Class #2

Friday, January 24, 2025 10:20 AM

* An optimization algorithm is a numerical method that helps us numerically find a value of x close to an optimal x^* in a reasonable amount of time as a function of 'n' and other problem parameters.

↳ Mathematical programming

* All optimization problems that are min or max can be thought of as minimization problems only.

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

$\max_x$ ∩

$$\arg\max_x f(x) = \arg\min_x -f(x)$$

$= \min_x$ ∪

$\min_x \max_y f(x,y)$ $\Rightarrow$ not the focus of this course

Constrained optimization

$$\min_{x \in \text{dom} f} \boxed{f_0(x)}$$ → objective function

Subject to → Constraint functions

$$f_i(x) \leq b_i, \quad i = 1, \dots, m$$

b_i → known constraint values

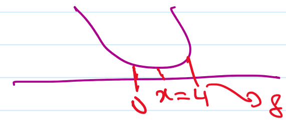
We have 'm' constraints

$\iff \min_{x \in \mathcal{X}} f_0(x)$: $\mathcal{X}$ is a subset of dom f

$$\mathcal{X} = \{ x : f_i(x) \leq b_i, \ i = 1, \dots, m \}$$

e.g., $\min \ (x-4)^2$

e.g., $\min_{0 \le x \le 8} (x-4)^2$



$$\mathcal{X} = \{x : x \in [0,8]\}$$

$0 \le x \le 8$

$\Rightarrow \quad x \le 8$

$\Rightarrow \quad -x \le 0$

$f_1(x) = x \Rightarrow b_1 = 8$

$f_2(x) = -x \Rightarrow b_2 = 0$

$$\min_{x \in \mathrm{dom} f} (x-4)^2$$

s.t. $\quad x \le 8$

$\quad -x \le 0$

Any $x \in \mathcal{X}$ is termed a 'feasible solution'.

In unconstrained optimization, all x's are feasible.

Classes of optimization Problems

Depending on $\boxed{f_0(x)}$, $f_i(x)$, $i = 1, \ldots, m$, and the domain of $f_0(x)$, we characterize an optimization problem to be from a certain class of problems, all of which can be solved using a particular set of numerical tools.

① Combinatorial Optimization

$\Rightarrow$ x takes on values from a discrete set (i.e., dom f corresponds to a discrete set).

$$x \in \{0,1\}^n$$

② Linear programming

The objective function $f_0(x)$ as well as the constraint functions $f_i(x), i=1,\dots,m$, are linear function of x.

$\Downarrow$

Integer programming

↳ Linear programs in which x takes only integer values.

③ Convex programming

↳ The $\boxed{\text{function } f_0(x)}$, the constraint functions $\boxed{f_i(x)}$ $i, i=1,\dots,m$, are all convex.

A function f is convex if and only if

$$f(\alpha x + \beta y) \boxed{\leq} \alpha f(x) + \beta f(y)$$

$\forall$ $\alpha, \beta \geq 0$, $\alpha + \beta = 1$ $\Longrightarrow$ $\alpha = 1 - \beta$

$\Longleftrightarrow$ $f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$

$\forall$ $\theta \in [0,1]$

$x, y \in \operatorname{dom} f$ and $\theta x + (1-\theta) y \in \operatorname{dom} f$

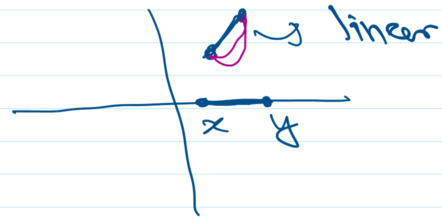
$\Longleftrightarrow$ $\operatorname{dom} f$ has to satisfy the convexity property that $\forall$ $x, y \in \operatorname{dom} f$

$$\theta x + (1-\theta) y \in \operatorname{dom} f \quad \forall \; \theta \in [0,1]$$

In other words, dom f has to be a convex set.

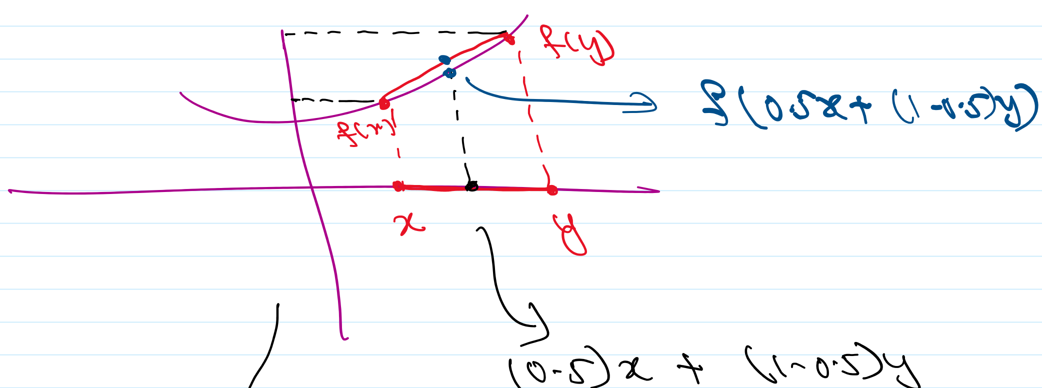
Linear function : $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

All linear functions are convex

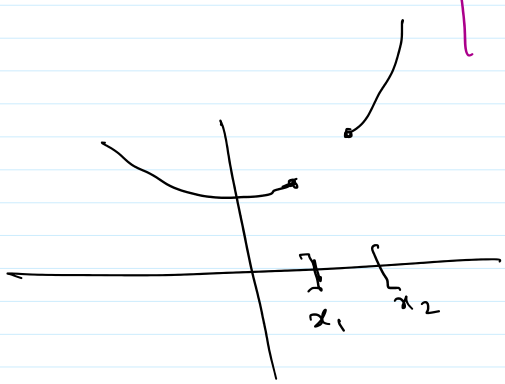


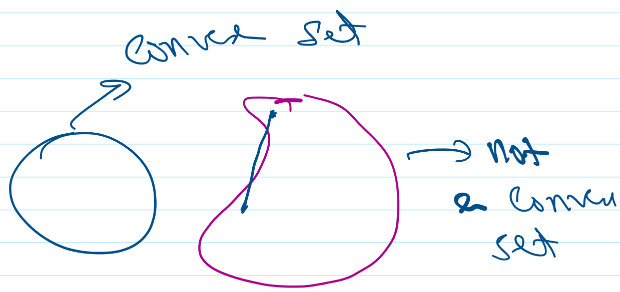
$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$

$\theta \in [0,1]$



E.g.;





④ Non Convex optimization

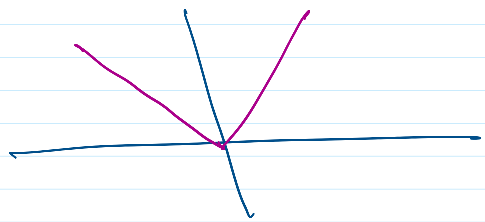
↳ Any optimization that is not convex

⑤ Nonlinear programming

↳ Any optimization that is not linear.

(6) Nonsmooth optimization

↳ when $f_0(x)$ has discontinuous derivatives



(7) Stochastic optimization

↳ When $f_0(x)$ or $f_i(x)$ have randomness involved in them.

Global optimization method

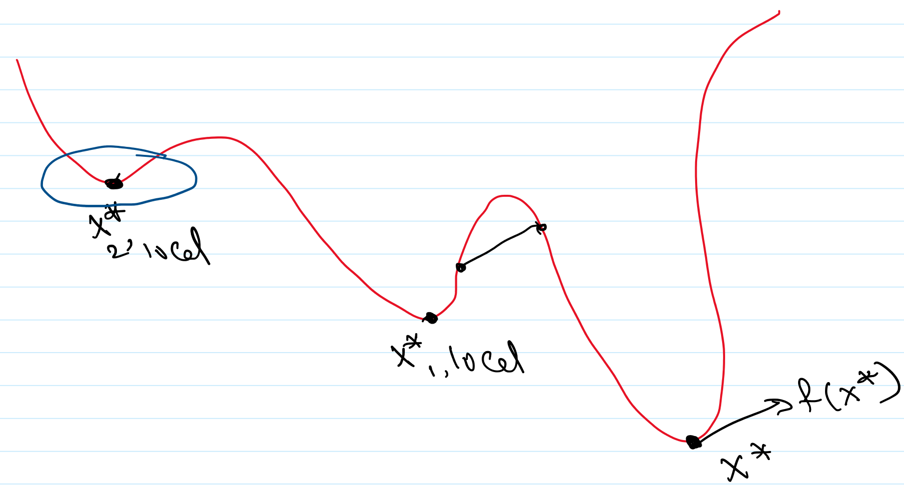
↳ A Solver that provides an optimal solution x^*.

Convex optimization ⇒ we will see that global optimization is straight forward.

Nonconvex optimization ⇒ This is not computationally feasible in most cases (NP-hard).

Local optimization

↳ Find a Solution x^*_{local} that is optimal in some neighborhood of x^*_{local}.



Why Convex Optimization?

1. Many real-world problems tend to be convex.
2. Convex optimization leads to global solutions.
3. The solvers are "efficient".
4. Many nonconvex problems can be 'relaxed' to convex problems and then solved globally.
5. Convex optim. can be used as an 'initialization' scheme for many nonconvex problems.

Class of Solvers

1. Zeroth-order methods
   ↳ Solve the problem by having access to function values only.

2. First-order methods
   ↳ Make use of first-order partial derivatives

↳ Make use of first-order partial derivatives of $f(x)$,

$$\frac{\partial f(x)}{\partial x_i}, \quad i = 1, \ldots, n$$

↳ Use gradient information of the function

③ Second order methods

↳ Use both gradient information and second-order partial derivatives.

$$\frac{\partial f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n$$

↳ Use Hessian information

Examples of convex optimization

① Least Squares problem

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

↳ Quadratic program

$A : \mathbb{R}^{k \times n}$ matrix

$b : \mathbb{R}^{k \times 1}$ vector

$b \Rightarrow$ observations vector (K observations)

$A \Rightarrow$ known matrix

Find closest x s.t. $\boxed{\underline{A}x \approx \underline{b}}$

If $K \geq n$ and A is full rank

$x^* = (A^T A)^{-1} A^T b$

↳ Analytical solution

↳ we can use convex optimization tools to solve this in $O(n^2 K)$ time.