Tutorial 1

05 January 2022 19:11

Last time Matrix completion Problem

movies

$$\begin{array}{c|ccccc} & 1 & 2 & 3 & \cdots & M \\ \hline 1 & R_{11} & & R_{13} & & R_{1M} \\ 2 & & R_{22} & & & R_{2M'} \\ \vdots & & & & & \\ i & & & R_{ij} & & \\ \vdots & & & & & \\ N & & & & & \end{array}$$

customers

i^{th} customer's rating of j^{th} movie

- Goal: Estimate missing entries

Assume: Each customer has a feature vector of length k

Each movie also has a feature vector of length k

→ u_i: feature vector of customer i

→ v_j: feature vector of movie j

$$R_{ij} = u_i^T v_j = \sum_{l=1}^{k} u_i(l)\, v_j(l).$$

$l = 1$ action

$f_1 =$ Action $\quad f_2 =$ Comedy $\quad \cdots \quad f_k$

→ $v_j(1)$ $\quad v_j(2)$ $\quad v_j(k)$

→ $u_i(1)$ $\quad u_i(2)$ $\quad u_i(k)$

$u_1 \ldots u_N, \; v_1 \ldots v_M$

$R_{ij} \approx u_i^T v_j$

$R_{23} \approx u_2^T v_3$

$R_{210} = u_2^T v_{10}$

$$\min_{u_i, v_j} \sum_{(i,j) \in \Lambda} \left(R_{ij} - u_i^T v_j \right)^2$$

$R_{ij} - u_i^T v_j$

indices of the known entries of the matrix

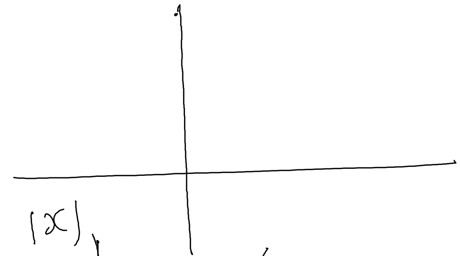
u_i^*, v_j^*

$R_{ij} - u_i^T v_j$

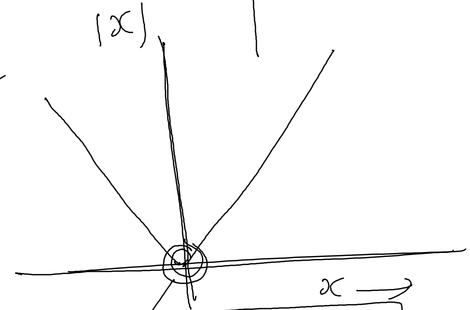
u_i^*, v_j^*

$i=1,\dots,m, \quad j=1,\dots,n$

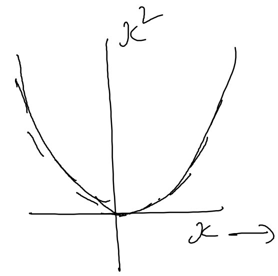
$R_{ij} = u_i^T v_j$



$$\min_{u_i, v_j} \sum_{(i,j)\in\Omega} \left((R_{ij} - u_i^T v_j)\right)^2$$



$\boxed{\text{Not-differentiable}}$



$$\frac{d(x^2)}{dx} = 2x \quad \forall x$$

$$\frac{d|x|}{dx} = ?$$

$$\begin{cases} x>0 & \frac{d(x)}{dx} = 1 \\ x<0 & \frac{d(-x)}{dx} = -1 \end{cases}$$

$x=0 \quad \frac{d|x|}{dx}$ not defined

non-differentiable → sub-differential

Convex fn

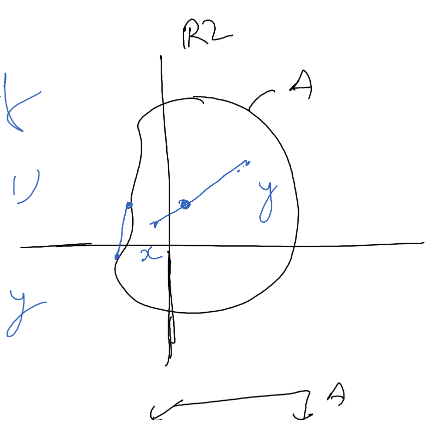
$\mathbb{R}^n \quad A \subseteq \mathbb{R}^n$

Convex set: A is called convex if

$\forall x, y \in A \quad \alpha x + (1-\alpha) y \in A \quad \forall \alpha \in (0,1)$

a convex combination of x and y

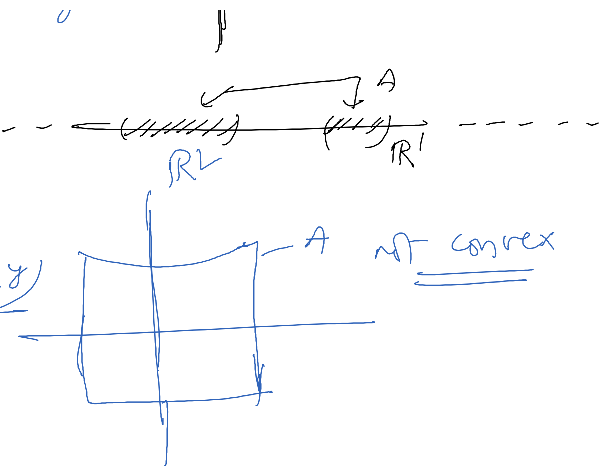
a convex set



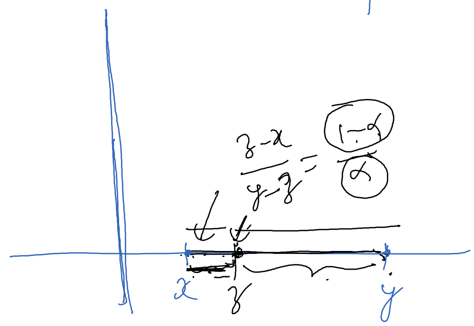
a convex set

$f : A \to \mathbb{R}$ is called a convex function if $\forall\, x, y \in A$, $\alpha \in (0,1)$

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



$f : \mathbb{R} \to \mathbb{R}$

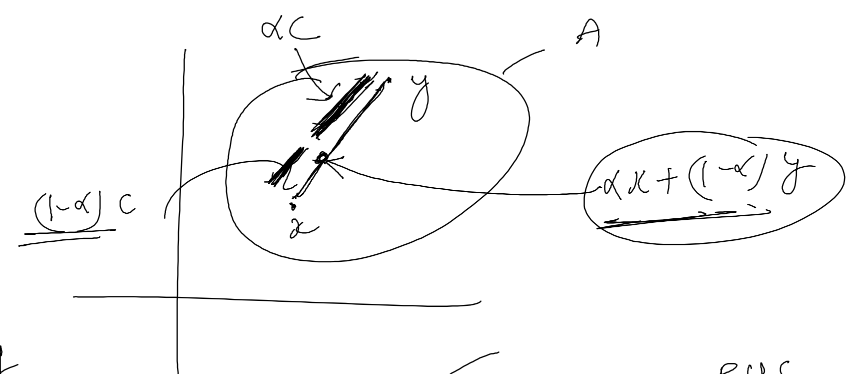


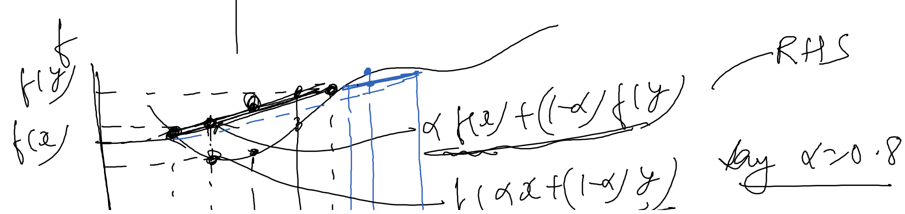
$$z = \alpha x + (1-\alpha) y$$

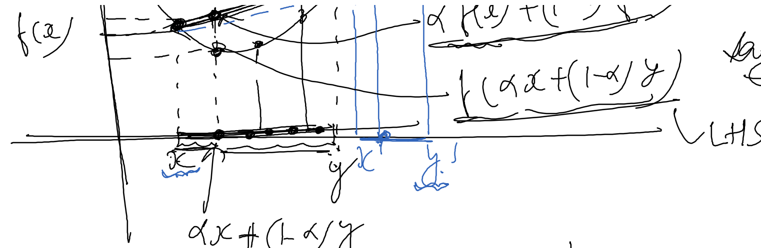
$$z - x = \alpha x + (1-\alpha) y - x = (y-x)(1-\alpha)$$

$$y - z = y - (\alpha x + (1-\alpha) y) = \alpha (y-x)$$

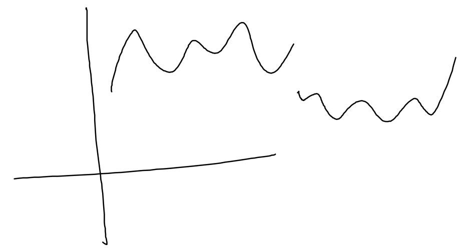
$$\boxed{\frac{z-x}{y-z} = \frac{1-\alpha}{\alpha}}$$

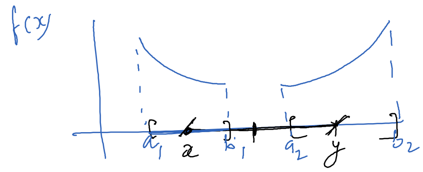






say $\alpha = 0.8$

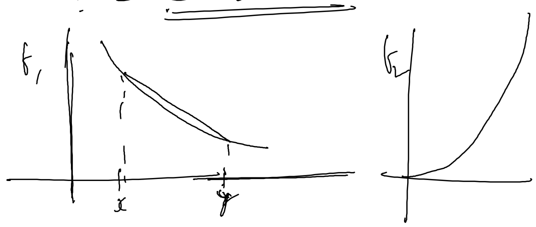




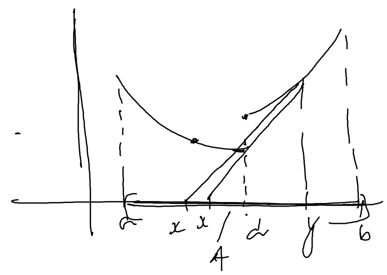
$A = [a_1, b_1) \cup [a_2, b_2)$

not convex

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



f_1, f_2 both are convex

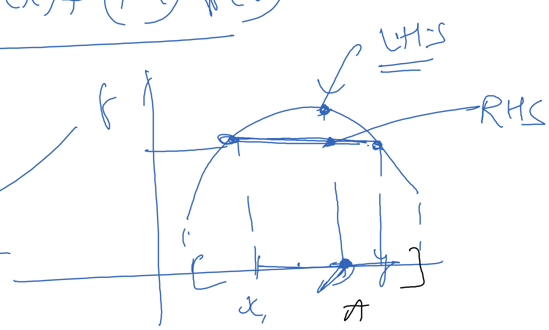


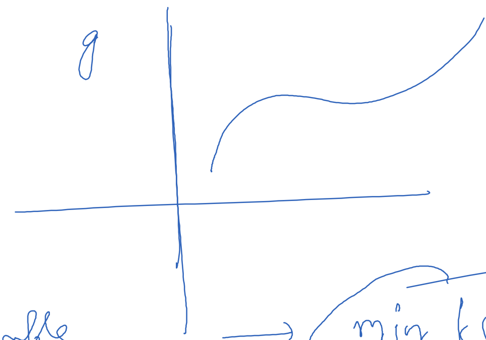
Concave fn A is convex set $f : A \to \mathbb{R}$

$\forall x, y \in A$

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

Concave

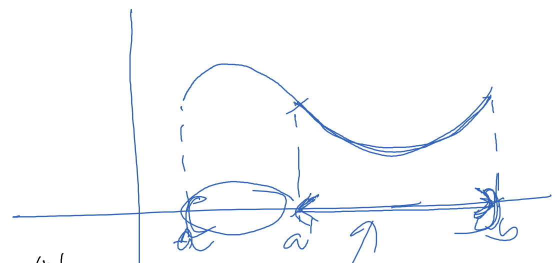


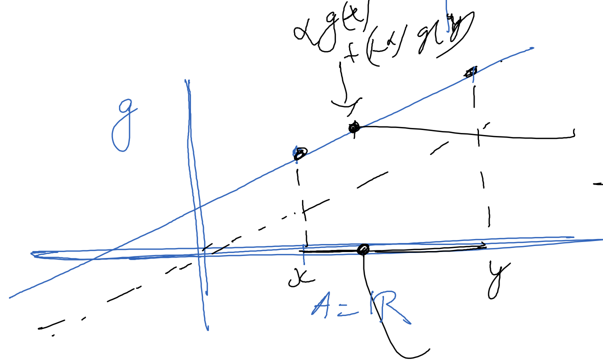


neither concave nor convex

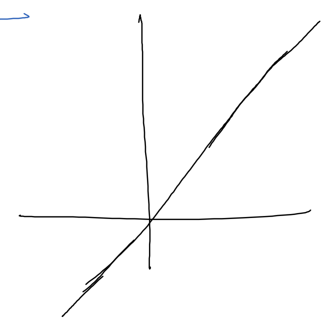
Example $\longrightarrow$ $\min\limits_{x \in A} f(x) = \min\limits_{x \in [a', b]} f(x)$

convex set

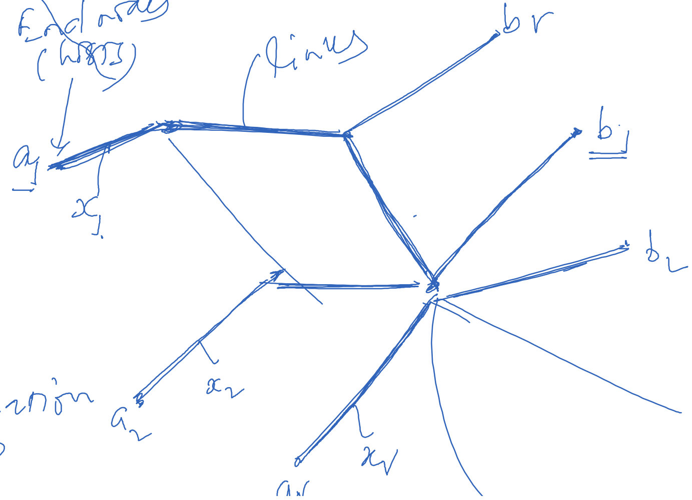




linear functions are convex as well as concave



Ex 3 "Networking" (Internet Resource Allocation)



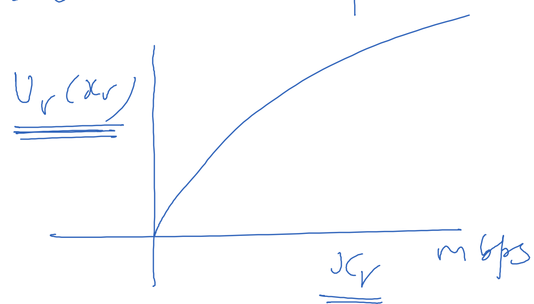
R source destination pairs

R source destination pairs a_r

a_r x_r

switches/routers

- same link may be shared by many source destination pairs ✓
- Each link has a given capacity.
- say link e has capacity c_e
- Each src-dest pair has a utility for the data rate
- $U_r(x_r)$ is the utility of src-destination pair r for data rate x_r



service provider

- ~~Controller~~ can tell users x_r
- controller wants to optimize

$$\boxed{\sum_{r \in R} U_r(x_r)}$$

$$\max_{\{x_r\}} \sum_{r \in R} U_r(x_r)$$

$$\text{s.t.} \quad \sum_{r;\, r \text{ passes through } e} x_r \le c_e \quad \forall e \in E$$

$$x_r \ge 0$$

- NUM problem
  (network Utility maximization problem)

Linear algebra basics

1. for $A, B \in \mathbb{R}^{n \times n}$

1. for $A, B \in \mathbb{R}^{n \times n}$

$$\det(AB) = \det(A)\det(B)$$

* $\det(BA) = \det(AB)$ (AB need not equal BA)

2. If $A, B \in \mathbb{R}^{n \times n}$ then eigenvalues of AB are same as eigenvalues of BA

3. If $A \in \mathbb{R}^{n \times n}$ is symmetric then all its eigenvalues are real

4. If $A \in \mathbb{R}^{n \times n}$ is symmetric, its eigenvectors can be chosen to be orthonormal, i.e. one can find eigenvectors

$v_1, \ldots\ldots, v_n$ such that

$$v_i^T v_j = 0 \quad \forall i \neq j$$
$$\|v_i\|_2 = 1 \quad \forall i$$

(λ, x) eigenvalue - eigenvector pair $\Rightarrow Ax = \lambda x$

* $v_1, \ldots\ldots, v_n$ form an orthonormal basis of $\mathbb{R}^n$,

Any $x \in \mathbb{R}^n$ can be represented as

$$x = \sum_{i=1}^{n} \alpha_i v_i$$

5. $A \in \mathbb{R}^{n \times n}$ is symmetric λ_{min} and λ_{max} are its smallest and largest eigenvalues then

$$\lambda_{min} \|x\|^2 \leq x^T A x \leq \lambda_{max} \|x\|^2 \quad \forall x \in \mathbb{R}^n$$