

Lecture 1 - Problem formulation

Dinesh Krishnamoorthy

Department of Engineering Cybernetics, NTNU

Spring 2026

Sources

Presentation based on material from:

- ▶ Nocedal & Wright (2006): Numerical Optimization
- ▶ Boyd & Vandenberghe (2004): Convex Optimization

Mathematical preliminaries

Given a scalar function $f(\mathbf{x})$ with $\mathbf{x} \in \mathbb{R}^n$, i.e., $f : \mathbb{R}^n \rightarrow \mathbb{R}$, its gradient is given by

$$\frac{\partial f}{\partial \mathbf{x}} = \nabla f(\mathbf{x})^\top = \left[\frac{\partial f}{\partial x_1} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right]$$

and its Hessian is given by

$$\frac{\partial^2 f}{\partial \mathbf{x}^2} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Vector functions

A vector function $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a vector of m functions

$$\mathbf{g} = [g_1 \quad \cdots \quad g_m]^\top$$

Given a vector function $\mathbf{g}(\mathbf{x})$ with $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, the $m \times n$ Jacobian matrix is given by

$$\frac{\partial \mathbf{g}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

Mathematical Optimization

Optimization is the minimization/maximization of a function subject to constraints on its variables.

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \leq 0 \end{aligned}$$

- ▶ $x \in \mathbb{R}^n$ - vector of decision variables
- ▶ $f : \mathbb{R}^n \rightarrow \mathbb{R}$ - objective function
- ▶ $h : \mathbb{R}^n \rightarrow \mathbb{R}^{m_h}$ - set of equality constraints
- ▶ $g : \mathbb{R}^n \rightarrow \mathbb{R}^{m_g}$ - set of inequality constraints

Equivalent Formulations

$$\begin{array}{ll}\min_x & f(x) \\ \text{s.t.} & h(x) = 0 \\ & g(x) \leq 0\end{array}$$

Minimization or maximization?

► $\min_x f(x) \Leftrightarrow \max_x -f(x)$

Different formulations based on inequality constraints

- negative null form ($g(x) \leq 0$)
- positive null form ($g(x) \geq 0$)
- negative unity form ($g(x) \leq 1$)
- positive unity form ($g(x) \geq 1$)

In this course, we will use minimization with negative null form for the inequality constraints.

Unconstrained Vs. Unconstrained Optimization

Constrained Optimization

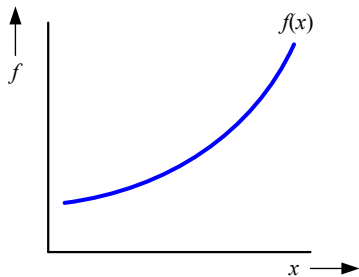
$$\begin{aligned} \min_x & f(x) \\ \text{s.t. } & h(x) = 0 \\ & g(x) \leq 0 \end{aligned}$$

Unconstrained Optimization

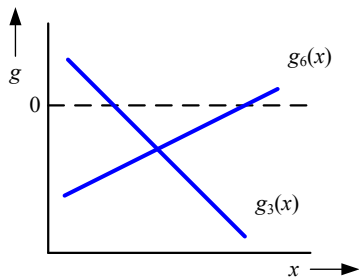
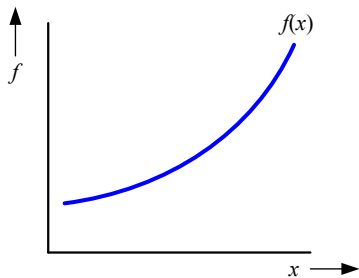
$$\min_{x \in \mathbb{R}^n} f(x)$$

- ▶ We have different tailored algorithms to solve constrained and unconstrained optimization problems.
- ▶ Sometime we can also transform constrained problems into unconstrained problems (penalty methods).

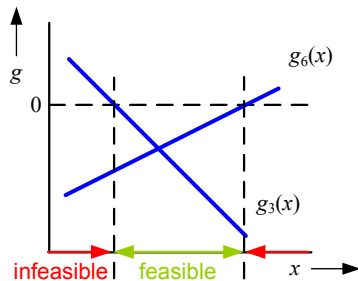
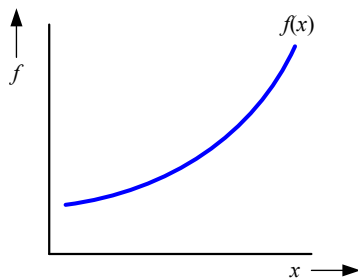
Example: 1D



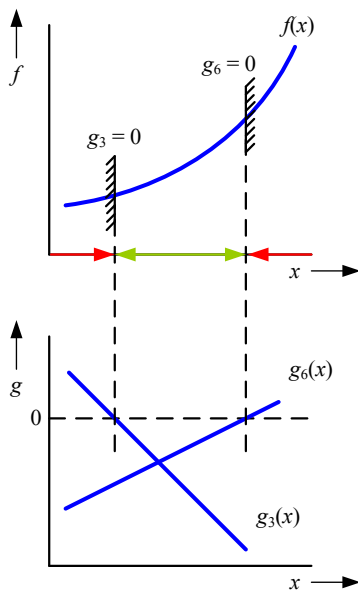
Example: 1D



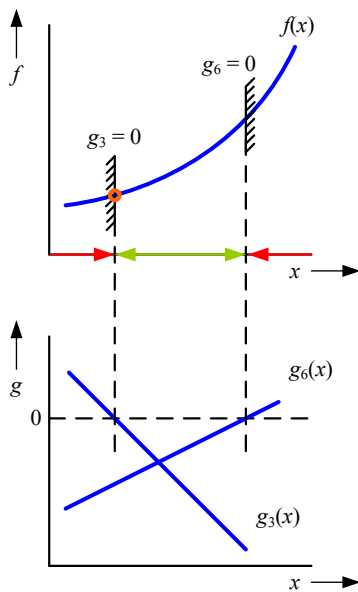
Example: 1D



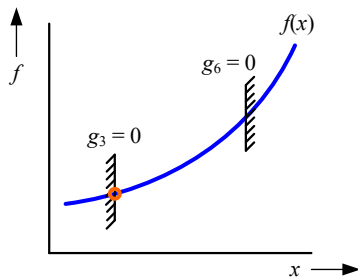
Example: 1D



Example: 1D

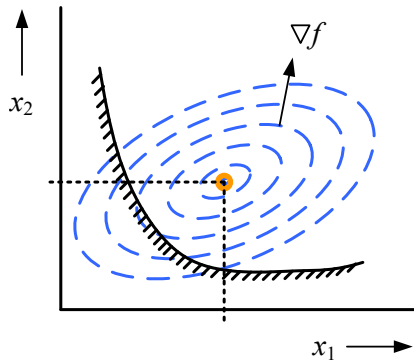
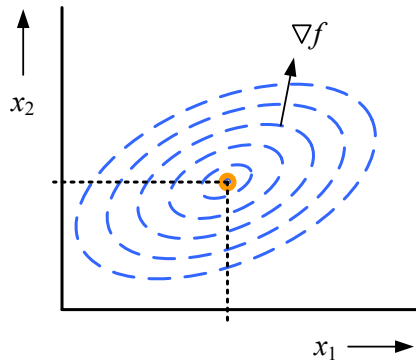


Example: 1D



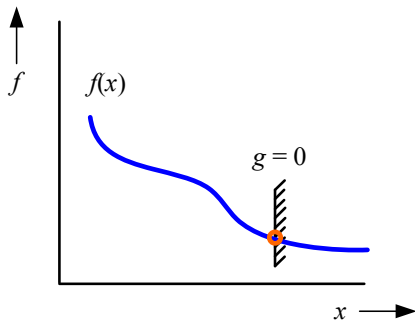
Example 2D

2D examples

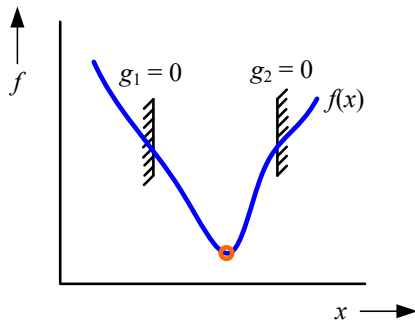


unconstrained cases

Constrained versus unconstrained optimum



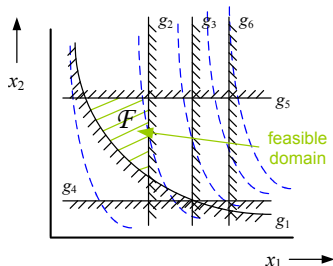
- ▶ constrained optimum
- ▶ bounded optimum



- ▶ unconstrained optimum
- ▶ interior optimum

Feasible domain

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0, \\ & g(x) \leq 0 \end{aligned}$$



Feasible domain - The set of all points that satisfy all the constraints

$$\mathcal{F} := \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}$$

Equivalently

$$\mathcal{F} := \{x \in \mathbb{R}^n \mid g(x) \leq 0\} \cap \{x \in \mathbb{R}^n \mid h(x) = 0\}$$

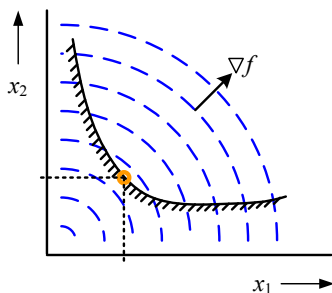
Active constraints and active set

A constraint $g_j(x) \leq 0$ being active at the optimum implies that

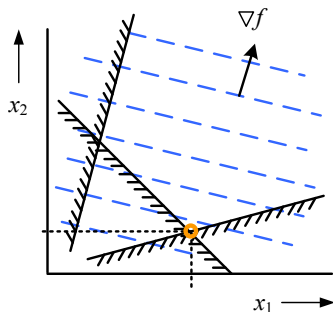
1. the constraint is satisfied with strict equality: $g_j(x) = 0$
2. if strongly active, if $g_j(x) \leq 0$ is left out, the location of the optimum changes

Overview well-constrained problems

2-D examples



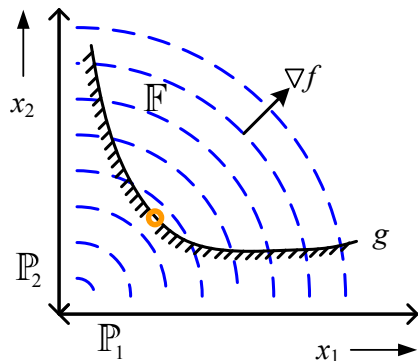
The number of active constraints is less than the number of variables.



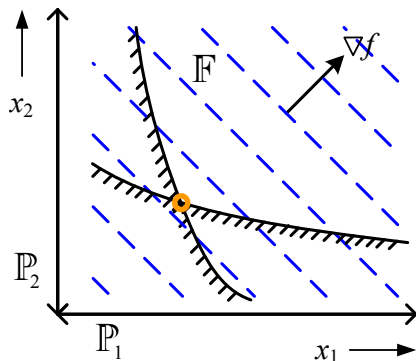
The number of active constraints equals the number of variables.

constrained cases (boundary optima)

Constrained optimum - 2D examples

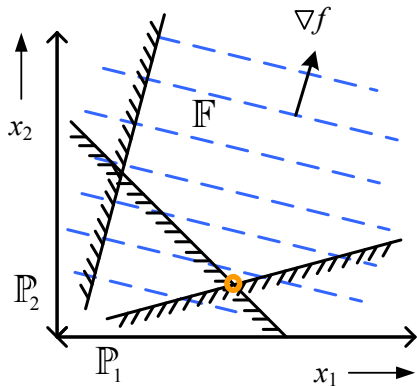


well-constrained over $\mathbb{P}_1 \times \mathbb{P}_2$

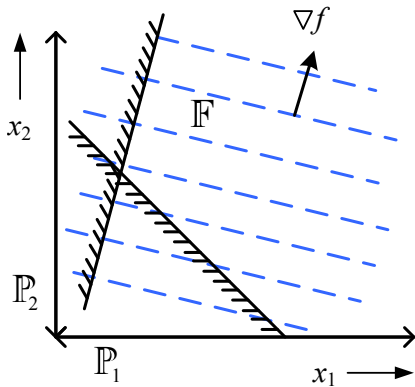


well-constrained over $\mathbb{P}_1 \times \mathbb{P}_2$

Constrained optimum - 2D examples

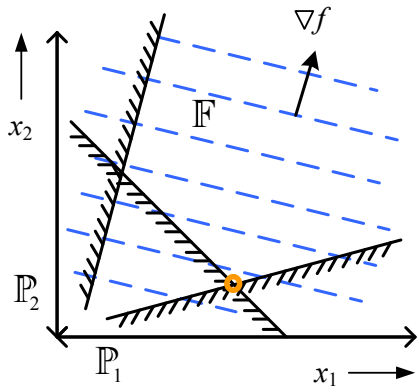


well-constrained over $\mathbb{P}_1 \times \mathbb{P}_2$

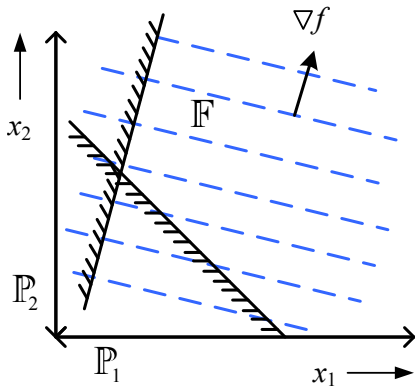


Is it well-constrained?

Constrained optimum - 2D examples

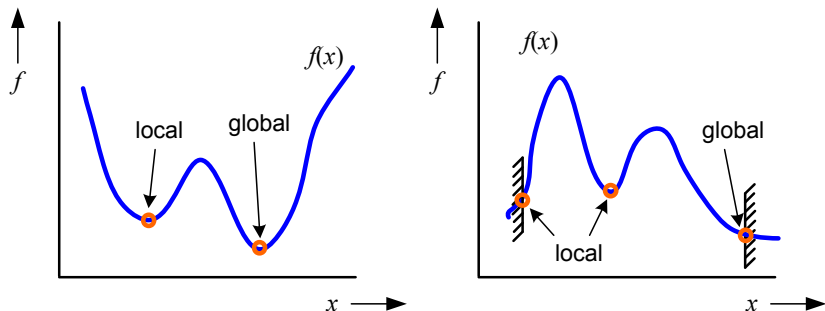


well-constrained over $P_1 \times P_2$



Is it well-constrained?
not well-constrained (a.k.a.
degenerate)

Multiple local minima: multi-modality



- For many problems finding global solution is not easy. Often we seek only a local solution.
- For some classes of problems (e.g., Linear programs and convex problems) local solutions are also global solutions.

NB! In this course we will focus only on local solutions (if multiple solutions exist).

Local Vs. Global optimum

Local Optimum

A point x^* is a local optimum if it yields the lowest cost of its feasible neighborhood.

$$f(x^*) \leq f(x), \forall x \in N_\delta(x^*) \cap \mathcal{F}$$

where the neighborhood $N_\delta(x^*) := \{x : \|x - x^*\| < \delta\}$

Global Optimum

$x^* \in \mathcal{F}$ is a global optimum with the objective function lower than any other point in the feasible set.

$$f(x^*) \leq f(x), \forall x \in \mathcal{F}$$

For convex problems, any solution is a **global** optimum.

Problem properties: smoothness

Function $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ is

- differentiable if all $\frac{\partial f}{\partial x_i}$ exist
- continuously differentiable if all $\frac{\partial f}{\partial x_i}$ continuous
- twice differentiable if all $\frac{\partial^2 f}{\partial x_i \partial x_j}$ exist
- twice continuously differentiable if all $\frac{\partial^2 f}{\partial x_i \partial x_j}$ continuous

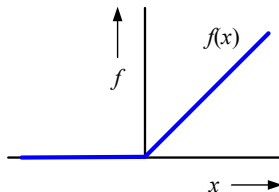
Smoothness assumption $\equiv$ twice continuously differentiable

Non-smoothness Reformulation - Max

- ▶ Discontinuities? → Try to reformulate
- ▶ Example: max-operator in objective

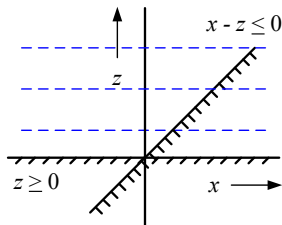
Non-smooth problem:

$$\min_{x \in \mathbb{R}} \max(0, x)$$



Reformulated:

$$\begin{aligned} \min_{(x,z) \in \mathbb{R}^2} \quad & z \\ \text{s.t.} \quad & x - z \leq 0 \\ & z \geq 0 \end{aligned}$$

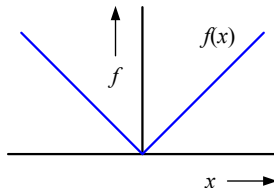


Non-smoothness reformulation - abs

- ▶ Discontinuities? → Try to reformulate
- ▶ Example: absolute value operator in objective

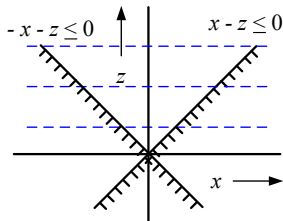
Non-smooth problem:

$$\min_{x \in \mathbb{R}} |x|$$



Reformulated:

$$\begin{array}{ll} \min_{(x,z) \in \mathbb{R}^2} & z \\ \text{s.t.} & x - z \leq 0 \\ & -x - z \leq 0 \end{array}$$

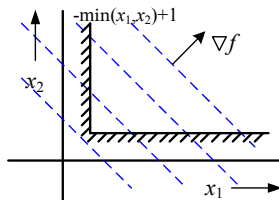


Non-smoothness reformulation - min

- Discontinuities? → Try to reformulate
- Example: min operator in constraint

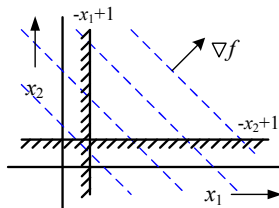
Non-smooth problem:

$$\begin{array}{ll}\min_{(x_1, x_2) \in \mathbb{R}^2} & x_1 + x_2 \\ \text{s.t.} & -\min(x_1, x_2) + 1 \leq 0\end{array}$$



Reformulated:

$$\begin{array}{ll}\min_{(x_1, x_2) \in \mathbb{R}^2} & x_1 + x_2 \\ \text{s.t.} & -x_1 + 1 \leq 0 \\ & -x_2 + 1 \leq 0\end{array}$$



Continuous versus discrete optimization

Continuous optimization problems

- ▶ All decision variables are allowed to be real numbers $x \in \mathbb{R}^n$
- ▶ i.e., feasible set is uncountable finite

Discrete optimization problems / (mixed) integer programming problem

- ▶ some decision variables must be integers, e.g. $x_i \in \mathbb{Z}$, $x_i \in \{0, 1\}$

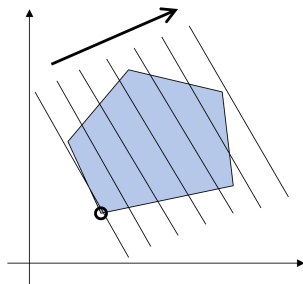
Linear programming (LP)

An optimization problem is an LP problem iff $f(x)$, $g(x)$, $h(x)$ are affine

$$\min_x f(x) := c^\top x$$

$$\text{s.t. } h(x) := A_{eq}^\top x - b_{eq} = 0,$$

$$g(x) := A^\top x - b \leq 0$$



Example: LP

- ▶ A farmer wants to grow apples (A) and bananas (B)
- ▶ He has a field of size 100000m^2
- ▶ Growing 1 tonne of A requires an area of 4000m^2 , growing 1 tonne of B requires an area of 3000m^2
- ▶ A requires 60 kg fertilizer per tonne, B requires 80 kg fertilizer per tonne
- ▶ The profit for A is 7000 per tonne (including fertilizer cost), the profit for B is 6000 per tonne (including fertilizer cost)
- ▶ The farmer can legally use up to 2000 kg of fertilizer
- ▶ The farmer wants to maximize his profits

Example: LP

Variables:

- ▶ x_1 tonnes of A
- ▶ x_2 tonnes of B

Objective: $\max 7000x_1 + 6000x_2$

Constraints:

- ▶ Land limit
 $4000x_1 + 3000x_2 \leq 100000$
- ▶ Fertilizer limit
 $60x_1 + 80x_2 \leq 2000$

$$\min_{x_1, x_2} -7000x_1 - 6000x_2$$

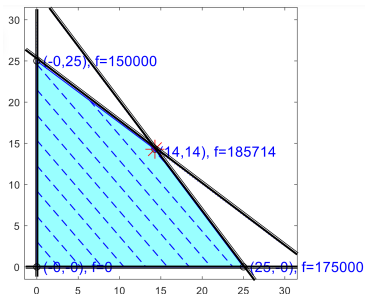
$$\text{s.t. } 4000x_1 + 3000x_2 - 100000 \leq 0$$

$$60x_1 + 80x_2 - 2000 \leq 0$$

$$-x_1 \leq 0$$

$$-x_2 \leq 0$$

Example: LP



$$\min_{x_1, x_2} -7000x_1 - 6000x_2$$

$$\text{s.t. } 4000x_1 + 3000x_2 - 100000 \leq 0$$

$$60x_1 + 80x_2 - 2000 \leq 0$$

$$-x_1 \leq 0$$

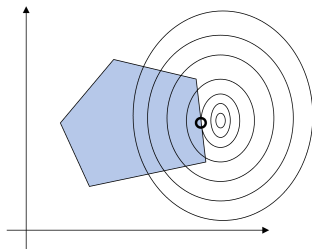
$$-x_2 \leq 0$$

Quadratic programming (QP)

An optimization problem is an QP problem if $f(x)$ is quadratic and $g(x), h(x)$ are affine

$$\begin{aligned} \min_x f(x) &:= \frac{1}{2}x^\top Qx + c^\top x \\ \text{s.t. } h(x) &:= A_{eq}^\top x - b_{eq} = 0, \\ g(x) &:= A^\top x - b \leq 0 \end{aligned}$$

This QP is **convex** if $Q \succeq 0$

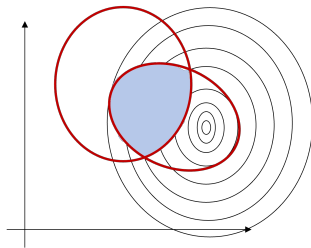


Quadratically constrained quadratic programming (QCQP)

In QCQP, quadratic (or affine) expressions are allowed in both objective and constraints.

$$\begin{aligned} \min_x \quad & f(x) := \frac{1}{2}x^\top Q_0 x + c_0^\top x \\ \text{s.t.} \quad & h_i(x) := \frac{1}{2}x^\top Q_i x + c_i^\top x + d_i = 0, \\ & g_j(x) := \frac{1}{2}x^\top Q_j x + c_j^\top x + d_j \leq 0 \end{aligned}$$

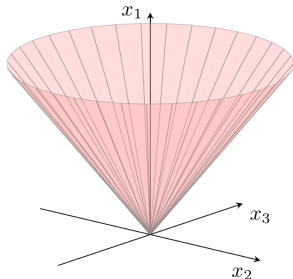
Convex QCQP requires $Q \succeq 0$ and $Q_j \succeq 0$



Second order cone program (SOCP)

SOCP is characterized by affine cost $f(x)$ and equality constraint $h(x)$, and a 2-norm cone inequality constraint $g(x)$.

$$\begin{aligned} \min_x \quad & f(x) := c^\top x \\ \text{s.t.} \quad & h(x) := A_{eq}^\top x - b_{eq} = 0, \\ & g(x) := \|A^\top x - b\|_2 - (c_j^\top x + d_j) \leq 0 \end{aligned}$$



SOCP is **nonsmooth** because the Euclidean norm is not differentiable at the origin.

Semi-definite program (SDP)

SDP is characterized by affine cost $f(x)$ and equality constraint $h(x)$, and Linear Matrix Inequality (LMI) constraint $G(x)$.

$$\begin{aligned} \min_x \quad & f(x) := c^\top x \\ \text{s.t.} \quad & h(x) := A_{eq}^\top x - b_{eq} = 0, \\ & G(x) := G_0 + x_1 G_1 + \cdots + x_n G_n \succeq 0 \quad (\text{LMI}) \end{aligned}$$

with symmetric matrices G_i

- ▶ SDP is **nonsmooth** because positive semi-definite constraints involve eigenvalues, which are not differentiable at eigenvalue multiplicities.
- ▶ An SDP can also have matrix decision variables.

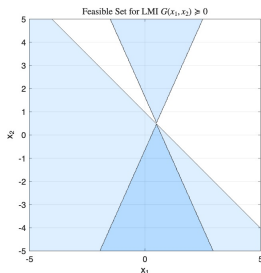
Example: Linear Matrix inequality (LMI)

$$G(x) := G_0 + x_1 G_1 + x_2 G_2$$

with

$$G_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, G_1 = \begin{bmatrix} -1 & -1 \\ -1 & 4 \end{bmatrix}, G_2 = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix}$$

$$G(x) \succeq 0 \Rightarrow G_{11} \geq 0 \text{ and } \det(G) \geq 0$$

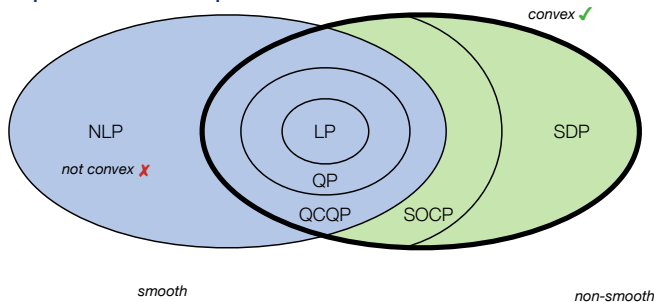


Nonlinear Program (NLP)

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \\ & g(x) \leq 0 \end{aligned}$$

The problem is a nonlinear program if *at least one* of the functions $f(x)$, $h(x)$, or $g(x)$ is nonlinear in x .

Types of optimization problems



- ▶ If all functions are affine $\Rightarrow$ LP.
- ▶ If quadratic or conic structure appears $\Rightarrow$ QP / QCQP / SOCP / SDP.
- ▶ If **no special structure** (e.g., polynomials, exponentials, trigonometric terms) $\Rightarrow$ **NLP**.

"The great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity"

-R. Tyrell Rockafellar

Importance of convexity

Convex optimization problems

- ▶ have only globally optimal solutions (unique if strongly convex)
- ▶ can be solved in polynomial time (fast)
- ▶ can be formulated in a solver-friendly convex form (LP, QP, ...)