ACM 113: Lecture 1

Agenda

- What is a mathematical optimization problem?
- Linear and nonlinear programming
- First examples
- Least squares
- Convex optimization
- Nonlinear optimization
- What is this course about?

What is a mathematical optimization problem?

A *mathematical optimization problem* or *optimization problem* is of the form

$$
\begin{array}{lll}
(P) & \text{minimize} & f_0(x) \\
& \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m
\end{array}
$$

- $x \in \mathbb{R}^n$ is the optimization variable
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function
- functions f_i are the constraints

What is a mathematical optimization problem?

A *mathematical optimization problem* or *optimization problem* is of the form

$$\begin{array}{lll} (P) & \text{minimize} & f_0(x) \\ & \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m \end{array}$$

- $x \in \mathbb{R}^n$ is the optimization variable
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function
- functions f_i are the constraints

A vector $x^\star$ is solution to (P) if for all x such that

$$f_i(x) \leq b_i, \quad \forall i$$

then

$$f_0(x^\star) \leq f_0(x)$$

Linear and nonlinear programming

- Linear programming: The objective and the constraints are linear

  A function f is linear if

$$f(x + \lambda y) = f(x) + \lambda f(y)$$

  for all $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

Linear and nonlinear programming

- Linear programming: The objective and the constraints are linear

  A function f is linear if

  $$f(x + \lambda y) = f(x) + \lambda f(y)$$

  for all $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

- Nonlinear programming: The objective or one of the constraints is not linear

Linear and nonlinear programming

- Linear programming: The objective and the constraints are linear

  A function f is linear if

  $$f(x + \lambda y) = f(x) + \lambda f(y)$$

  for all $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

- Nonlinear programming: The objective or one of the constraints is not linear

- Convex programming: The objective functional and the constraint functionals are convex

  A function f is convex if

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

  for all $x, y \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$

Linear and nonlinear programming

- Linear programming: The objective and the constraints are linear

  A function f is linear if

  $$f(x + \lambda y) = f(x) + \lambda f(y)$$

  for all $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

- Nonlinear programming: The objective or one of the constraints is not linear

- Convex programming: The objective functional and the constraint functionals are convex

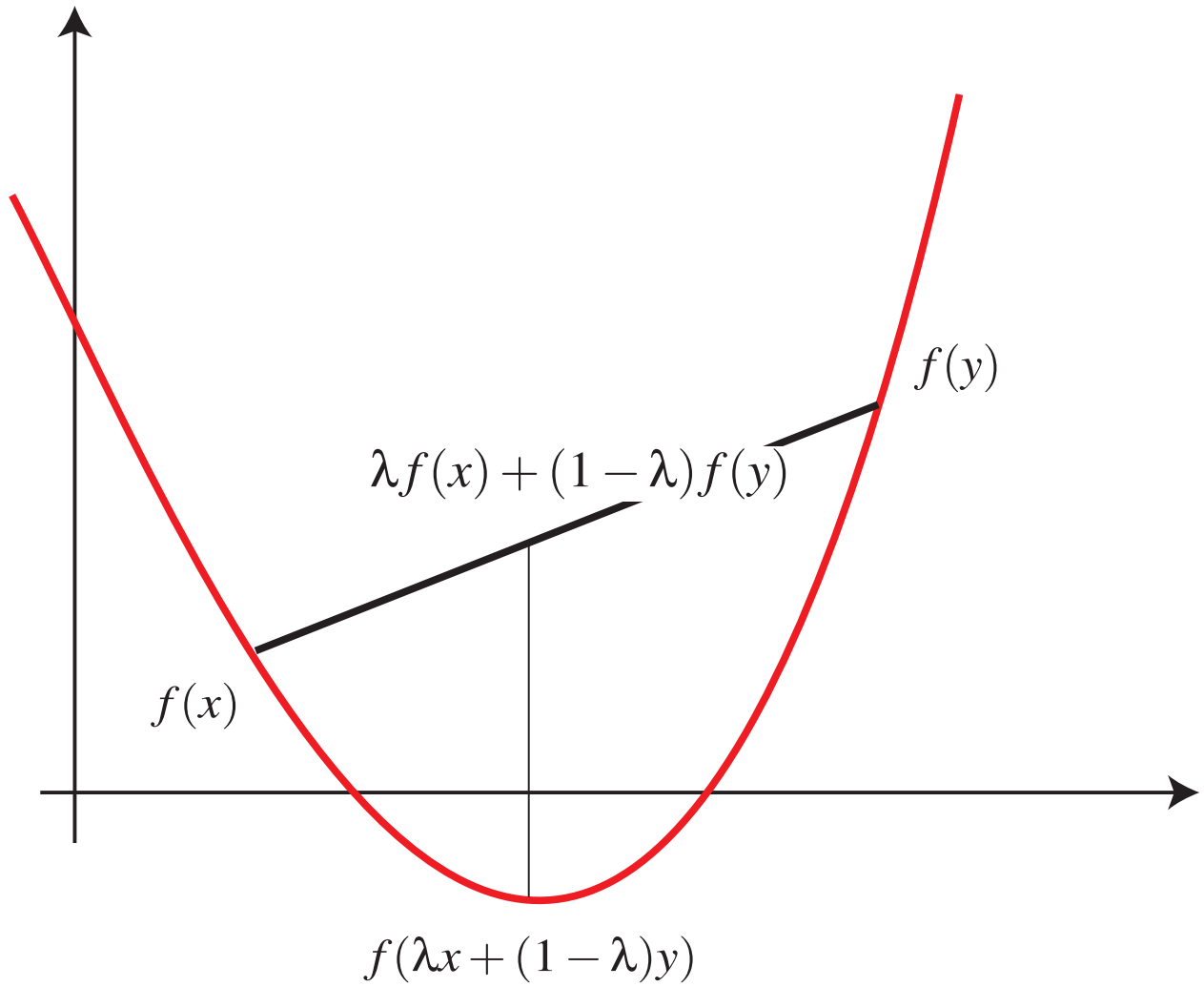
  A function f is convex if

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

  for all $x, y \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$

In this course, emphasis on convex programming/optimization

Convexity is more general than linearity



Applications of mathematical optimization

- Nearly all fields of science and engineering
- Examples
  - Portfolio management
  - Data fitting
  - Regularization in signal processing

Examples of optimization problems

Least Squares

Examples of optimization problems

Least Squares

Examples of optimization problems

Least Squares

- No constraints

Examples of optimization problems

Least Squares

- No constraints
- Objective functional

$$f_0(x) = \|b - Ax\|_2^2 = \sum_{i=1}^{k} (a_i^T x - b_i)^2$$

where $A \in \mathbb{R}^{k \times n}$ (typically $n \leq k$) and $A = \text{row}(a_1, \ldots, a_k)$

Examples of optimization problems

Least Squares

- No constraints
- Objective functional

$$f_0(x) = \|b - Ax\|_2^2 = \sum_{i=1}^{k} (a_i^T x - b_i)^2$$

where $A \in \mathbb{R}^{k \times n}$ (typically $n \leq k$) and $A = \text{row}(a_1, \ldots, a_k)$

- Solution

$$(A^T A)x = A^T b \quad \Rightarrow \quad x^\star = (A^T A)^{-1} A^T b$$

assuming $A^T A$ is invertible (if not?)

Examples of optimization problems

Least Squares

- No constraints
- Objective functional

$$f_0(x) = \|b - Ax\|_2^2 = \sum_{i=1}^{k} (a_i^T x - b_i)^2$$

where $A \in \mathbb{R}^{k \times n}$ (typically $n \leq k$) and $A = \text{row}(a_1, \ldots, a_k)$

- Solution

$$(A^T A)x = A^T b \quad \Rightarrow \quad x^\star = (A^T A)^{-1} A^T b$$

assuming $A^T A$ is invertible (if not?)

- Many methods for solving least squares problems: fast and stable
- Mature technology
- Important and straightforward to recognize a least squares problem

Another example

$$f_0(x) = \|b - Ax\|_2^2 + \lambda \|x\|_2^2$$

- Connected with statistical regularization (ridge regression)
- Trade-off between fit and size of the x_j's
- E.g. prior distribution on parameters in Bayesian statistics

Linear programming

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

Linear programming

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

- No analytical solution

Linear programming

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

- No analytical solution
- Mature technology: fast and stable solvers
  - Simplex method (Dantzig, 47)
  - Interior point methods (Karmakar, 84)

Linear programming

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

- No analytical solution
- Mature technology: fast and stable solvers
  - Simplex method (Dantzig, 47)
  - Interior point methods (Karmakar, 84)

Using linear programming

- Many problems can be recast as LP's
- Not as easy to recognize as least squares problems
- Example

$$\text{minimize} \|b - Ax\|_\infty = \text{minimize} \max_i |a_i^T x - b_i|$$

$$\text{minimize } \|b - Ax\|_\infty = \text{minimize } \max_i |a_i^T x - b_i|$$

$$\text{minimize } \|b - Ax\|_\infty = \text{minimize } \max_i |a_i^T x - b_i|$$

- Objective is maximum absolute deviation, not mean square absolute deviation
- Objective functional is not differentiable

$$\text{minimize } \|b - Ax\|_\infty = \text{minimize } \max_i |a_i^T x - b_i|$$

- Objective is maximum absolute deviation, not mean square absolute deviation
- Objective functional is not differentiable

Recast as

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & a_i^T x - b_i \leq t \\ & a_i^T x - b_i \geq -t \end{aligned}$$

$$\text{minimize } \|b - Ax\|_\infty = \text{minimize } \max_i |a_i^T x - b_i|$$

- Objective is maximum absolute deviation, not mean square absolute deviation
- Objective functional is not differentiable

Recast as

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & a_i^T x - b_i \leq t \\ & a_i^T x - b_i \geq -t \end{aligned}$$

Important to recognize LP's

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} |x_j| \\ \text{subject to} \quad & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{aligned}$$

Is this is an LP?

Convex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \in C \quad \text{or} \quad f_i(x) \leq b_i, \end{array}$$

with f_0 and f_i convex (or C convex set)

Convex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \in C \quad \text{or} \quad f_i(x) \leq b_i, \end{array}$$

with f_0 and f_i convex (or C convex set)

- Solving convex problems
  - In general, no analytical solution
  - Many reliable and efficient algorithms exist
  - Almost mature technology

Convex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \in C \quad \text{or} \quad f_i(x) \leq b_i, \end{array}$$

with f_0 and f_i convex (or C convex set)

- Solving convex problems
  - In general, no analytical solution
  - Many reliable and efficient algorithms exist
  - Almost mature technology

- Using convex optimization
  - Often go unrecognized
  - Many problems can be formulated as convex problems
  - Important to learn the skills to cast nonlinear programs into convex programs
  - Important to distinguish between convex and nonconvex problems

Solving optimization problems

- General optimization problem
  - Very difficult to solve
  - Typical methods are compromising; e.g. extremely long (prohibitive) computation time or solution is not found

Solving optimization problems

- General optimization problem
  - Very difficult to solve
  - Typical methods are compromising; e.g. extremely long (prohibitive) computation time or solution is not found

- Exceptions
  - Least squares problems
  - LP's
  - Convex optimization problems

Solving optimization problems

- General optimization problem
  - Very difficult to solve
  - Typical methods are compromising; e.g. extremely long (prohibitive) computation time or solution is not found

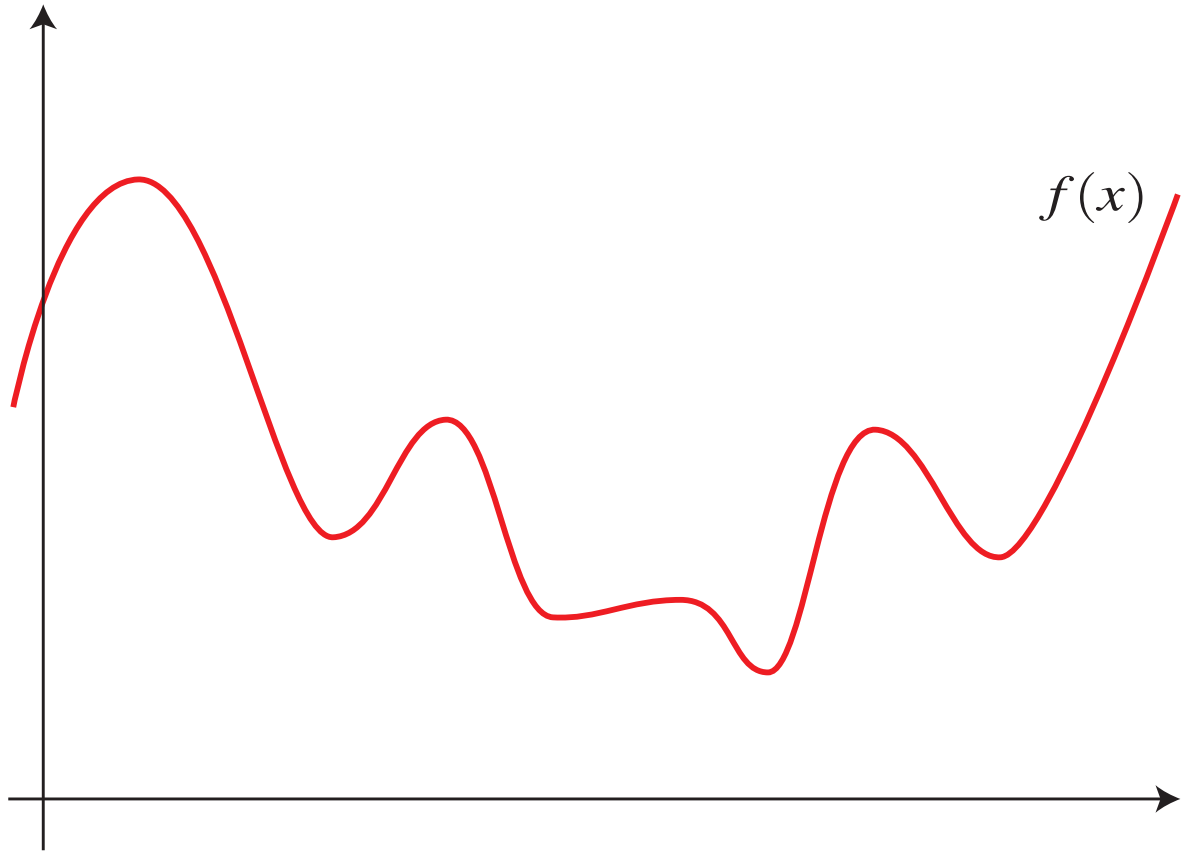
- Exceptions
  - Least squares problems
  - LP's
  - Convex optimization problems

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

Rockafellar, *SIAM Review*, 1993

Nonlinear optimization

1. *Nonconvex objective*



2. Nonconvex constraints

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

- Easy
- Feasible for $n \sim 100,000$

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \end{aligned}$$

- Easy
- Feasible for $n \sim 100,000$

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \text{ and } x_j \in \{0, 1\} \end{aligned}$$

- Extremely hard
- Infeasible for $n \geq 40$?

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \end{aligned}$$

- Easy
- Feasible for $n \sim 100,000$

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \text{ and } x_j \in \{0, 1\} \end{aligned}$$

- Extremely hard
- Infeasible for $n \geq 40$?

- Not much is known about general nonlinear problems
- 'Art' more than a science and technology

Methods

- Local optimization
  - Finds a local minimum
  - Often gives no idea whether this is a global minimum or how far this is from a global minimum
  - Requires initial guess
  - Usually efficient methods

Methods

- Local optimization
  - Finds a local minimum
  - Often gives no idea whether this is a global minimum or how far this is from a global minimum
  - Requires initial guess
  - Usually efficient methods

- Global optimization
  - Worst case complexity is exponential with problem size
  - Small-size problems are not even practical

Insights from convex optimization

- Initialization for local opt. methods
- Convex heuristics for nonconvex problems

Insights from convex optimization

- Initialization for local opt. methods
- Convex heuristics for nonconvex problems

  Example:

$$\begin{array}{ll} \text{minimize} & \#\{j : x_j \neq 0\} \\ \text{subject to} & Ax = b \end{array}$$

  Combinatorial problem (practically impossible)

Insights from convex optimization

- Initialization for local opt. methods
- Convex heuristics for nonconvex problems

Example:

$$\begin{array}{ll} \text{minimize} & \#\{j : x_j \neq 0\} \\ \text{subject to} & Ax = b \end{array}$$

Combinatorial problem (practically impossible)

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} |x_j| \\ \text{subject to} & Ax = b \end{array}$$

Convex problem. Sometimes, same solution (Candes et al., 2004)

Insights from convex optimization

- Initialization for local opt. methods
- Convex heuristics for nonconvex problems

  Example:

$$\begin{array}{ll} \text{minimize} & \#\{j : x_j \neq 0\} \\ \text{subject to} & Ax = b \end{array}$$

  Combinatorial problem (practically impossible)

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} |x_j| \\ \text{subject to} & Ax = b \end{array}$$

  Convex problem. Sometimes, same solution (Candes et al., 2004)

- Bounds for global optimization: dual of an optimization problem is always convex $\rightarrow$ gives a lower bound on the optimal value

Insights from convex optimization

- Initialization for local opt. methods
- Convex heuristics for nonconvex problems

  Example:

$$\begin{array}{ll} \text{minimize} & \#\{j : x_j \neq 0\} \\ \text{subject to} & Ax = b \end{array}$$

  Combinatorial problem (practically impossible)

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} |x_j| \\ \text{subject to} & Ax = b \end{array}$$

  Convex problem. Sometimes, same solution (Candes et al., 2004)

- Bounds for global optimization: dual of an optimization problem is always convex $\rightarrow$ gives a lower bound on the optimal value

- Relaxation: nonconvex constraints replaced with looser convex constraints

This course

- Recognize and formulate convex problems
- Develop algorithms and efficient code
- Learn about the useful theory

This course

- Recognize and formulate convex problems
- Develop algorithms and efficient code
- Learn about the useful theory

Topics

- Convexity: convex sets, functions, optimization
- Algorithms
- Examples and applications