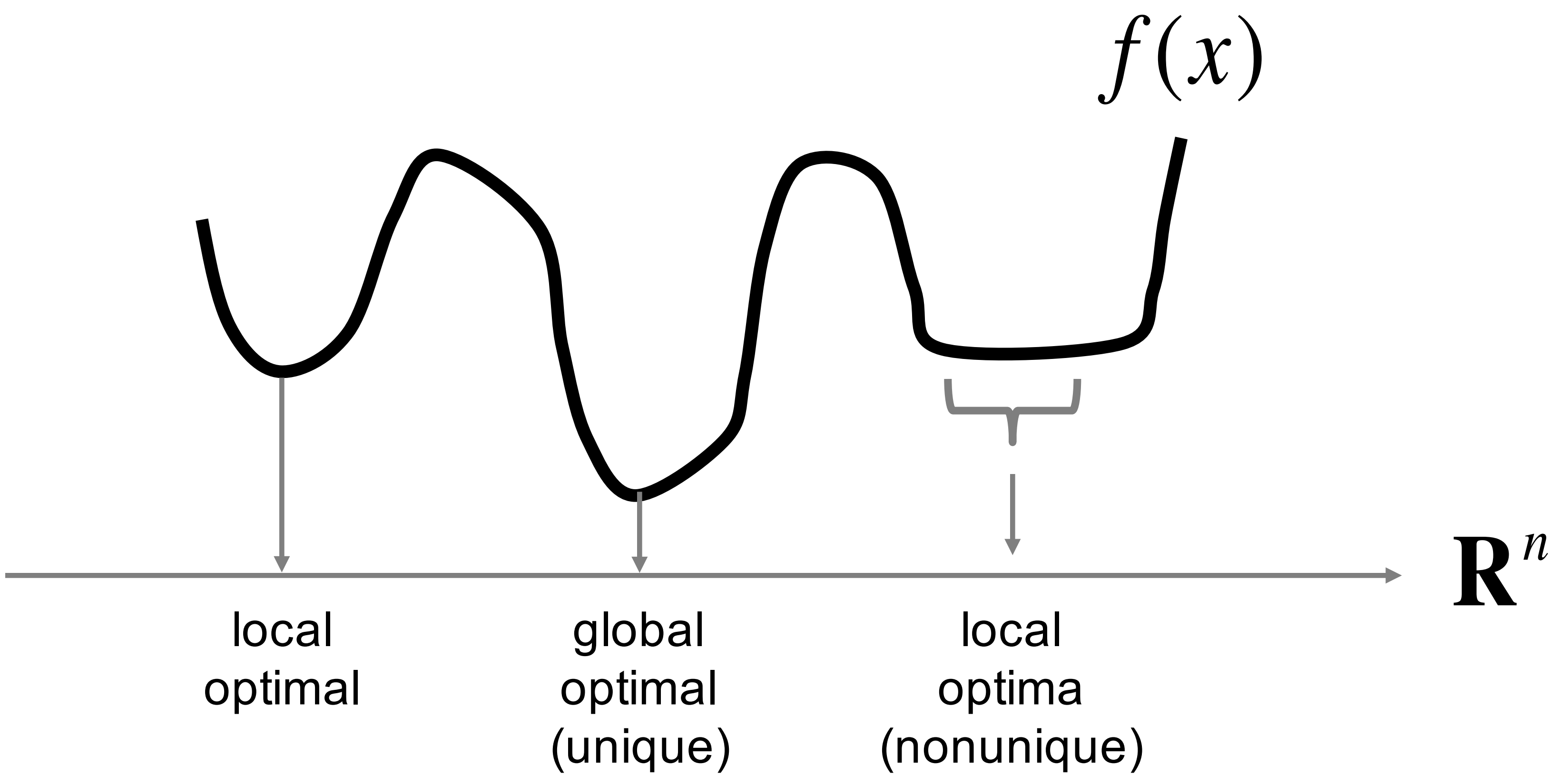
Convex optimization

References:
Boyd and Vandenberghe, Convex optimization, 2004
Ben-Tal and Nemirovski, Lectures on modern convex optimization, 2013

Unconstrained optimization

$$\min_{x \in \mathbf{R}^n} \quad f(x)$$

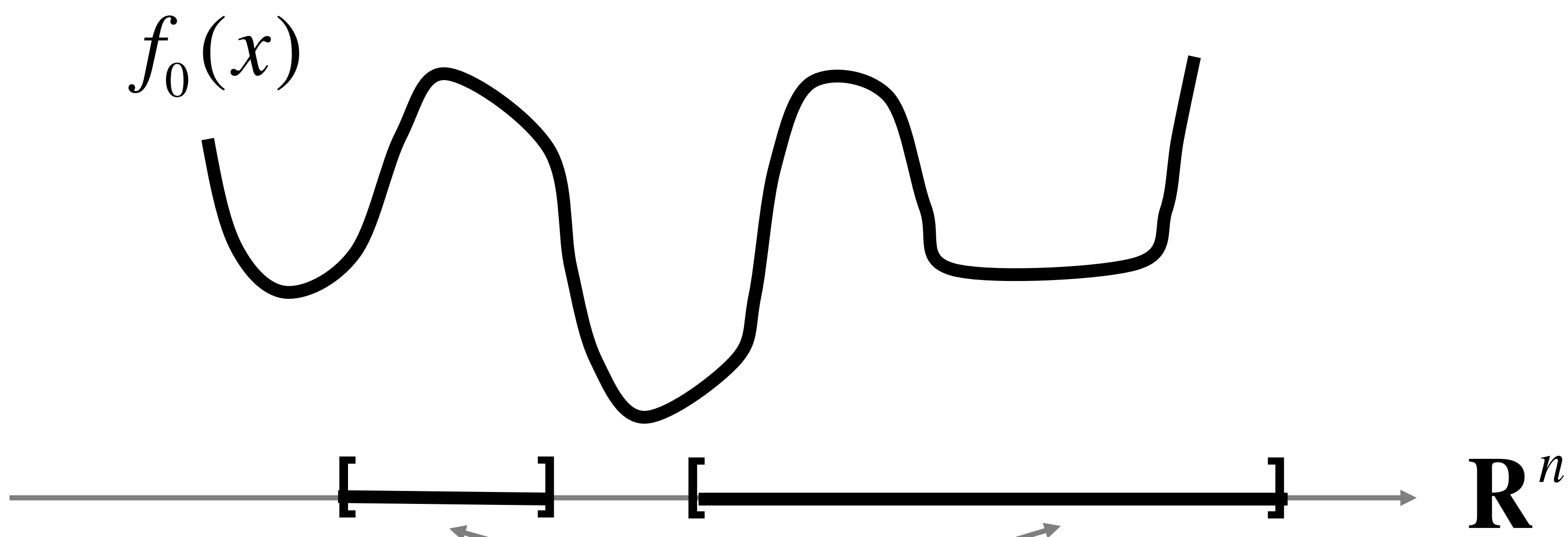


Constrained optimization

$$\min_{x \in \mathbf{R}^n} \quad f_0(x)$$

$$\text{s. t.} \quad f_k(x) \le 0, \qquad k = 1, \ldots, K$$

$$g_k(x) = 0, \qquad k = 1, \ldots, M$$

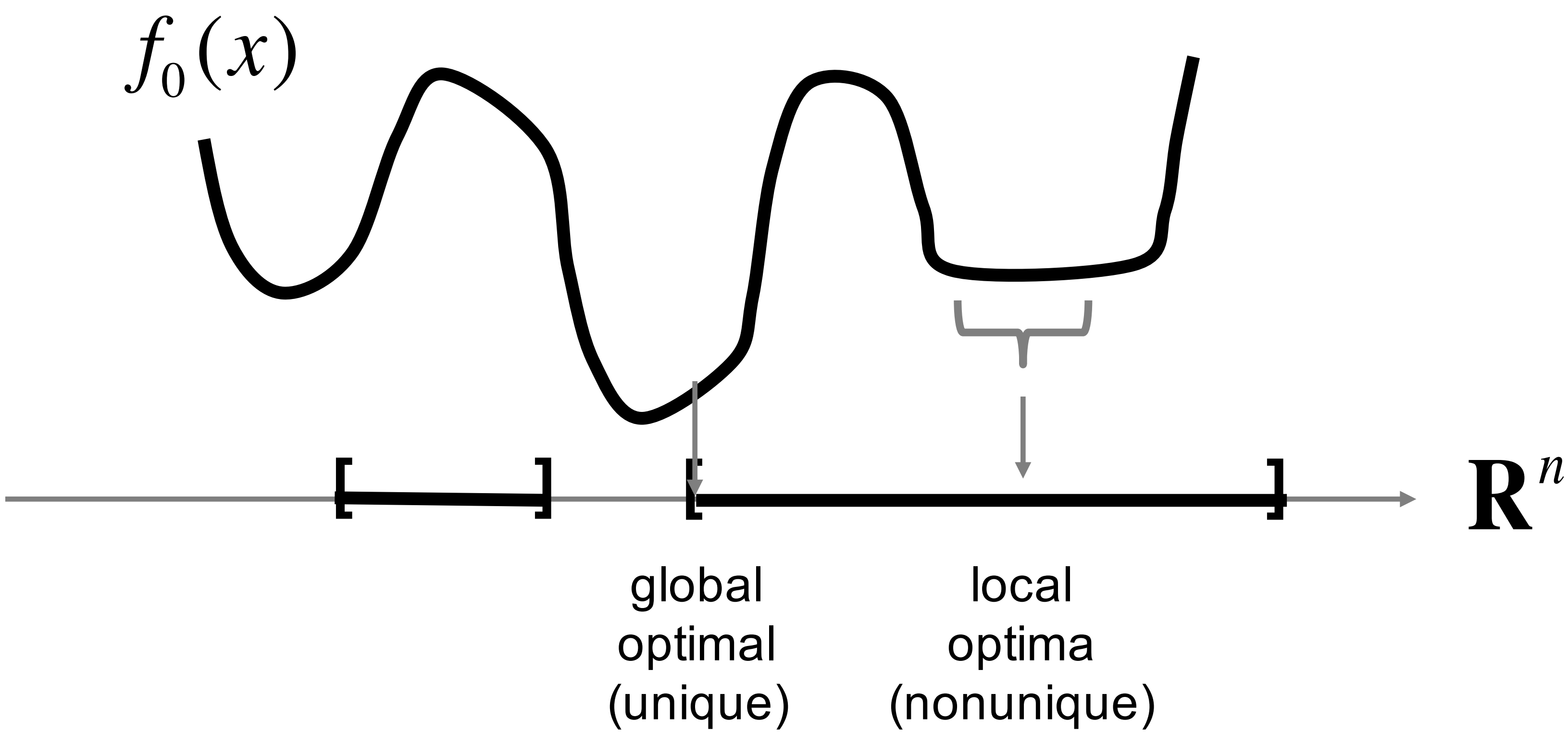


feasible set

$$\mathbf{X} := \left\{ x \in \mathbf{R}^n \,\middle|\, f_k(x) \le b_k,\ k = 1, \ldots, K \right\}$$

Constrained optimization

$$
\begin{aligned}
\min_{x \in \mathbf{R}^n} \quad & f_0(x) \\
\text{s. t.} \quad & f_k(x) \le 0, \quad k = 1, \dots, K \\
& g_k(x) = 0, \quad k = 1, \dots, M
\end{aligned}
$$



Constrained optimization

$$\begin{aligned} &\min_{x \in \mathbf{R}^n} && f_0(x) \\ &\text{s. t.} && f_k(x) \le 0, \quad k = 1, \ldots, K \\ & && g_k(x) = 0, \quad k = 1, \ldots, M \end{aligned}$$

Definition

- (Global) minimizers/optima:

$$\mathbf{X}^* := \left\{ x^* \in \mathbf{X} \,\middle|\, f_0(x^*) \le f_0(x) \quad \forall x \in \mathbf{X} \right\}$$

- A minimizer $x^* \in \mathbf{X}$ is *unique* if

$$f_0(x^*) < f_0(x) \quad \forall x \in \mathbf{X}$$

Convex optimization

$$\begin{aligned} \min_{x \in \mathbf{R}^n} \quad & f_0(x) \\ \text{s. t.} \quad & f_k(x) \le 0, \quad k = 1, \ldots, K \\ & Ax = b \end{aligned}$$

Definition

Convex optimization if

- $f_k(x)$ are *convex functions* for $k = 0, 1, \ ..., K$

The feasible set $\mathbf{X}$ is a *convex set*

Convex optimization is polynomial-time

Convex optimization

$$
\begin{aligned}
\min_{x \in \mathbf{R}^n} \quad & f_0(x) \\
\text{s. t.} \quad & f_k(x) \le 0, \qquad k = 1, \ldots, K \\
& Ax = b
\end{aligned}
$$

Questions

- How to recognize a convex program
- How to characterize minimizers?
- How to compute a minimizer?

Convex optimization

$$\begin{aligned} \min_{x \in \mathbf{R}^n} \quad & f_0(x) \\ \text{s. t.} \quad & f_k(x) \le 0, \quad k = 1, \ldots, K \\ & Ax = b \end{aligned}$$

Questions

- How to recognize a convex program
  - Convex set, convex function
- How to characterize minimizers?
  - KKT condition, duality theorem
- How to compute a minimizer?
  - First-order algorithms, Newton algorithms
  - Distributed algorithms

Convex optimization

- Convex set
- Convex function
- Duality and KKT condition
- Algorithms

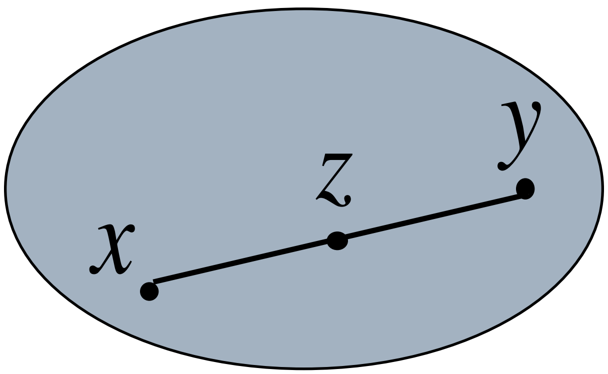
References:

Boyd and Vandenberghe, Convex optimization, 2004
Ben-Tal and Nemirovski, Lectures on modern convex optimization, 2013

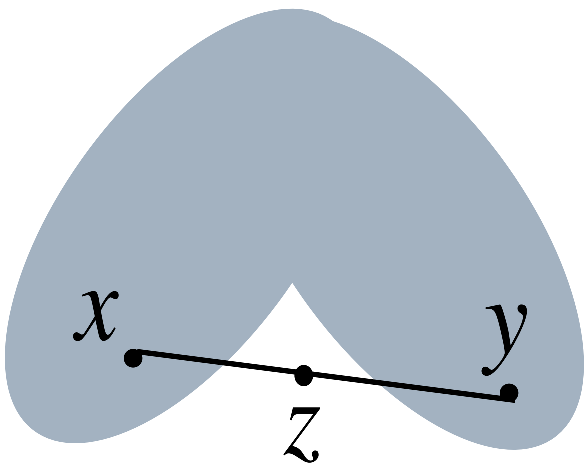
Convex set

Definition

A set S is convex if for all $x, y \in S$
for all $\alpha \in [0,1]$, $z := \alpha x + (1-\alpha)y \in S$



S is convex



S is nonconvex

S can be in an arbitrary space, not necessarily in $\mathbf{R}^n$

Convex set

Examples

- Half-plane or half-space

$$S := \left\{ x \in \mathbf{R}^n \,\middle|\, a^T x = b \right\} \qquad (a \neq 0)$$

$$S := \left\{ x \in \mathbf{R}^n \,\middle|\, a^T x \leq b \right\} \qquad (a \neq 0)$$

Convex set

Examples

- Norm ball (any norm)

$$S := \left\{ x \in \mathbf{R}^n \;\middle|\; \|x - c\| \le r \right\} = \left\{ c + ru \;\middle|\; \|u\| \le 1 \right\}$$

- Ellipsoid

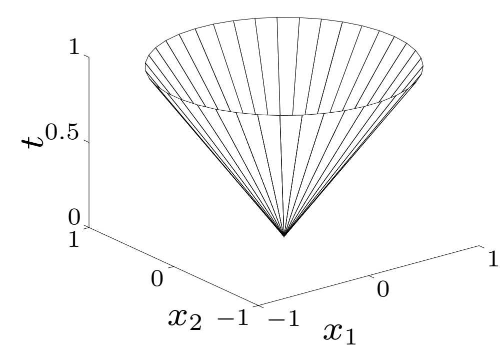
$$S := \left\{ x \in \mathbf{R}^n \;\middle|\; (x - c)^T P^{-1} (x - c) \le 1 \right\} \quad (P \text{ pd})$$

$$S := \left\{ c + Au \;\middle|\; \|u\|_2 \le 1 \right\} \quad (A \text{ nonsingular})$$

- Norm cone (any norm)

$$S := \left\{ (x,t) \in \mathbf{R}^{n+1} \;\middle|\; \|x\| \le t \right\}$$

$\|.\|_2$: second-order cone



Convex set

Examples: polyhedra

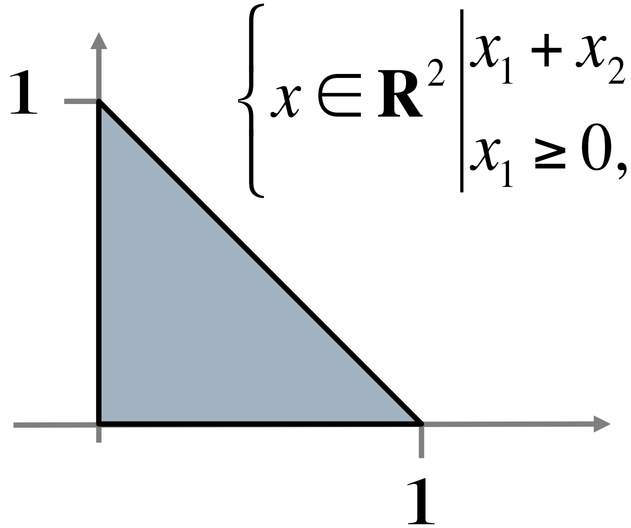
- Linear equality and inequality:

$$S := \left\{ x \in \mathbf{R}^n \,\middle|\, A_1 x = b_1,\ \mathrm{A}_2 x \geq b_2 \right\}$$

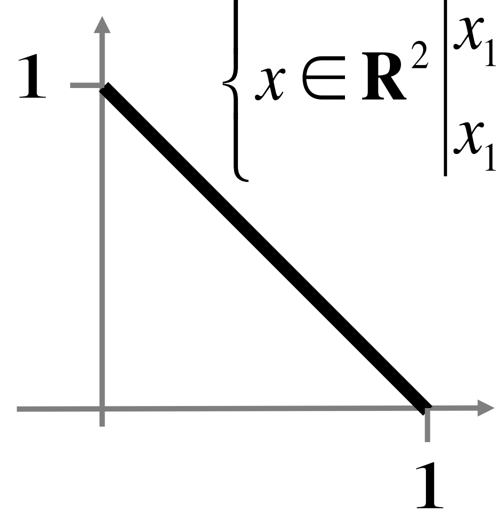
for some $A_j \in \mathbf{R}^{m \times n},\ b_j \in \mathbf{R}^m$

Polyhedron: finite intersection of halfplanes and halfspaces

$$\left\{ x \in \mathbf{R}^2 \,\middle|\, \begin{array}{l} x_1 + x_2 \leq 1, \\ x_1 \geq 0, x_1 \geq 0 \end{array} \right\}$$



$$\left\{ x \in \mathbf{R}^2 \,\middle|\, \begin{array}{l} x_1 + x_2 = 1, \\ x_1 \geq 0, x_1 \geq 0 \end{array} \right\}$$



Convex set

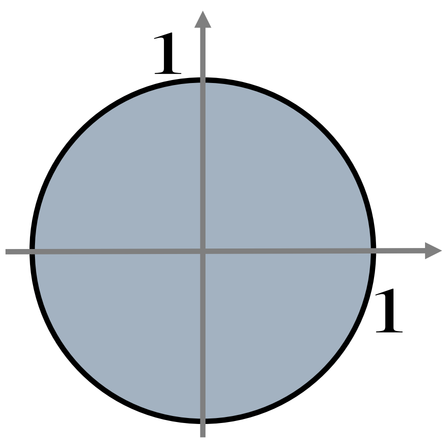
Examples

- Nonlinear convex inequality:

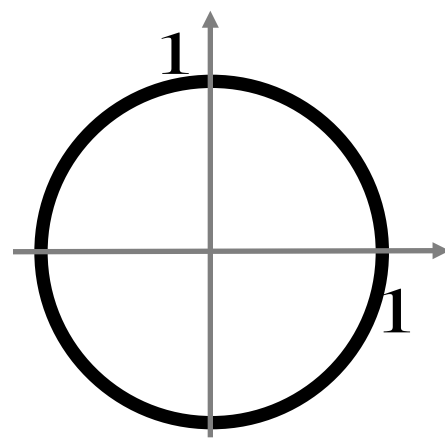
$$S := \left\{ x \in \mathbf{R}^n \,\middle|\, g(x) \le 0 \right\}$$

for some convex function $g(x)$ [see below]

$$\left\{ x \in \mathbf{R}^2 \,\middle|\, x_1^2 + x_2^2 \le 1 \right\}$$



$$\left\{ x \in \mathbf{R}^2 \,\middle|\, x_1^2 + x_2^2 = 1 \right\}$$



nonconvex !

Convex set

Examples: positive semedifinite cones

- Hermitian matrices

$$\mathbf{S}^n := \left\{ A \in \mathbf{C}^{n\times n} \middle| A = A^H \right\}$$

- Positive semidefinite (psd) matrices

$$\mathbf{S}_+^n := \left\{ A \in \mathbf{S}^n \middle| x^H A x \geq 0 \ \text{ for all } x \in \mathbf{C}^n \right\}$$

- Positive definite (pd) matrices

$$\mathbf{S}_{++}^n := \left\{ A \in \mathbf{S}^n \middle| x^H A x > 0 \ \text{ for all } x \in \mathbf{C}^n \right\}$$

Convex set

To recognize a convex set S

- Verify definition

- Show S is obtained from simple convex sets (polyhedra, balls, ellipsoids, cones, …) by operations that preserve convexity
  - intersection
  - affine functions
  - perspective function
  - linear fractional functions

Convex set

Convexity-preserving operations

Suppose $A_k \subseteq \mathbf{R}^n, k = 1,\ldots,K,$ are convex.

Then the following sets are convex:

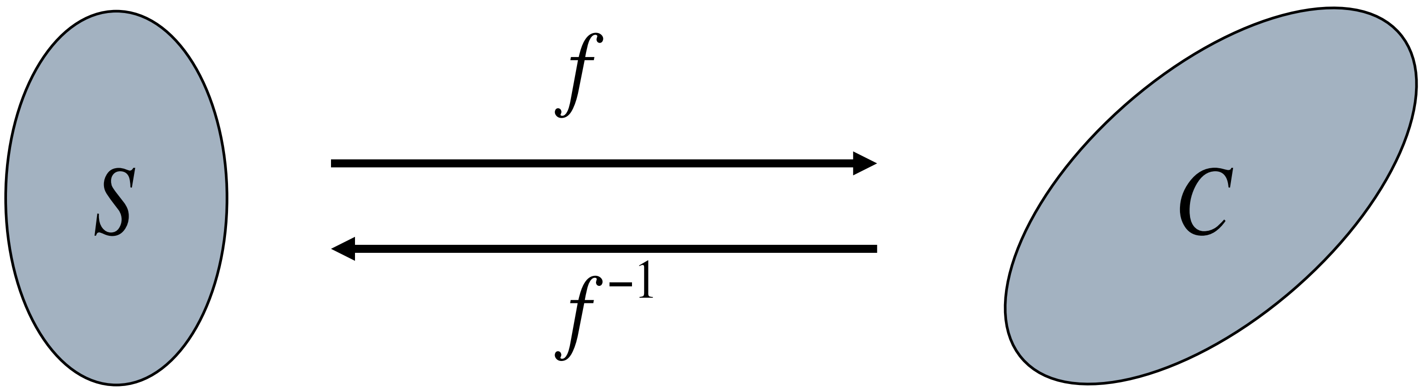
set of sd matrices $\mathbf{S}_+^n := \bigcap_{\substack{z \neq 0 \\ z \in \mathbf{R}^n}} \left\{ X \in \mathbf{S}^n : z^T X z \geq 0 \right\}$

i.e., arbitrary intersection of convex sets is convex

- $B := \bigcap_k A_k$
- $B := A_1 \times \cdots \times A_K$
- $B := \sum_k A_k \; := \; \left\{ \sum_k x_k \middle| x_k \in A_k \right\}$

Convex set

Affine function: $f(x) := Ax + b$



$$S = f^{-1}(C) := \{x \mid Ax + b \in C\} \qquad C = f(S) := \{Ax + b \mid x \in S\}$$

S is convex iff C is convex

Application

- Scaling, translation, projection

Convex set

Affine function: $f:\mathbf{R}^n \to \mathbf{S}_+^m$

- Solution set of LMI

$$S := \left\{ x \middle| x_1 A_1 + \cdots + x_n A_n \le B \right\} \quad (A, B \in \mathbf{S}^m)$$ symmetric matrices

because:

$$f:\mathbf{R}^n \to \mathbf{S}_+^m : f(x) = B - A(x)$$

$$C: \text{ psd cone } \mathbf{S}_+^m \; := \left\{ f(x) \ge 0 \right\}$$

$$S := f^{-1}(C)$$

Convex set

Affine function: $f : \mathbf{R}^n \to \mathbf{R}^{n+1}$

- Hyperbolic cone

$$S := \left\{ x \middle| x^T P x \le (c^T x)^2 \right\} \quad (P \in \mathbf{S}^m, c \in \mathbf{R}^n)$$

because:

$$f : \mathbf{R}^n \to \mathbf{R}^{n+1} : f(x) = \left(P^{1/2} x, c^T x \right)$$

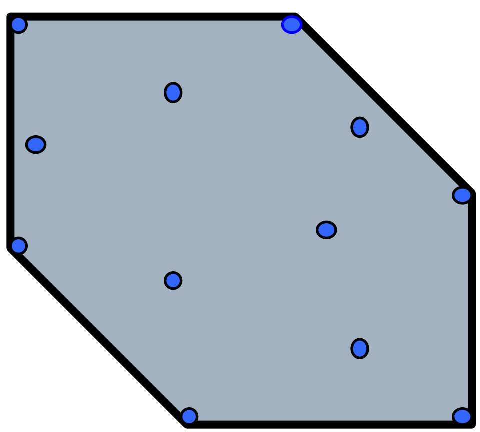
$$C := \left\{ (y,t) \in \mathbf{R}^{n+1} \middle| y^T y \le t^2 \right\} \quad \text{second-order cone}$$

$$S := f^{-1}(C)$$

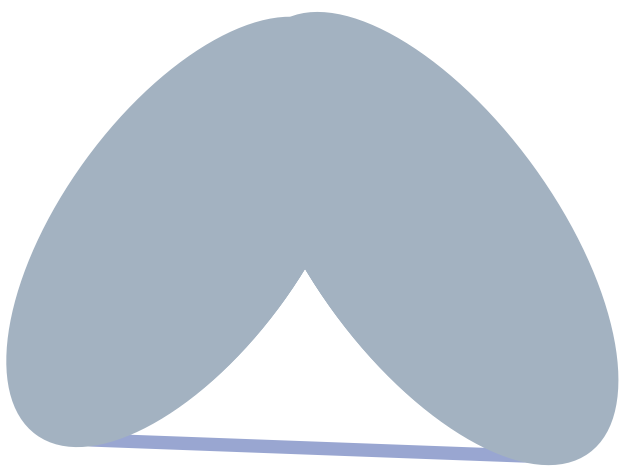
Convex set

Definition

The convex hull $\mathrm{conv}\, S$ of an arbitrary set S is the smallest convex set that contains S



conv S



conv S

Convex set

Examples

- $S := \left\{ A \in \mathbf{C}^{n\times n} \,\middle|\, \text{psd, rank } A \le 1 \right\}$

conv S = $\mathbf{S}_+^n$ (the set of psd matrices)

Convex optimization

- Convex set
- Convex function
- Duality and KKT condition
- Algorithms

References:
Boyd and Vandenberghe, Convex optimization, 2004
Ben-Tal and Nemirovski, Lectures on modern convex optimization, 2013

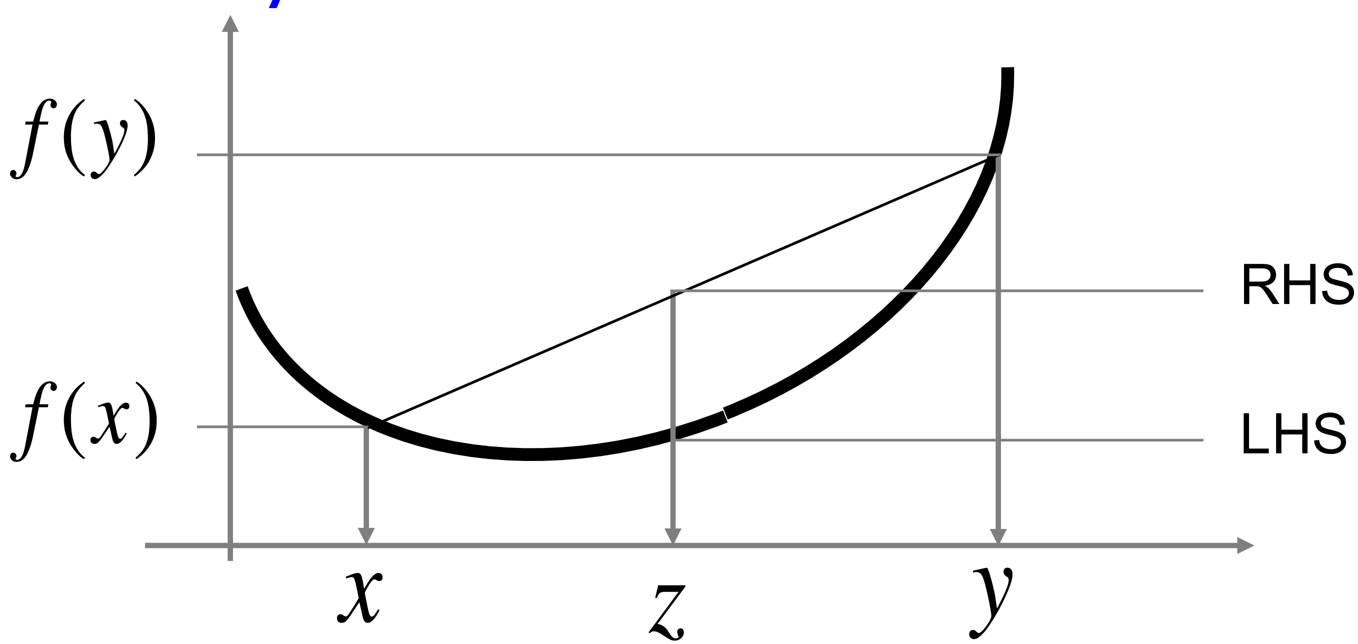
Convex function

Definition

$f(x):\mathbf{R}^n \rightarrow \mathbf{R}$ is convex if dom f is a convex set and for all $x, y \in \mathrm{dom}\, f$, $\alpha \in [0,1]$

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

f is strictly convex if "<"



Convex function

Characterization:

$f(x): \mathbf{R}^n \to \mathbf{R}$ is convex iff for all $x, y \in \mathbf{R}^n$ and $t \in \mathbf{R}$ s.t. $x + ty \in \operatorname{dom} f$

$$g(t) := f(x + ty) \quad \text{is convex}$$

Note: $g(t)$ is a function of a scalar t. It says that start from any point x, go in any direction y, the function g is convex.

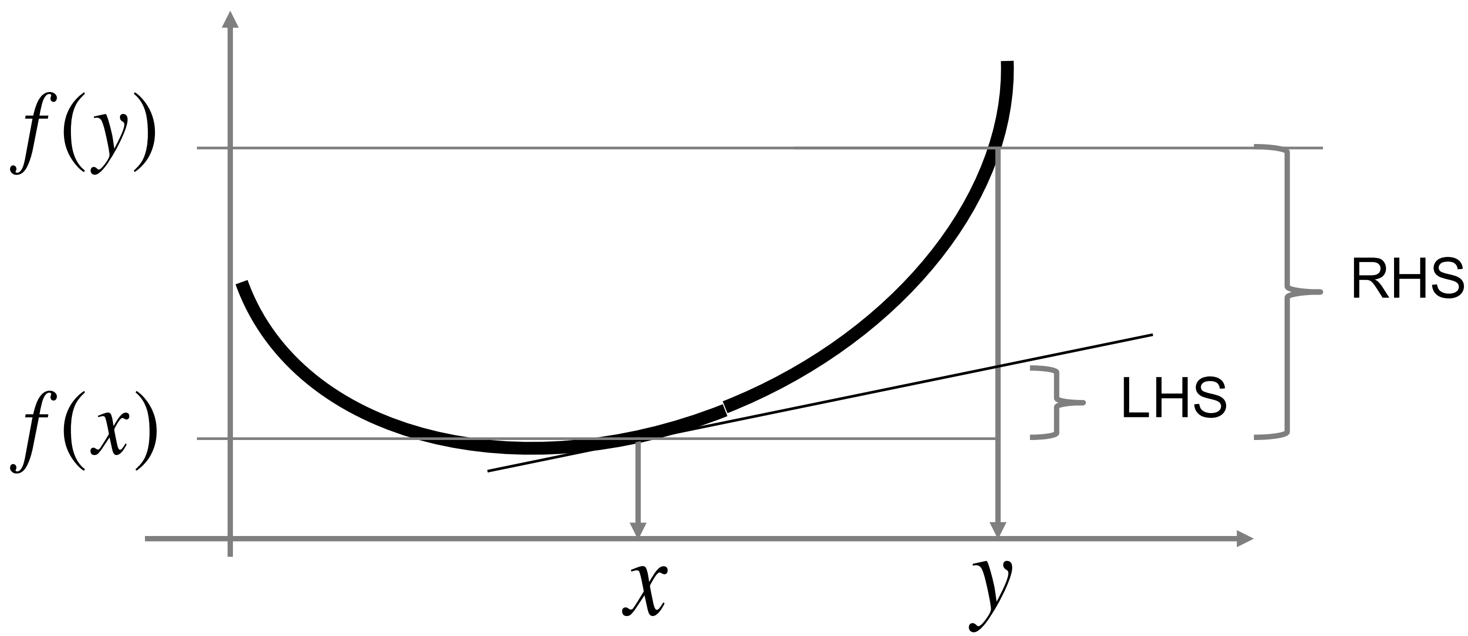
Convex function

Characterization: $\nabla f(x)$ exists, convex dom f

$f(x) : \mathbf{R}^n \to \mathbf{R}$ is convex iff for all $x, y \in \operatorname{dom} f$

$$\left(\nabla f(x)\right)^T (y - x) \le f(y) - f(x)$$

strictly convex iff "<"



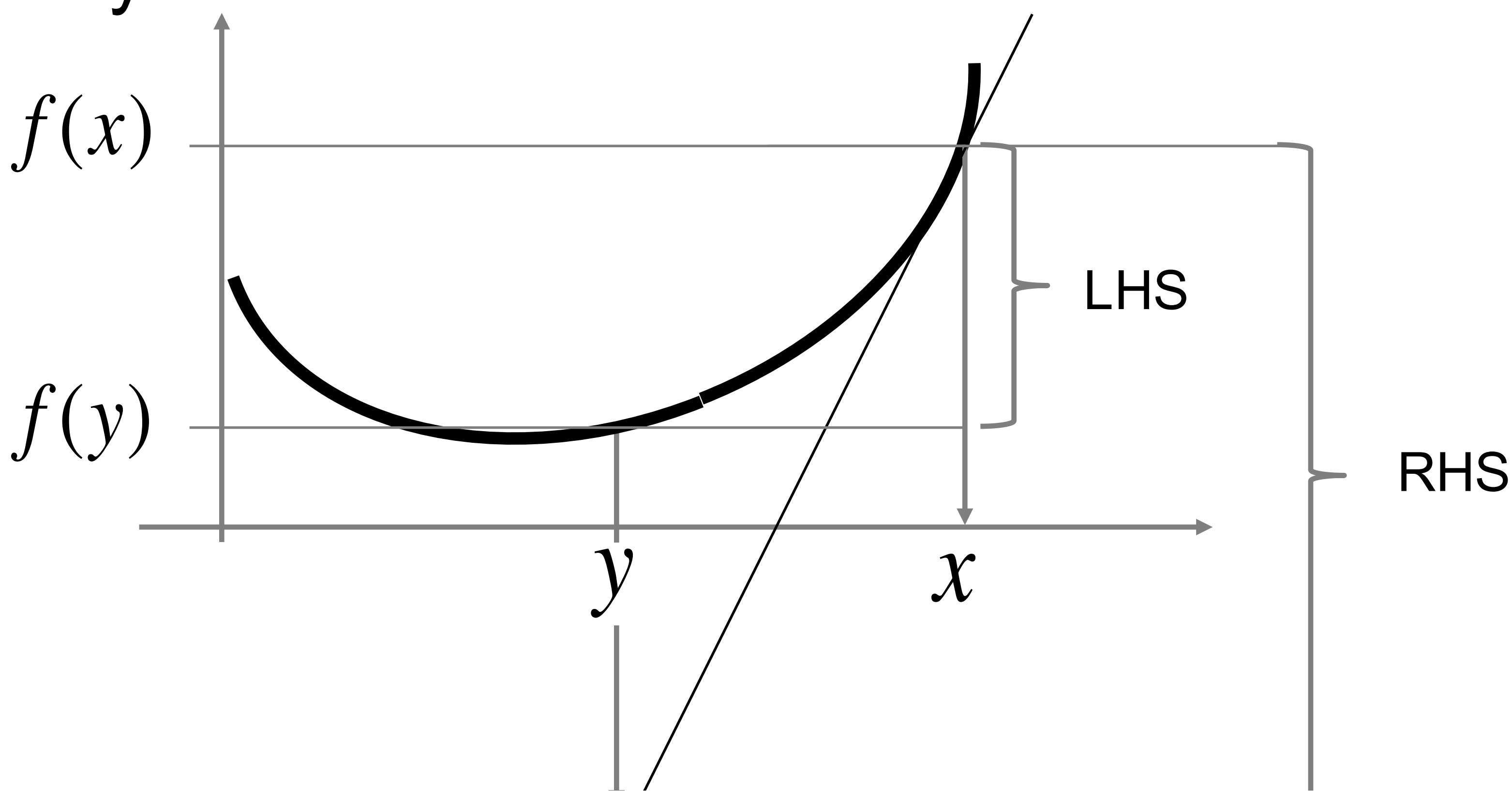
Convex function

Characterization: $\nabla f(x)$ exists, convex dom f

$f(x): \mathbf{R}^n \to \mathbf{R}$ is *convex* iff for all $x, y \in \mathbf{R}^n$

$$\left(\nabla f(x)\right)^T (x-y) \geq f(x) - f(y)$$

strictly convex iff "<"



Convex function

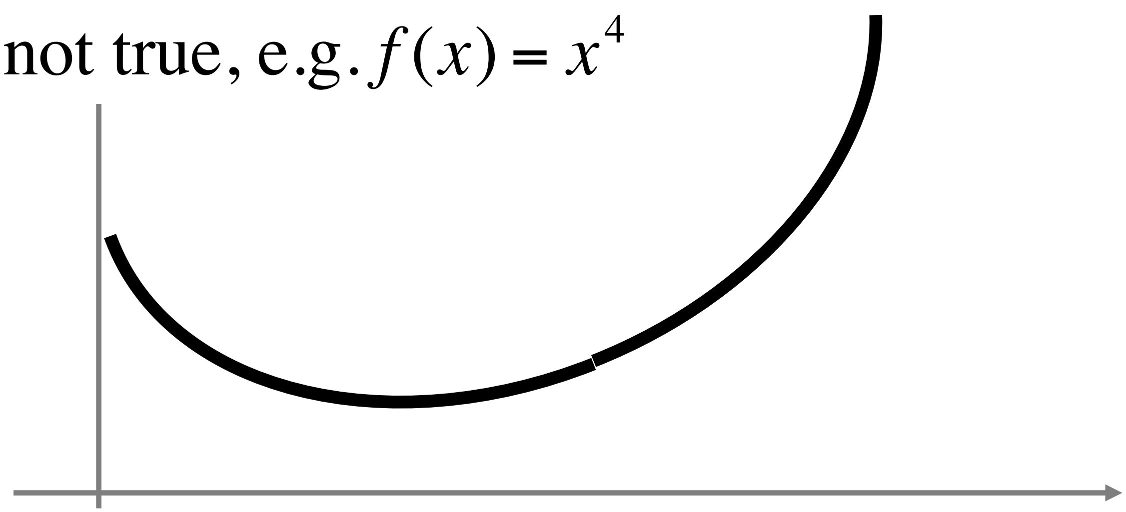
Characterization: $\dfrac{\partial^2 f}{\partial x^2}(x)$ exists, convex dom f

$f(x): D \to \mathbf{R}$ is *convex* on D iff for all $x \in D$

$$\frac{\partial^2 f}{\partial x^2}(x) \geq 0$$ (positive semidefinite)

strictly convex iff ">" (positive definite)

converse not true, e.g. $f(x) = x^4$



Common mistake

Note that $\forall x \in D \quad \frac{\partial^2 f}{\partial x^2}(x) \geq 0$ means

fix any $x \in D \quad \forall y \in \mathbf{R}^n \quad y^T \frac{\partial^2 f}{\partial x^2}(x) y \geq 0$

and is different from

$$\forall x \in D \quad x^T \frac{\partial^2 f}{\partial x^2}(x) x \geq 0$$

Common mistake

Example

$$f(x) = x_1 x_2 \qquad \forall x \in D := \left\{ (x_1, x_2) \middle| x_1 \geq 0, x_2 \geq 0 \right\}$$

Then $\dfrac{\partial^2 f}{\partial x^2}(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$$\text{ev-ev's}: \left(1,\ y_1 = [1\ 1]^T\right), \left(-1,\ y_2 = [1\ -1]^T\right)$$

1. f is not convex on D $\left(\because \dfrac{\partial^2 f}{\partial x^2}(x) \text{ not psd bc } y_2^T \dfrac{\partial^2 f}{\partial x^2} y_2 \leq 0 \right)$

2. $x^T \dfrac{\partial^2 f}{\partial x^2}(x) x = [x_1\ x_2] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1 x_2 \ \geq\ 0 \ \text{ over } D$

Convex function

Examples

- $f(x) = x^2$

Definition:

$$\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y)$$

$$= \left(\alpha x^2 + (1-\alpha) y^2\right) - \left(\alpha^2 x^2 + (1-\alpha)^2 y^2 + 2\alpha(1-\alpha) xy\right)$$

$$= \alpha(1-\alpha)\left(x-y\right)^2 \quad \geq \quad 0$$

Convex function

Examples

- $f(x) = x^2$

Characterization 1:

$$g(t) := f(x+ty) = \left(y \cdot t + x\right)^2$$

which is clearly convex in t

Convex function

Examples

- $f(x) = x^2$

Characterization 2:

$$\left(f(y) - f(x)\right) \; - \; \left(\nabla f(x)\right)^T (y - x)$$

$$= \left(y^2 - x^2\right) - 2x(y - x)$$

$$= (x - y)^2 \;\; \geq \;\; 0$$

Convex function

Examples

- $f(x) = x^2$

Characterization 3:

$$\frac{\partial^2 f}{\partial f^2}(x) = 2 \geq 0$$

Convex function

Examples

- $f(x) = x^2$
- $f(x) = e^x$
- $f(x) = -\log x$
- $f(x) = \dfrac{1}{x}$
- any norm $\|x\|$

Convex function

To recognize a convex function f

- Verify definition
- Apply one of 3 characterizations
- Show f is obtained from simple convex functions (linear, quadratic, exp, -log, …) by operations that preserve convexity
  - nonnegative weighted sum
  - composition with affine function
  - pointwise supremum
  - composition

Convex function

Convexity-preserving operations

Suppose $f_k : \mathbf{R}^n \to \mathbf{R},\ k = 1, \ldots, K$ are convex.
Then the following functions are convex:

- $g(x) := \sum_k a_k f_k(x),\ a_k \geq 0$
- $g(x) := \max_k f_k(x)$ $\qquad g(x) := \sup_{y \in Y} f(x, y)$
- $g(x) := h\left(f_1(x)\right)$ provided $h : \mathbf{R} \to \mathbf{R}$ is convex and h is nondecreasing

$$\tilde{h}(x) := h(x),\ \ x \in \operatorname{dom} h,\ \ \tilde{h}(x) := \infty,\ \ x \notin \operatorname{dom} h$$

Convex function

Composition

- Composition with affine function

$$f(Ax+b) \text{ is convex if } f \text{ is convex}$$

Example:

- Any norm of affine function $\|Ax+b\|$

Convex function

Composition

- Composition with affine function

$$f(Ax+b) \text{ is convex if } f \text{ is convex}$$

Example:

Log barrier for linear inequalities

$$f(x) := -\sum_i \log\left(b_i - a_i^T x\right)$$

$$\operatorname{dom} f := \left\{ x \middle| a_i^T x < b_i \quad \forall i \right\}$$

Convex function

Composition

- Composition with max function

$$\sup_{y \in Y} f(x;y) \text{ is convex if } f(\cdot;y) \text{ is convex } \forall y$$

Example:

- max eigenvalue of matrix $X \in \mathbf{S}^n$

$$\lambda_{\max}(X) = \max_{\|y\|_2=1} y^T X y = \max_{\|y\|_2=1} \underbrace{\mathrm{tr}\left(yy^T\right)X}_{\text{linear (convex) in } X}$$

Convex optimization

- Convex set
- Convex function
- Duality and KKT condition
- Algorithms

References:
Boyd and Vandenberghe, Convex optimization, 2004
Ben-Tal and Nemirovski, Lectures on modern convex optimization, 2013

Convex optimization

$$\begin{aligned} \min_{x \in \mathbf{R}^n} \quad & f_0(x) \\ \text{s. t.} \quad & f_k(x) \le 0, \quad k = 1, \ldots, K \\ & Ax = b \end{aligned}$$

$f_k(x)$: convex functions for $k = 0, 1, \ \ldots, K$

Convex optimization

Advantages

Local optimality implies global optimality

- Sufficient to focus on local optimal

First-order optimality condition is sufficient

- Not only necessary

Polynomial-time computable

- Nonconvex programs are NP-hard in general

Duality theory and Lagrange multipliers

- Important for both structure and computation

Duality theory

$$
\begin{aligned}
\min_{x \in \mathbf{R}^n} \quad & f_0(x) \\
\text{s. t.} \quad & f_k(x) \le 0, \qquad k = 1, \dots, K \\
& Ax = b, \qquad A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m
\end{aligned}
$$

Dual problem

$$
\begin{aligned}
\min_{x \in \mathbf{R}^n} \quad & f_0(x) \\
\text{s. t.} \quad & f_k(x) \le 0, \qquad k = 1, \ldots, K \\
& Ax = b, \qquad A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m
\end{aligned}
$$

Lagrange multipliers

$\lambda \in \mathbf{R}_+^K$

$\mu \in \mathbf{R}^m$

Dual problem

$$\min_{x \in \mathbf{R}^n} \quad f_0(x)$$

$$\text{s. t.} \quad f_k(x) \le 0, \qquad k = 1, \ldots, K$$

$$Ax = b, \qquad A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$$

Lagrange multipliers

$\lambda \in \mathbf{R}_+^K$

$\mu \in \mathbf{R}^m$

Lagrangian

$$L(x; \lambda, \mu) := f_0(x) + \sum_{k \ge 1} \lambda_k f_k(x) + \mu^T (Ax - b)$$

Convert constraints into penalties !

- one Lagrange multiplier per constraint
- inequality constraints → $\lambda \ge 0$
- equality constraints → μ

Dual problem

$$\min_{x \in \mathbf{R}^n} \quad f_0(x)$$

$$\text{s. t.} \quad f_k(x) \le 0, \quad k = 1, \ldots, K$$

$$Ax = b, \quad A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$$

Lagrange multipliers

$$\lambda \in \mathbf{R}_+^K$$

$$\mu \in \mathbf{R}^m$$

Lagrangian

$$L(x; \lambda, \mu) := f_0(x) + \sum_{k \ge 1} \lambda_k f_k(x) + \mu^T (Ax - b)$$

Dual objective function

$$D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$$

unconstrained min

Dual problem

$$\min_{x \in \mathbf{R}^n} \quad f_0(x)$$

$$\text{s. t.} \quad f_k(x) \le 0, \quad k = 1, \ldots, K$$

$$Ax = b, \quad A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$$

Lagrange multipliers

$\lambda \in \mathbf{R}_+^K$

$\mu \in \mathbf{R}^m$

Lagrangian

$$L(x; \lambda, \mu) := f_0(x) + \sum_{k \ge 1} \lambda_k f_k(x) + \mu^T (Ax - b)$$

Dual objective function

$$D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$$

unconstrained min

Dual problem: $\max_{\lambda \ge 0, \mu} D(\lambda, \mu)$

Weak duality

Primal: $\min_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s. t.} \quad f_k(x) \le 0, \quad Ax = b$

Dual: $\max_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

Weak duality

Primal: $\min_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s. t.} \quad f_k(x) \le 0, \quad Ax = b$

Dual: $\max_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

Theorem: weak duality

For any primal feasible x and dual feasible (λ, μ)

$$D(\lambda, \mu) \le f_0(x)$$

In particular

$$D(\lambda^*, \mu^*) \le f_0(x^*)$$

weak duality holds for nonconvex programs

Strong duality

Primal: $\min_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s. t.} \quad f_k(x) \le 0, \quad Ax = b$

Dual: $\max_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

Slater's condition: $\exists x \in \text{relint}\ \mathbf{D} := \bigcap_{k \ge 0} \text{dom}\, f_k$ s.t.

$$f_k(x) < 0 \quad \text{if } f_k \text{ is not affine}$$

Strong duality

Primal: $\min\limits_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s. t.} \quad f_k(x) \le 0, \quad Ax = b$

Dual: $\max\limits_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min\limits_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

Slater's condition: $\exists x \in \text{relint } \mathbf{D} := \bigcap\limits_{k \ge 0} \text{dom } f_k$ s.t.

$$f_k(x) < 0 \quad \text{if } f_k \text{ is not affine}$$

Theorem: strong duality

Suppose Primal is convex and Slater's cond holds

- $D(\lambda^*, \mu^*) = f_0(x^*)$
- If dual is feasible then dual is attained

Optimality condition

Theorem (KKT condition)

Suppose Primal is convex and Slater's cond holds

x^* is optimal if and only if there exist $\left(\lambda^*, \mu^*\right)$ s.t.

- primal feasible: $f_k(x^*) \le 0, \quad Ax^* = b$
- dual feasible: $\lambda^* \ge 0$
- first-order cond: $\nabla f_0(x^*) + \sum_{k\ge 1} \lambda_k^* \nabla f_k(x^*) + \mu^{*T} A = 0$
- complementary slackness: $\lambda_k^* f_k(x^*) = 0$

Convex optimization

- Convex set
- Convex function
- Duality and KKT condition
- Algorithms

References:
Boyd and Vandenberghe, Convex optimization, 2004
Ben-Tal and Nemirovski, Lectures on modern convex optimization, 2013

Algorithms

Primal: $\min_{x\in\mathbf{R}^n} \quad f_0(x) \quad \text{s.t.} \quad f_k(x)\le 0, \quad Ax=b$

Dual: $\max_{\lambda\ge 0,\mu} \quad D(\lambda,\mu) := \min_{x\in\mathbf{R}^n} L(x;\lambda,\mu)$

- Many algorithms compute solutions to KKT condition

$$\nabla f_0(x^*) + \sum_{k\ge 1} \lambda_k^* \nabla f_k(x^*) + \mu^{*T} A = 0$$

- Primal algorithm (if projection to feasible set is easy)
- Dual algorithm (if unconstrained primal is easy)
- Primal-dual algorithm
- Second-order algorithm

Primal algorithm

Primal: $\min_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s.t.} \quad f_k(x) \le 0, \quad Ax = b$

Dual: $\max_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

$$x(t+1) = \text{Proj}\left(x(t) - \gamma_t \nabla f_0(x(t))\right)$$

stepsize

- Steepest descent followed by projection to feasible set
- First-order algorithm

Dual algorithm

Primal: $\min_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s. t.} \quad f_k(x) \le 0, \quad Ax = b$

Dual: $\max_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

$$y(t) := (\lambda(t), \mu(t))$$

$$y(t+1) = \mathrm{Proj}\Big(y(t) + \gamma_t \nabla_y L(\underbrace{x(y(t))}_{\underset{x \in \mathbf{R}^n}{\operatorname{argmin}} L(x; \lambda(t), \mu(t))}; y(t)) \Big)$$

- Lagrangian is concave in y
- First-order algorithm

Primal-dual algorithm

Primal: $\min_{x \in \mathbf{R}^n} \quad f_0(x) \quad \text{s. t.} \quad f_k(x) \le 0, \quad Ax = b$

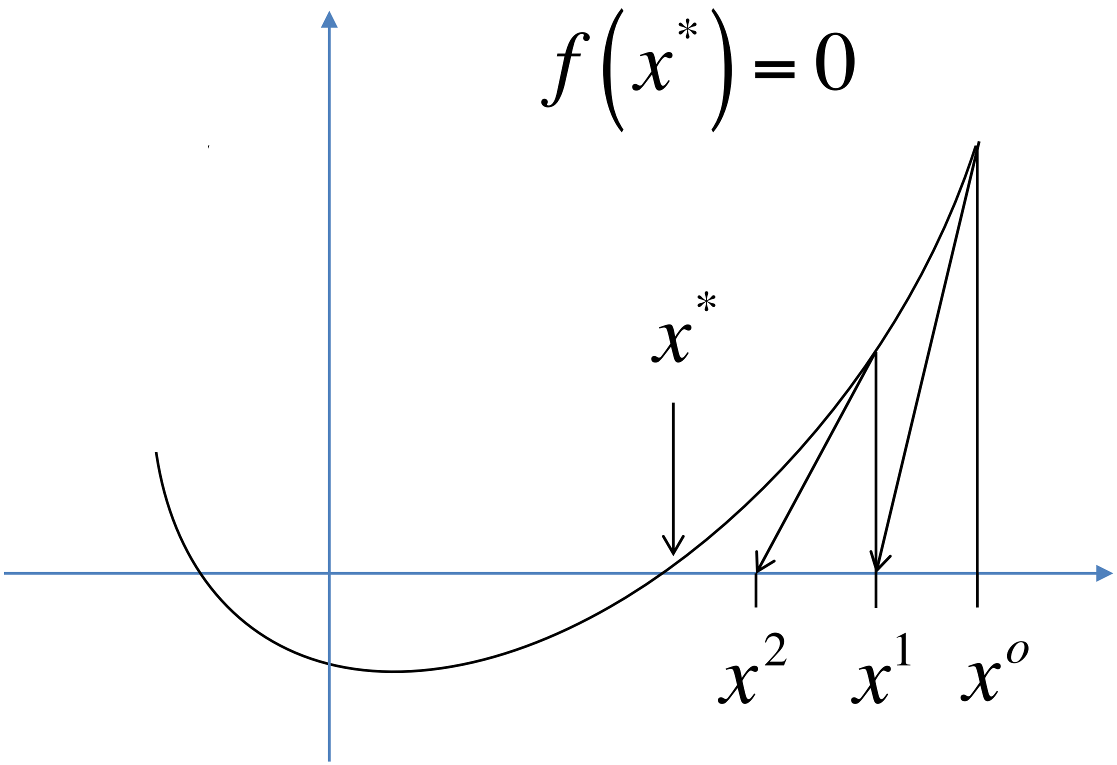
Dual: $\max_{\lambda \ge 0, \mu} \quad D(\lambda, \mu) := \min_{x \in \mathbf{R}^n} L(x; \lambda, \mu)$

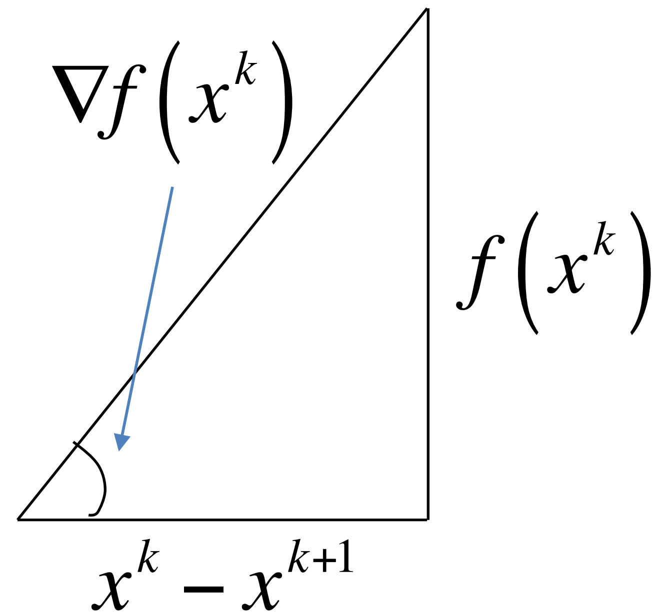
$$x(t+1) = x(t) - \gamma_t \nabla_x L(x(t); y(t))$$

$$y(t+1) = \text{Proj}\Big(y(t) + \gamma_t \nabla_y L(x(t); y(t)) \Big)$$

- Do not have to project to primal feasible set nor compute min x
- Lagrangian is convex in x and concave in y
- First-order algorithm to approach a saddle point

Newton-Raphson method





$$\nabla f\left(x^k\right)\left(x^k - x^{k+1}\right) = f\left(x^k\right)$$

N-R iteration: $$x^{k+1} = x^k - \left[\nabla f\left(x^k\right)\right]^{-1} f\left(x^k\right)$$

More preliminaries

- Semidefinite programs
- QCQP and semidefinite relaxations
- Partial matrices and completions
- Chordal relaxation

Convex optimization

$$\begin{aligned} \min_{x} \quad & f_0(x) \\ \text{s. t.} \quad & f_k(x) \le 0, \quad k = 1, \ldots, K \\ & Ax = b \end{aligned}$$

Definition

Convex optimization if

- $f_k(x)$ are *convex functions* for $k = 0, 1, \ \ldots, K$

The feasible set $\mathbf{X}$ is a *convex set*

Convex optimization is polynomial-time

Convex optimization

$$
\begin{aligned}
\min_x \quad & f_0(x) \\
\text{s.t.} \quad & f_k(x) \le 0, \qquad k = 1, \ldots, K \\
& Ax = b
\end{aligned}
$$

Questions

- How to recognize a convex program
- How to characterize minimizers?
- How to compute a minimizer?

Convex optimization

$$\begin{aligned} \min_{x} \quad & f_0(x) \\ \text{s. t.} \quad & f_k(x) \le 0, \quad k = 1, \dots, K \\ & Ax = b \end{aligned}$$

Questions

- How to recognize a convex program
  - Convex set, convex function
- How to characterize minimizers?
  - KKT condition, duality theorem
- How to compute a minimizer?
  - First-order algorithms, Newton algorithms
  - Distributed algorithms

2^{nd} order cone program (SOCP)

$$\begin{array}{lll} \min & c_0^H x & \\ \text{s.t.} & \left\| C_k x + b_k \right\| \le c_k^H x + d_k & k \ge 1 \end{array}$$

If c_k are complex, how to ensure
C_k^H x is real ?

- $C_k \in \mathbf{C}^{(n_k - 1)\times n}, \;\; b_k \in \mathbf{C}^{n_k - 1}, \;\; c_k \in \mathbf{C}^n, \; d_k \in \mathbf{R}$

- || || : Euclidean norm
- Feasible set is 2^{nd} order cone and convex
- Includes LP, convex QP as special cases
- Special case of SDP, but much simpler computationally

2nd order cone program (SOCP)

$$\begin{array}{lll} \min & c_0^H x & \\ \text{s.t.} & \|C_k x + b_k\| \leq c_k^H x + d_k & k \geq 1 \end{array}$$

- $C_k \in \mathbf{R}^{(n_k-1)\times n},\ \ b_k \in \mathbf{R}^{n_k-1},\ \ c_k \in \mathbf{C}^n,\ d_k \in \mathbf{R}$

- || || : Euclidean norm
- Feasible set is 2nd order cone and convex
- Includes LP, convex QP as special cases
- Special case of SDP, but much simpler computationally

SOCP in rotated form

$$\begin{array}{ll} \min & c_0^H x \\ \text{s.t.} & \left\| C_k x + b_k \right\|^2 \;\le\; \left(c_k^H x + d_k \right)\left(\hat{c}_k^H x + \hat{d}_k \right) \end{array}$$

- Useful for OPF:

$$\begin{array}{llll} \min & c_0^H x & & \\ \text{s.t.} & C_k x \;=\; b_k & & k \ge 1 \\ & \left\| w_m \right\|^2 \;\le\; y_m z_m & & m \ge 1 \end{array}$$

- Transformation:

$$\left\| w \right\|^2 \le yz,\; y \ge 0,\; z \ge 0 \quad \Leftrightarrow \quad \left\| \begin{bmatrix} 2w \\ y - z \end{bmatrix} \right\| \le y + z$$

Semidefinite program (SDP)

Primal: $$\min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s.t.} \quad A_0 + \sum_{i=1}^{n} x_i A_i \leq 0$$

Linear SDP

- What is the Lagrangian?
- What is the dual problem?
- What is KKT condition?

… let's first review the case of linear program

Linear program

Primal: $$\min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s. t.} \quad a_0 + \sum_{i=1}^{n} x_i a_i \leq 0$$

Linear program

Primal: $$\min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s.t.} \quad a_0 + \sum_{i=1}^{n} x_i a_i \leq 0$$

Lagrangian:

$$L(x;\lambda) := \sum_{i=1}^{n} c_i x_i + \lambda \left(a_0 + \sum_{i=1}^{n} x_i a_i \right), \quad \lambda \geq 0$$

$$= a_0 \lambda + \sum_{i=1}^{n} x_i \left(a_i \lambda + c_i \right)$$

Linear program

Primal: $$\min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s. t.} \quad a_0 + \sum_{i=1}^{n} x_i a_i \le 0$$

Lagrangian:

$$L(x;\lambda) \;:=\; a_0\lambda + \sum_{i=1}^{n} x_i\left(a_i\lambda + c_i\right), \quad \lambda \ge 0$$

Dual:

$$D(\lambda) := \min_{x \in \mathbf{R}^n} L(x;\lambda) \;=\; a_0\lambda + \min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} x_i\left(a_i\lambda + c_i\right)$$

$$= \begin{cases} a_0\lambda & \text{if} \;\; a_i\lambda + c_i = 0 \;\; \text{for all } i \\ -\infty & \text{else} \end{cases}$$

Linear program

Primal:

$$\boxed{\min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s. t.} \quad a_0 + \sum_{i=1}^{n} x_i a_i \leq 0}$$

Lagrangian:

$$L(x;\lambda) \;:=\; a_0\lambda + \sum_{i=1}^{n} x_i \left(a_i\lambda + c_i\right), \quad \lambda \geq 0$$

Dual:

$$D(\lambda) = = \begin{cases} a_0\lambda & \text{if} \quad a_i\lambda + c_i = 0 \quad \text{for all } i \\ -\infty & \text{else} \end{cases}$$

$$\boxed{\max_{\lambda \geq 0} \; a_0\lambda \quad \text{s.t.} \quad a_i\lambda + c_i = 0 \quad \text{for all } i}$$

Linear program

Lagrangian: $$L(x;\lambda) \; := \; a_0\lambda + \sum_{i=1}^{n} \boxed{x_i\left(a_i\lambda + c_i\right)}$$

Primal: $$\min_{x\in\mathbf{R}^n} \; \max_{\lambda\geq 0} \; L(x;\lambda)$$

Dual: $$\max_{\lambda\geq 0} \min_{x\in\mathbf{R}^n} \; L(x;\lambda)$$

Complementary Slackness:

$$x_i\left(a_i\lambda + c_i\right) = 0 \quad \forall i$$

Semidefinite program (SDP)

Primal: $$\min_{x\in\mathbf{R}^n} \sum_{i=1}^n c_i x_i \quad \text{s.t.} \quad A_0 + \sum_{i=1}^n x_i A_i \le 0$$

Lagrangian: for $\Lambda \ge 0$

$$L(x;\Lambda) := \sum_{i=1}^n c_i x_i + \operatorname{tr} \Lambda \left(A_0 + \sum_{i=1}^n x_i A_i \right)$$

$$= \operatorname{tr}\left(A_0 \Lambda\right) + \sum_{i=1}^n \left(\operatorname{tr}\,(A_i \Lambda) + c_i\right) x_i$$

$$D(\Lambda) = \begin{cases} \operatorname{tr}\left(A_0\Lambda\right) & \text{if} \quad \operatorname{tr}\,(A_i\Lambda) + c_i = 0 \quad \forall i \\ -\infty & \text{else} \end{cases}$$

Semidefinite program (SDP)

Primal: $$\min_{x\in\mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s. t.} \quad A_0 + \sum_{i=1}^{n} x_i A_i \le 0$$

Dual: $$\max_{\Lambda\ge 0} \operatorname{tr}\left(A_0\Lambda\right) \quad \text{s.t.} \quad \operatorname{tr}\left(A_i\Lambda\right) + c_i = 0 \quad \forall i$$

We will later use an inequality form:

$$\max_{\Lambda\ge 0} \operatorname{tr}\left(A_0\Lambda\right)$$

$$\text{s.t.} \quad \operatorname{tr}\left(A_i\Lambda\right) \le c_i \quad \forall i$$

equivalent to equality form through slack variables

PSD cones are convex

Examples: positive semedifinite cones

- Hermitian matrices

$$\mathbf{S}^n := \left\{ A \in \mathbf{C}^{n\times n} \,\middle|\, A = A^H \right\}$$

- Positive semidefinite (psd) matrices

$$\mathbf{S}^n_+ := \left\{ A \in \mathbf{S}^n \,\middle|\, x^H A x \geq 0 \ \text{ for all } x \in \mathbf{C}^n \right\}$$

- Positive definite (pd) matrices

$$\mathbf{S}^n_{++} := \left\{ A \in \mathbf{S}^n \,\middle|\, x^H A x > 0 \ \text{ for all } x \in \mathbf{C}^n \right\}$$

Semidefinite program (SDP)

Primal: $$\min_{x \in \mathbf{R}^n} \sum_{i=1}^{n} c_i x_i \quad \text{s. t.} \quad A_0 + \sum_{i=1}^{n} x_i A_i \leq 0$$

Dual: $$\max_{\Lambda \geq 0} \operatorname{tr}\left(A_0 \Lambda\right) \quad \text{s.t.} \quad \operatorname{tr}\left(A_i \Lambda\right) + c_i = 0 \quad \forall i$$

Theorem: strong duality

primal optimal value = dual optimal value

Semidefinite program (SDP)

Theorem: The following are equivalent

- (x^*, Λ^*) is primal-dual optimal
- (x^*, Λ^*) is a saddle pt of Lagrangian

$$L(x^*, \Lambda) \le L(x^*, \Lambda^*) \le L(x, \Lambda^*) \quad \forall \text{feasible } x, \Lambda$$

- KKT: $$A_0 + \sum_{i=1}^{n} x_i^* A_i \le 0,$$

$$\Lambda^* \ge 0, \quad \mathrm{tr}\,(A_i \Lambda^*) + c_i = 0 \quad \forall i$$

$$\mathrm{tr}\,\Lambda^* \left(A_0 + \sum_{i=1}^{n} x_i^* A_i \right) = 0$$

Mathematical preliminaries

- Semidefinite programs
- QCQP and semidefinite relaxations
- Partial matrices and completions
- Chordal relaxation

QCQP

$$\begin{array}{lll} \text{min} & x^H C_0 x & \\ \text{over} & x \in \mathbf{C}^n & \\ \text{s.t.} & x^H C_k x \;\le\; b_k & k \ge 1 \end{array}$$

- $C_k, k \ge 0$, Hermitian $\Rightarrow$ $x^H C_k x$ is real
  $b_k \in \mathbf{R}^n$
- Convex problem if all C_k are psd
  Nonconvex otherwise

QCQP

$$\begin{array}{lll} \text{min} & x^H C_0 x & \\ \text{over} & x \in \mathbf{C}^n & \\ \text{s.t.} & x^H C_k x \leq b_k & k \geq 1 \end{array}$$

- $x^H C_k x = \operatorname{tr} x^H C_k x = \operatorname{tr} C_k \left(x x^H \right)$

QCQP

$$\begin{array}{lll} \min & \operatorname{tr} C_0\left(xx^H\right) & \\ \text{over} & x \in \mathbf{C}^n & \\ \text{s.t.} & \operatorname{tr} C_k\left(xx^H\right) \le b_k & k \ge 1 \end{array}$$

- $x^H C_k x = \operatorname{tr} x^H C_k x = \operatorname{tr} C_k\left(xx^H\right)$

QCQP

$$
\begin{array}{lll}
\min & \operatorname{tr} C_0\left(xx^H\right) & \\
\text{over} & x \in \mathbf{C}^n & \\
\text{s.t.} & \operatorname{tr} C_k \underbrace{\left(xx^H\right)}_{X \in \mathbf{S}_+^n} \le b_k & k \ge 1
\end{array}
$$

- $x^H C_k x = \operatorname{tr} x^H C_k x = \operatorname{tr} C_k\left(xx^H\right)$

QCQP

$$\begin{array}{lll} \min & \operatorname{tr} C_0 X & \\ \text{over} & X \in \mathbf{S}_+^n & \\ \text{s.t.} & \operatorname{tr} C_k X \le b_k & k \ge 1 \\ & \cancel{\operatorname{rank} X = 1} & \end{array}$$

$\longleftarrow$ only nonconvexity

- Any solution X yields a unique x through

$$X = xx^{\mathrm{H}}$$

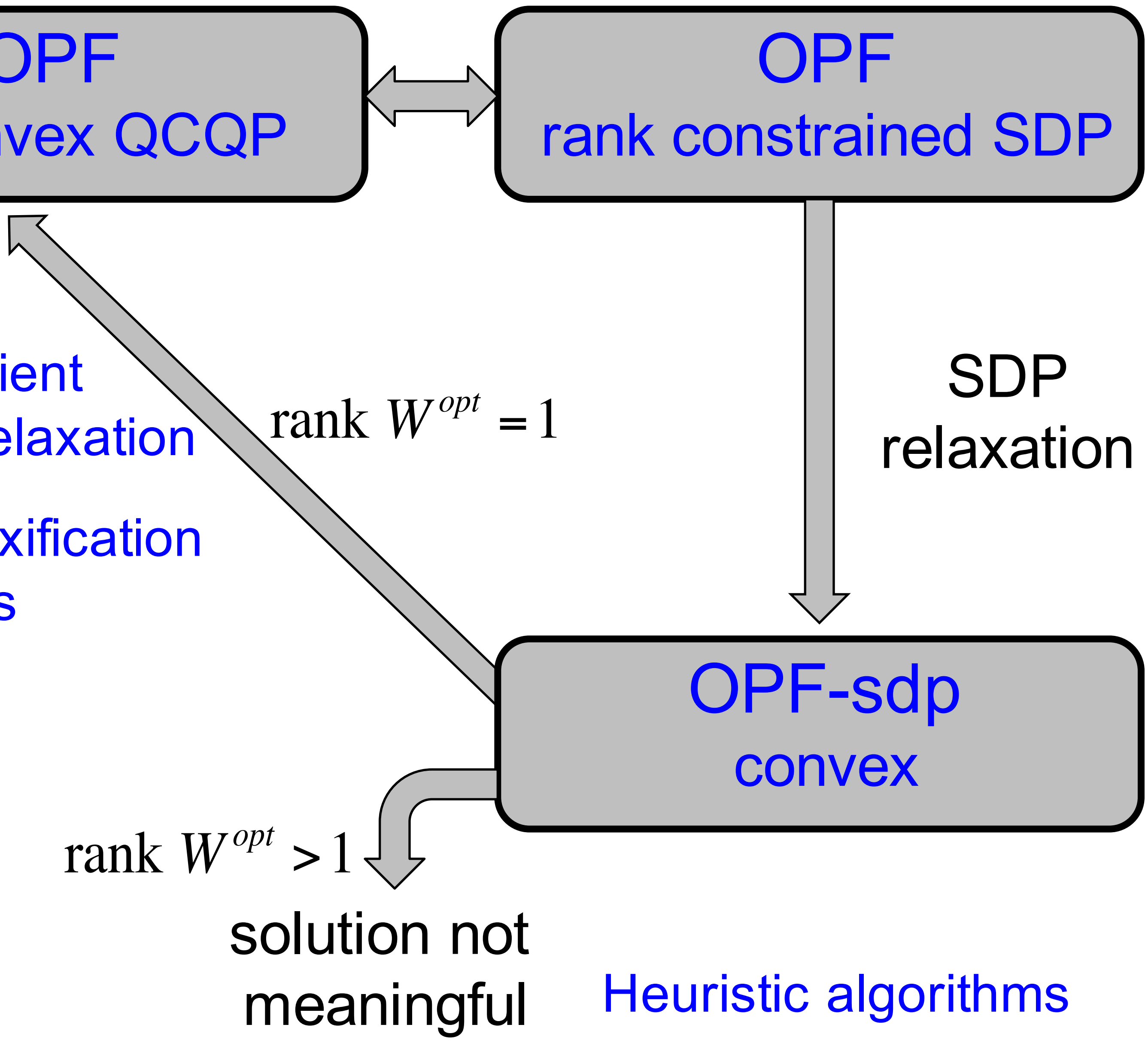
- Feasible sets are *equivalent*

Semidefinite program (SDP)

$$\begin{array}{lll} \min & \operatorname{tr} C_0 X & \\ \text{s.t.} & \operatorname{tr} C_k X \;\le\; b_k & k \ge 1 \\ & X \;\ge\; 0 & \end{array}$$

- Feasible set of QCQP is an *effective subset* of feasible set of SDP
- SDP is a *relaxation* of QCQP

Preview: solution strategy



Radial network: sufficient conditions for exact relaxation

Mesh network: convexification through phase shifters

Heuristic algorithms

SOCP in rotated form

$$\begin{array}{ll} \min & c_0^H x \\ \text{s.t.} & \left\| C_k x + b_k \right\|^2 \;\le\; \left(c_k^H x + d_k \right)\left(\hat{c}_k^H x + \hat{d}_k \right) \end{array}$$

- Useful for OPF:

$$\begin{array}{llll} \min & c_0^H x & & \\ \text{s.t.} & C_k x \;=\; b_k & & k \ge 1 \\ & \left\| w_m \right\|^2 \;\le\; y_m z_m & & m \ge 1 \end{array}$$

- Transformation:

$$\left\| w \right\|^2 \le yz, \; y \ge 0, \; z \ge 0 \quad \Leftrightarrow \quad \left\| \begin{bmatrix} 2w \\ y - z \end{bmatrix} \right\| \le y + z$$

Recap: QCQP, SDP, SOCP

QCQP

$$\begin{array}{lll} \min & x^H C_0 x & \\ \text{s.t.} & x^H C_k x \le b_k & k \ge 1 \end{array}$$

SDP

$$\begin{array}{lll} \min & \operatorname{tr} C_0 X & \\ \text{s.t.} & \operatorname{tr} C_k X \le b_k & k \ge 1 \\ & X \ge 0 & \end{array}$$

SOCP

$$\begin{array}{lll} \min & c_0^H x & \\ \text{s.t.} & C_k x = b_k & k \ge 1 \\ & \|w_m\|^2 \le y_m z_m & m \ge 1 \end{array}$$

Mathematical preliminaries

- Semidefinite programs
- QCQP and semidefinite relaxations
- Partial matrices and completions
- Chordal relaxation

Graphs

Graph $G = (V, E)$

Complete graph: all node pairs adjacent

Clique: complete subgraph of G

- An edge is a clique
- *Maximal* clique: a clique that is not a subgraph of another clique

Chordal graph: all minimal cycles have length 3

- *Minimal* cycle: cycle without chord

Chordal ext: chordal graph containing G

- Every graph has a chordal extension
- Chordal extensions are not unique

Partial matrices

Fix an undirected graph $G = (V, E)$

Partial matrix X_G :

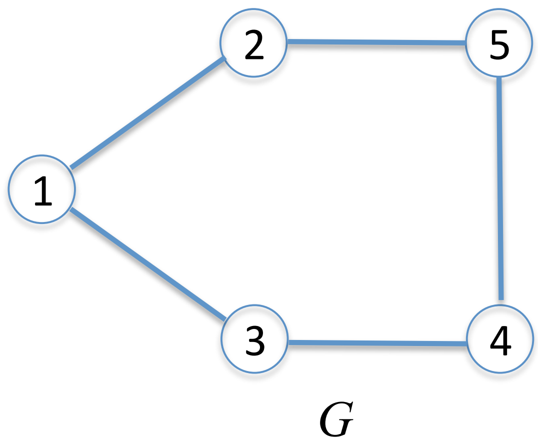
$$X_G := \left(\left[X_G \right]_{jj}, j \in V,\ \left[X_G \right]_{jk}, (j,k) \in E \right)$$

Completion X of a partial matrix X_G :

$$X = X_G \text{ on } G$$

Example

partial matrix X_G := { complex numbers on G }



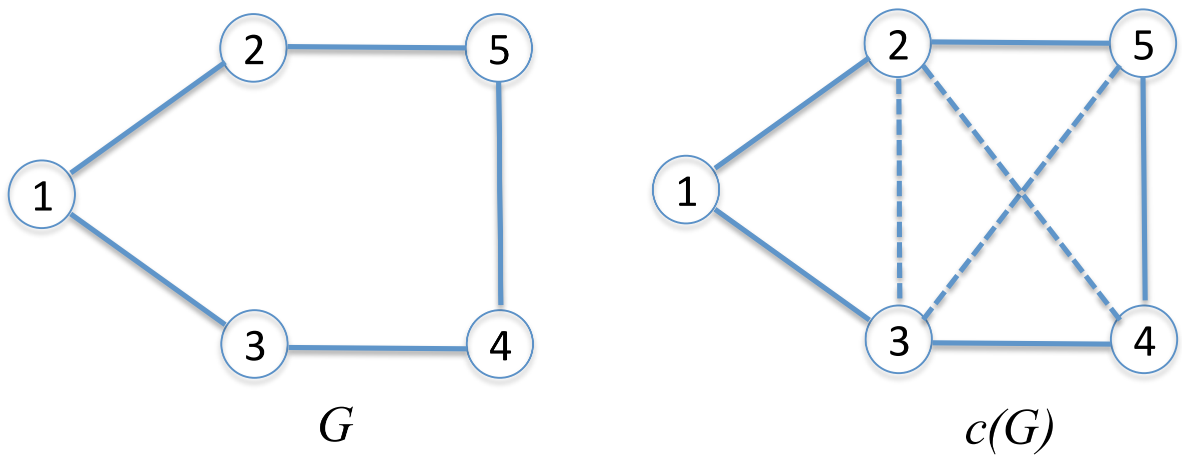
n-vertex complete graph

$$X_G = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & & & x_{25} \\ x_{31} & & x_{33} & x_{34} & \\ & & x_{43} & x_{44} & x_{45} \\ & x_{52} & & x_{54} & x_{55} \end{bmatrix}$$

completion: full matrix X that agrees with X_G on G

Example

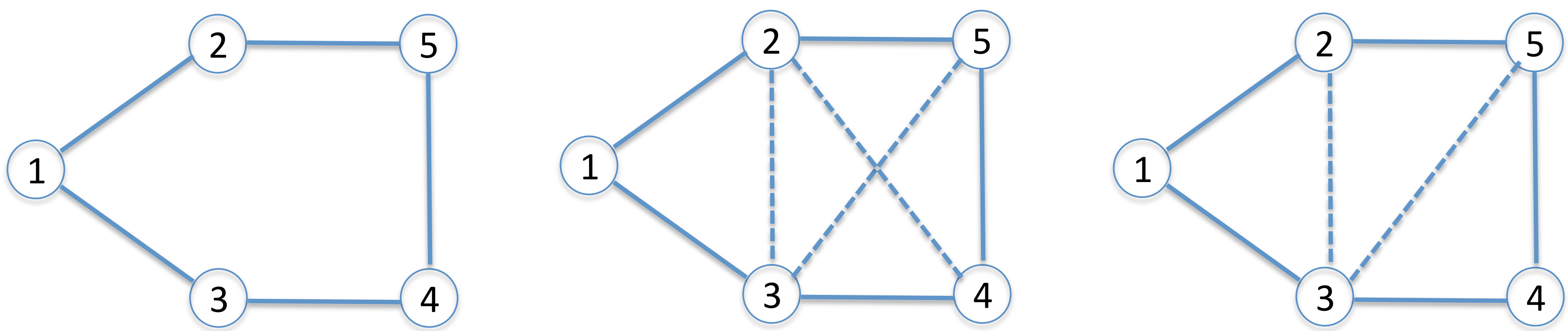
chordal ext $X_{c(G)} :=$ { complex numbers on $c(G)$ }



$$X_G = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & & & x_{25} \\ x_{31} & & x_{33} & x_{34} & \\ & & x_{43} & x_{44} & x_{45} \\ & x_{52} & & x_{54} & x_{55} \end{bmatrix} \quad X_{c(G)} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ & x_{42} & x_{43} & x_{44} & x_{45} \\ & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix}$$

Example

chordal ext $X_{c(G)} :=$ { complex numbers on $c(G)$ }



G $\qquad$ $c(G)$ $\qquad$ $c(G)$

$$X_G = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & & & x_{25} \\ x_{31} & & x_{33} & x_{34} & \\ & & x_{43} & x_{44} & x_{45} \\ & x_{52} & & x_{54} & x_{55} \end{bmatrix} \quad X_{c(G)} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ & x_{42} & x_{43} & x_{44} & x_{45} \\ & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix} \quad X_{c(G)} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & x_{23} & & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ & & x_{43} & x_{44} & x_{45} \\ & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix}$$

Partial matrices

Fix an undirected graph $G = (V, E)$

A partial matrix X_G is *psd* if

$$X_G(q) \geq 0 \quad \text{for all maximal cliques } q$$

A partial matrix X_G is *rank-1* if

$$\operatorname{rank} X_G(q) = 1 \quad \text{for all maximal cliques } q$$

Matrix completion

<u>Theorem</u> [Grone et al 1984]

Every psd partial matrix X_G has a psd completion if and only if G is chordal

- Motivates chordal relaxation

Chordal relaxation

QCQP

$$\begin{aligned} \min \quad & x^H C_0 x \\ \text{s.t.} \quad & x^H C_k x \le b_k \qquad k \ge 1 \end{aligned}$$

SDP

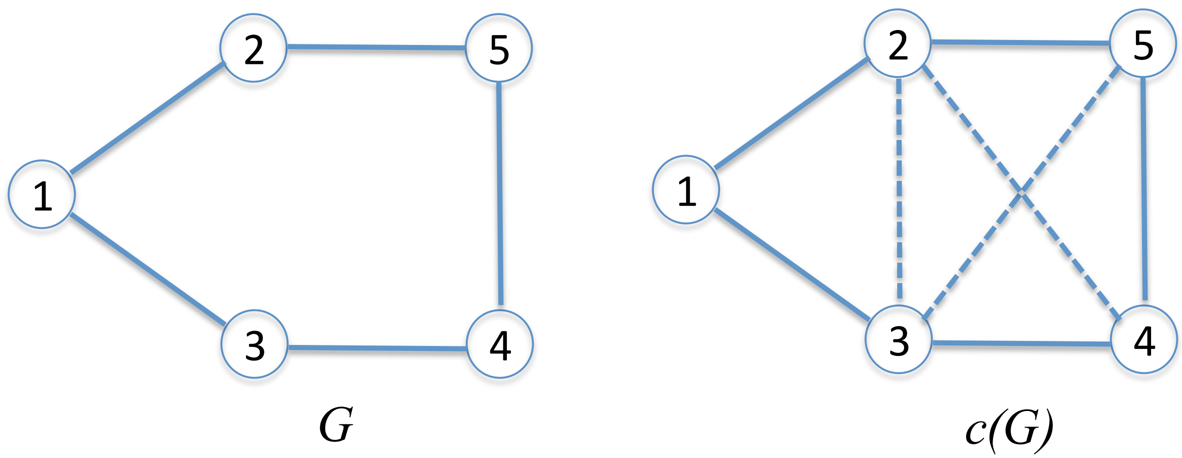
$$\begin{aligned} \min \quad & \text{tr } C_0 X \\ \text{s.t.} \quad & \text{tr } C_k X \le b_k \qquad k \ge 1 \\ & X \ge 0 \end{aligned}$$

Chordal

$$\begin{aligned} \min_{X_{c(G)}} \quad & \text{tr } C_0 X_G \\ \text{s.t.} \quad & \text{tr } C_k X_G \le b_k \qquad k \ge 1 \\ & X_{c(G)} \ge 0 \end{aligned}$$

Example

chordal ext $X_{c(G)} :=$ { complex numbers on $c(G)$ }



$$X_G = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & & & x_{25} \\ x_{31} & & x_{33} & x_{34} & \\ & & x_{43} & x_{44} & x_{45} \\ & x_{52} & & x_{54} & x_{55} \end{bmatrix} \qquad X_{c(G)} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & & \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ & x_{42} & x_{43} & x_{44} & x_{45} \\ & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix}$$

Chordal relaxation

$$\min_{X_{c(G)}} \quad \operatorname{tr} C_0 X_G$$

$$\text{s.t.} \quad \operatorname{tr} C_k X_G \;\leq\; b_k \qquad k \geq 1$$

$$X(q_1) \;\geq\; 0, \quad X(q_2) \;\geq\; 0$$

$$X(q_1) = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & \boxed{x_{22}} & \boxed{x_{23}} \\ x_{31} & \boxed{x_{32}} & \boxed{x_{33}} \end{bmatrix} \qquad X(q_2) = \begin{bmatrix} \boxed{x_{22}} & \boxed{x_{23}} & x_{24} & x_{25} \\ \boxed{x_{32}} & \boxed{x_{33}} & x_{34} & x_{35} \\ x_{42} & x_{43} & x_{44} & x_{45} \\ x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix}$$

Chordal relaxation

$$\min_{X_{c(G)}} \quad \operatorname{tr} C_0 X_G$$

$$\text{s.t.} \quad \operatorname{tr} C_k X_G \leq b_k \qquad k \geq 1$$

$$X'(q_1) \succeq 0, \quad X(q_2) \succeq 0$$

$$\boxed{u_{jk} = x_{jk}, \quad j,k = 2,3}$$

$$X'(q_1) = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & u_{22} & u_{23} \\ x_{31} & u_{32} & u_{33} \end{bmatrix} \qquad X(q_2) = \begin{bmatrix} x_{22} & x_{23} & x_{24} & x_{25} \\ x_{32} & x_{33} & x_{34} & x_{35} \\ x_{42} & x_{43} & x_{44} & x_{45} \\ x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix}$$

Chordal relaxation

$$\min_{X_{c(G)}} \quad \operatorname{tr} C_0' X'$$

$$\text{s.t.} \quad \operatorname{tr} C_k' X' \le b_k \qquad k \ge 1$$

$$X' \ge 0$$

$$\boxed{\operatorname{tr} C_r' X' = 0 \qquad r = 1,2,3,4}$$

$$X' = \begin{bmatrix} X'(q_1) & 0 \\ 0 & X(q_2) \end{bmatrix}$$

- This is SDP in standard form
- Size of X' and #equality constraints depend on $c(G)$

Chordal relaxation

$$\min_{X_{c(G)}} \quad \operatorname{tr} C_0' X'$$

$$\text{s.t.} \quad \operatorname{tr} C_k' X' \le b_k \qquad k \ge 1$$

$$X' \ge 0$$

$$\boxed{\operatorname{tr} C_r' X' = 0 \qquad r = 1,2,3,4}$$

$$X' = \begin{bmatrix} X'(q_1) & 0 \\ 0 & X(q_2) \end{bmatrix}$$

- Simpler than SDP for sparse graph G
- Equivalent to SDP in worst case