EE 5606 Convex Optimization

Course homepage

https://people.iith.ac.in/shashankvatedka/html/courses/2023/EE5606/course_details.html

Timetable slot

This course:

- math and programming
- requires linear algebra/matrix theory
- programming in python - some tutorials will be provided

Introduction

Why study this course?

Nearly every engineering problem is an optimization problem

Examples

1. Chip design spau

2. Wireless communication

Examples

3. Signal denoising

4. Object detection in images

Examples

5. Portfolio optimization

6. Industrial control

Formal definition of a minimization problem

1. Optimization variable $\underline{x} \in \mathbb{R}^n$
2. Objective function $f(\underline{x})$ $\quad f : \mathbb{R}^n \to \mathbb{R}$
3. Constraints

$$g_1(\underline{x}) \le a_1$$
$$g_2(\underline{x}) \le a_2$$

$$\boxed{\underline{x}^* = \underset{g(\underline{x}) \le 0}{\operatorname{argmin}} \; f(\underline{x})}$$

$$g(\underline{x}) \le 0 \rightarrow \begin{cases} g_1(\underline{x}) \le 0 \\ g_2(\underline{x}) \le 0 \end{cases}$$

Is this definition general enough?

① Maximization problems

$$\max \tilde{f}(x)$$
$$g(x) \leq 0$$

$$f(x) = -\tilde{f}(x)$$

$$1/\tilde{f}(x) \qquad e^{-\tilde{f}(x)}$$

② $$\min f(x)$$
$$g_1(x) \geq \alpha_1$$
$$g_2(x) \leq \alpha_2$$

$$\equiv$$

$$\min f(x)$$
$$-g_1(x) + \alpha_1 \leq 0$$
$$g_2(x) - \alpha_2 \leq 0$$

③

$$\min f(x)$$
$$g_1(x) \le 0$$
$$g_2(x) = 0$$

$$g_2(x) \le 0$$
$$g_2(x) \geq 0$$

$\equiv$

$$\min f(x)$$
$$g_1(x) \le 0$$
$$g_2(x) \le 0$$
$$-g_2(x) \le 0$$

④

$$\min f(x)$$
$$g_1(x) \le 0$$
$$g_2(x) < 0$$

$$+\frac{1}{g_2(x)} \le 0$$

$$\min f(x)$$
$$x < 0$$

$$x - \alpha \le 0$$
$$\alpha \le$$

$$f(x) = x^2$$

$$x < 1$$

Some more examples

The least squares solution for a system of linear equations

$A\underline{x} = \underline{b}$ $\quad$ A: $m \times n$

$BA = I_n$ $\quad$ $m > n$ — # of solutions =

$m < n$ - ∞, 0, $\quad$ 0/1

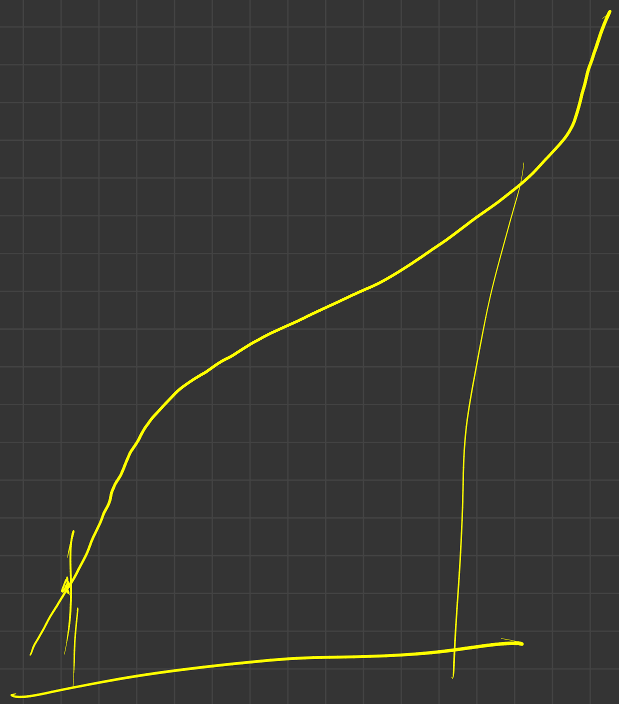
$f(\underline{x}) = \|A\underline{x} - \underline{b}\|_2^2$ $\quad$ $\underline{x} \in \mathbb{R}^n$

$\min f(\underline{x})$ $\quad$ $\underline{x}^* = (A^T A)^{-1} A^T \underline{b}$

Constrained least squares

$$f(\underline{x}) = \| A\underline{x} - \underline{b} \|^2$$

$$\| \underline{x} \| \leq 1$$



Power allocation in Gaussian channels

$$\begin{tikzcd} x_1 \arrow[r] & \oplus \arrow[r] & y_1 \\ & z_1 \arrow[u] & \end{tikzcd}$$

$$z_1 \sim \mathcal{N}(0, \sigma_1^2)$$

$$\begin{tikzcd} x_2 \arrow[r] & \oplus \arrow[r] & y_2 \\ & \arrow[u] & \end{tikzcd}$$

$$z_2 \sim \mathcal{N}(0, \sigma_2^2)$$

$$x_1^2 + x_2^2 + \cdots + x_m^2 \leq P$$

$$P_1 \quad P_2 \quad P_m$$

$$R = \sum_{i=1}^{m} \frac{1}{2} \log_2\left(1 + \frac{P_i}{\sigma_i^2}\right) = f(x)$$

$$P_1 + P_2 + \cdots + P_m \leq P$$

Empirical risk minimization

Given: Labeled data

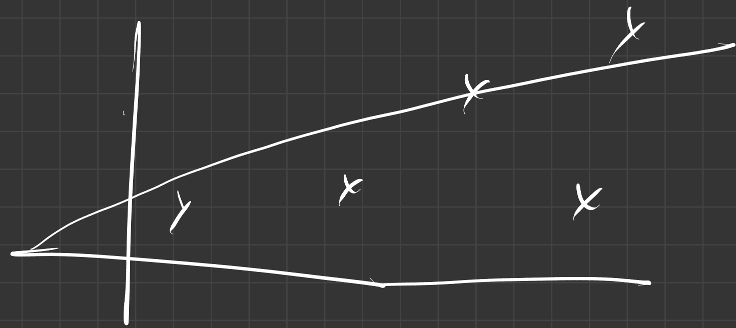
Ground truth $f: I \to \{0, 1\}$

No cat

cat

Data: $\{ (I_1, l_1), (I_2, l_2), \ldots \ldots (I_m, l_m) \}$

Take a subclass of functions $\{ f_{\underline{\alpha}} : \underline{\alpha} \in \mathbb{R}^n \}$



Loss function

$L : \{0, 1\} \times \{0, 1\} \to \mathbb{R}_{\geq 0}$

$L(l_1, l_2)$

0-1 loss:

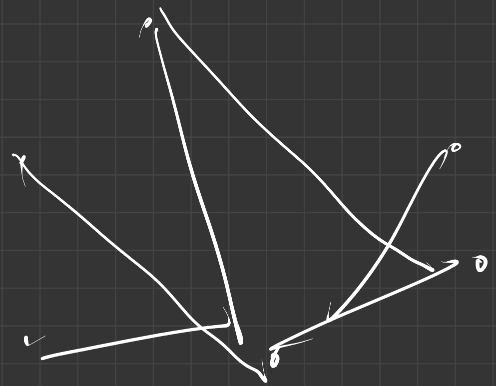
$$L(l, l') = \begin{cases} 1 & \text{if } l \neq l' \\ 0 & \text{else} \end{cases}$$

$$\underbrace{G(\underline{\alpha}) = \frac{1}{M} \sum_{i=1}^{M} \underbrace{L(l_i, f_{\underline{\alpha}}(I_i))}_{\text{loss for ith data pt}}}_{\text{Empirical risk}}$$

Compute

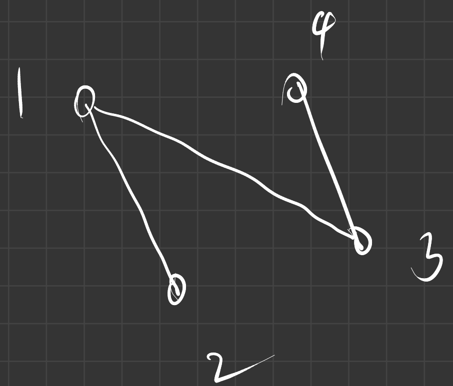
$$\underline{\alpha}^* = \arg\min_{\underline{\alpha}} G(\underline{\alpha})$$

Computing maximum cut of an undirected graph



Adjacency matrix : A $n \times n$ matrix

$$A_{ij} = \begin{cases} 1 & \text{if } i \,\&\, j \text{ are connected} \\ 0 & \text{else} \end{cases}$$



$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Cut: Given a graph G, and any partition of the vertex set V into $V_1 \cup V_2$, the # of edges going from V_1 to V_2 is the size of the cut (V_1, V_2)

The MAX-CUT of G is the size of the largest cut.

Optimization variable: $V_1 \subseteq V$

$\underline{x} \in \{+1, -1\}^n$

$$f(x) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\,|x_i - x_j| \qquad \Bigg| \; i \in V_1 \quad \sum_{i \in V_1}\sum_{j \in V_2} A_{ij}$$

$$\frac{1}{2}\,\frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\,(x_i - x_j)^2$$

$$\frac{1}{8}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\,(x_i^2 + x_j^2 - 2x_i x_j)$$

$$= \frac{1}{8}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\,2(1 - x_i x_j)$$

$$= \frac{1}{8}\sum_{i=1}^{n}\left[2\sum_{j=1}^{n} A_{ij} - 2\sum_{j=1}^{n} A_{ij}\,x_i x_j\right]$$

$$= \frac{1}{4}\left(\sum_{i=1}^{n} d_i - \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\, x_i\, x_j\right)$$

$$L = D - A$$

$$= \frac{1}{4}\, \underbrace{\underline{x}^T L\, \underline{x}}_{\sum_{i=1}^{n}\sum_{j=1}^{n} L_{ij}\, x_i\, x_j}$$

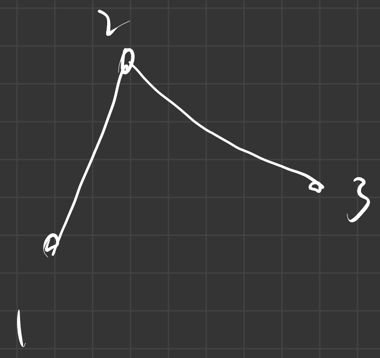
ith entry of $L\underline{x}$

$$y_i = \sum_{j=1}^{n} L_{ij}\, x_j$$

$$\underline{x}^* = \underset{\underline{x} \in \{\pm 1\}^n}{\operatorname{argmax}} \; \underline{x}^T L \underline{x}$$

Relaxation : Approximate this by an "easier" problem

- Relax the constraints
- Relax the fn : surrogate function



$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$Ax = \begin{bmatrix} x_2 \\ x_1 + x_3 \\ x_2 \end{bmatrix}$$

$$\min_{\underline{x} \in \{\pm 1\}^3} \underline{x}^T A \underline{x}$$

$$x^T A x =$$

$$x_1 x_2 + x_2(x_1 + x_3$$

$$+ x_2 x_3$$

$$2x_2(x_1 + x_3) = 2(x_1 x_2 + x_2 x_3) =$$

Convex Optimization in R

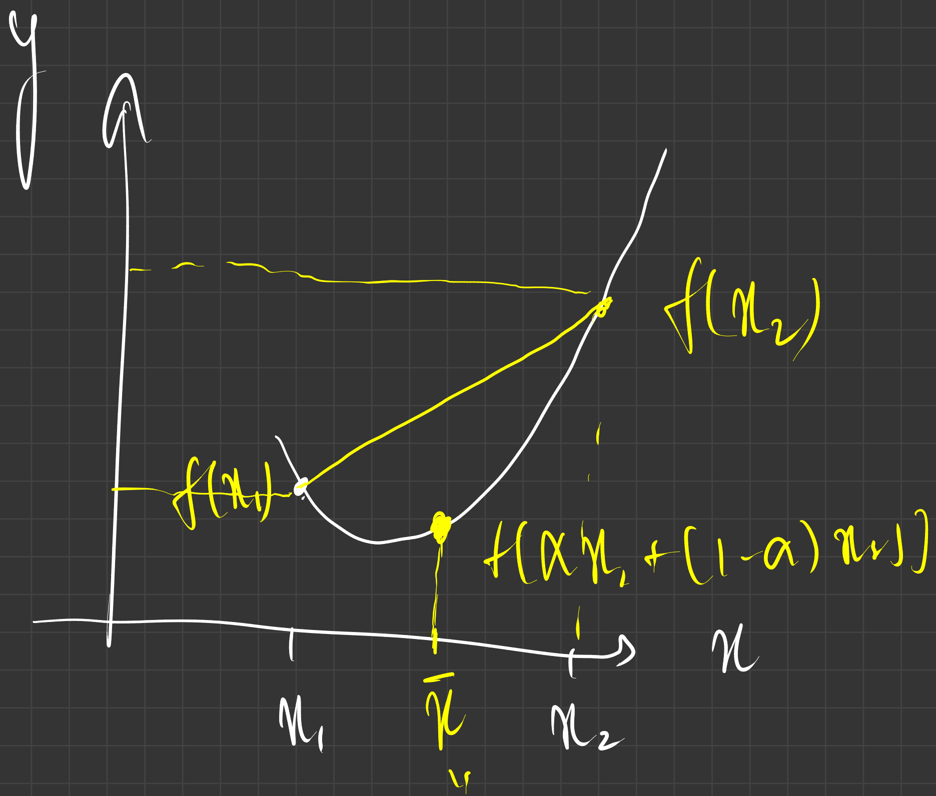
Convex function

Definition: $f: \mathbb{R} \to \mathbb{R}$ is convex if for all $x_1, x_2 \in \mathbb{R}$ & $0 \le \alpha \le 1$,

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)$$

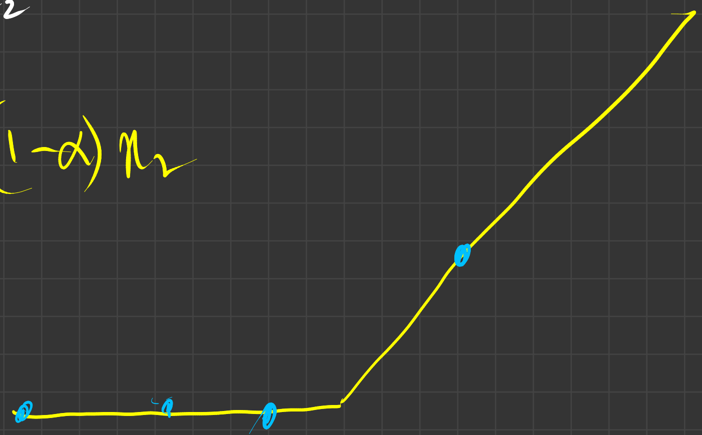
If equality above holds only for $\alpha = 0, 1$ then ~~strongly~~ strictly convex

f is concave if $-f$ is convex



$$\alpha x_1 + (1-\alpha)x_2$$

$$\alpha x_1 + (1-\alpha)x_2$$

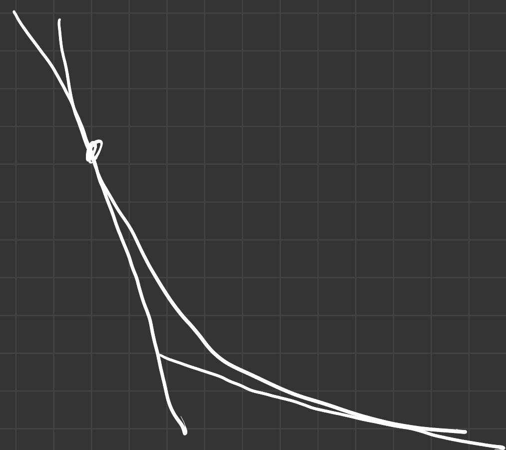


$$f(x) = \begin{cases} 0 & \text{if } x \le 0 \\ x & \text{if } x > 0 \end{cases}$$

Second derivative test

$\boxed{\text{A function } f: \mathbb{R} \to \mathbb{R} \text{ whose } f''(x) \text{ exists at all pts is convex if \& only if (iff) } f''(x) \geq 0 \ \forall x \in \mathbb{R}}$

① If $f''(x) \geq 0 \ \forall x$, then f is convex.



Take x_1, x_2, $0 \le \alpha \le 1$

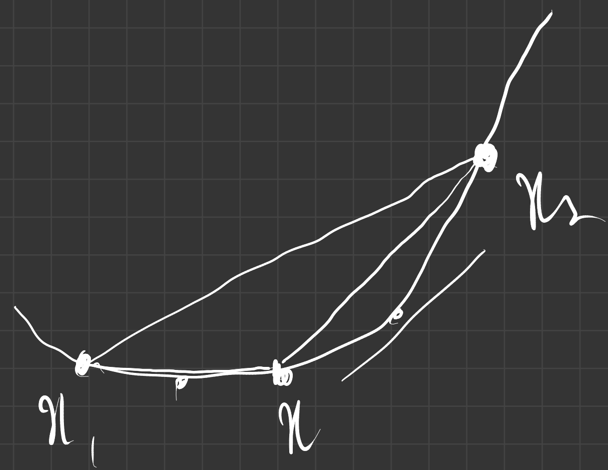
$$f(\alpha x_1 + (1-\alpha) x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)$$

Consider $\alpha f(x_1) + (1-\alpha) f(x_2) - f(\underbrace{\alpha x_1 + (1-\alpha) x_2}_{x})$

$$= \alpha \left(f(x_1) - f(x) \right) + (1-\alpha) \left(f(x_2) - f(x) \right)$$

$$= \alpha (x_1 - x) \frac{f(x_1) - f(x)}{x_1 - x}$$

$$+ (1-\alpha) (x_2 - x) \frac{f(x_2) - f(x)}{x_2 - x}$$



$$= \alpha (x_1 - x) f'(\beta_1)$$

MVT

$$+ (1-\alpha)(x_2 - x) f'(\beta_2)$$

$$x_1 < \beta_1 < x < \beta_2 < x_2$$

$$= \alpha \left[x_1 - \alpha x_1 - (1-\alpha) x_2 \right] f'(\beta_1)$$

$$+ (1-\alpha) \left[x_2 - \alpha x_1 - (1-\alpha) x_2 \right] f'(\beta_2)$$

$$= -\alpha (1-\alpha)(x_2 - x_1) f'(\beta_1) + (1-\alpha) \alpha (x_2 - x_1) f'(\beta_2$$

$$= \alpha (1-\alpha)(x_2 - x_1) \left(f'(\beta_2) - f'(\beta_1) \right) \geq 0$$

Want to ST if f is convex & $f''(x)$ exists for all x, then

$$f''(x) \geq 0 \quad \forall x \in \mathbb{R}$$

$$f''(x) = \lim_{t \downarrow 0} \frac{f'(x+t) - f'(x)}{t}$$

$$= \lim_{t \downarrow 0} \frac{\left(\frac{f(x+t) - f(x)}{t}\right) - \left(\frac{f(x) - f(x-t)}{t}\right)}{t}$$

$$= \lim_{t \downarrow 0} \frac{f(x+t) + f(x-t) - 2f(x)}{t^2}$$

$$f(x+t) + f(x-t) - 2f(x)$$

$$= 2\left[\frac{f(x+t)}{2} + \frac{f(x-t)}{t} - f(x) \right]$$

$$\alpha = \frac{1}{2} \qquad x_1 = x+t$$

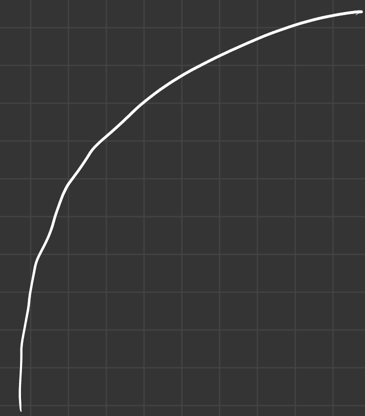
$$1-\alpha = \frac{1}{2} \qquad x_2 = x-t$$

$$\geq 2\left[\underbrace{f\left(\frac{x+t}{2} + \frac{x-t}{2} \right)}_{f(x)} - f(x) \right] \geq 0$$

<u>Examples</u>

① $f(x) = \log x$

$\rightarrow$ Concave



② $f(x) = e^{-x}$

Convex

③ $f(x) = x \log x$

$f'(x) = 1 + \ln x$

$f''(x) = \frac{1}{x} > 0$ for $x > 0$

④ $f(x) = x\ln x + (1-x)\ln(1-x)$, $0<x<1$

Convex

⑤ $f(x) = (1-x)\ln(1-x)$ $\quad 0<x<1$

Convex

⑥ $f(x) = |x|$

$$f(x) = \begin{cases} x, & x \geq 0 \\ -x+1, & x<0 \end{cases}$$

Take $x_1 < x_2$

$f(\alpha x_1 + (1-\alpha)x_2)$

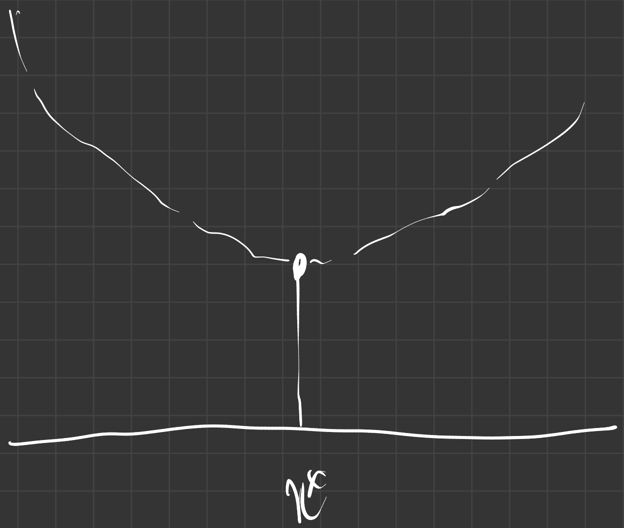
$= |\alpha x_1 + (1-\alpha)x_2| \leq |\alpha x_1| + |(1-\alpha)x_2|$

$$= \alpha|x_1| + (1-\alpha)|x_2|$$

$$= \alpha f(x_1) + (1-\alpha) f(x_2)$$

convex

$f(x)$: $f'(x^*) = 0$



$\forall x < x^*$, $f'(x) \leq f'(x^*) = 0$

$\Downarrow$

$f(x) \geq f(x^*)$

$\forall x > x^*$, $f'(x) \geq f'(x^*) = 0$

$f(x) \geq f(x^*)$