Course logistics

Instructors: Sandeep Kumar and Prof. Jayadeva

E-mail id: ksandeep@ee.iitd.ac.in and jayadeva@iitd.ac.in

Course website: https://sites.google.com/view/sandeepkr/teaching/ell-888-advanced-machine-learning

Grading Scheme The usual assessment of students taking this course shall be performed on the basis of performance in various components of the course.

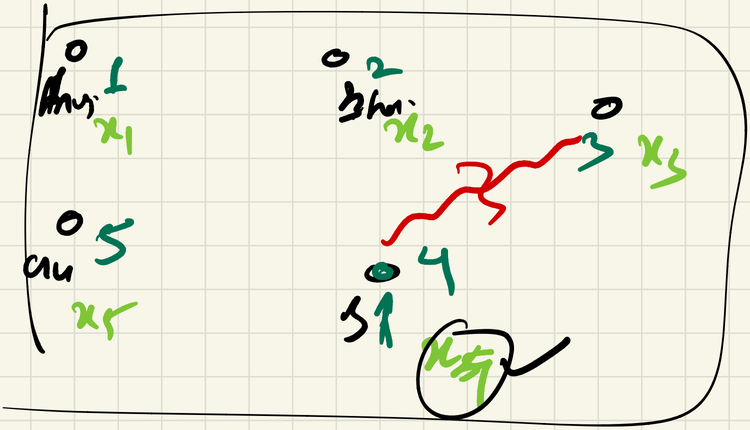
- 20 %: 15% **Lecture Scribbing, Slides, and Demo Example** +5%**variable component**
- 15%: **Minor Exam**
- 25%: **Major exam**
- 40 % **Research Project**

Audit Criterion will be Decided on Class Participation, Minor, and Major Exams.

Class timings: Slot J, Monday, Tuesday, and Friday 12:00 noon to 12:50 pm.

31/1/2022

lecture 2



$\mathbb{R}, \mathbb{I}, \mathbb{N}$

$\mathbb{Q}, \mathbb{Z}$

x_1 x

$X = [\underline{x_1}, \underline{x_2}, \underline{x_3}, \underline{x_4}, x_5]$

height → [178, 166.6 cm] 5ft, 5.6ft

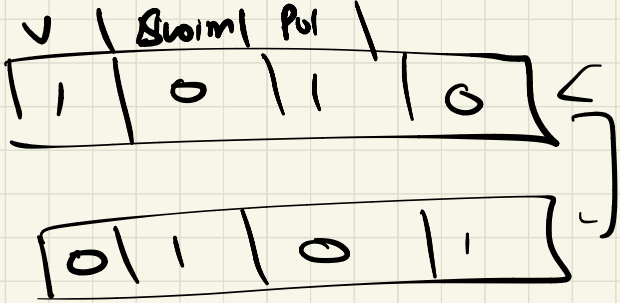
age → [39, 23, 21, 24]

weight

liking/disliking → [0|1]

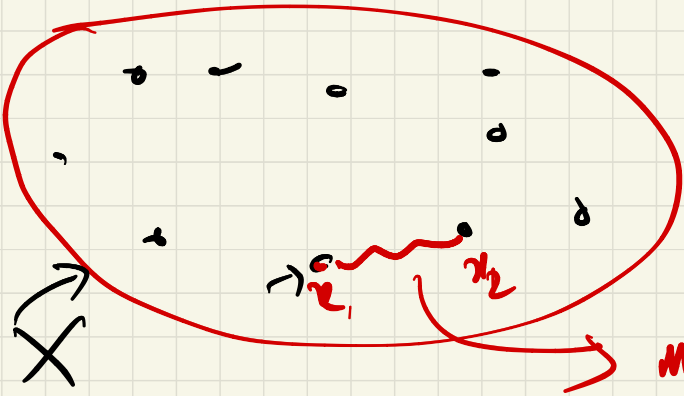
P1 →

P2



$\to x_1, x_2 \ldots x_n \leftarrow \mathbb{R}$

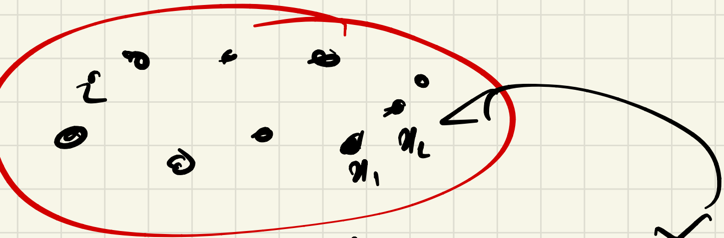
→



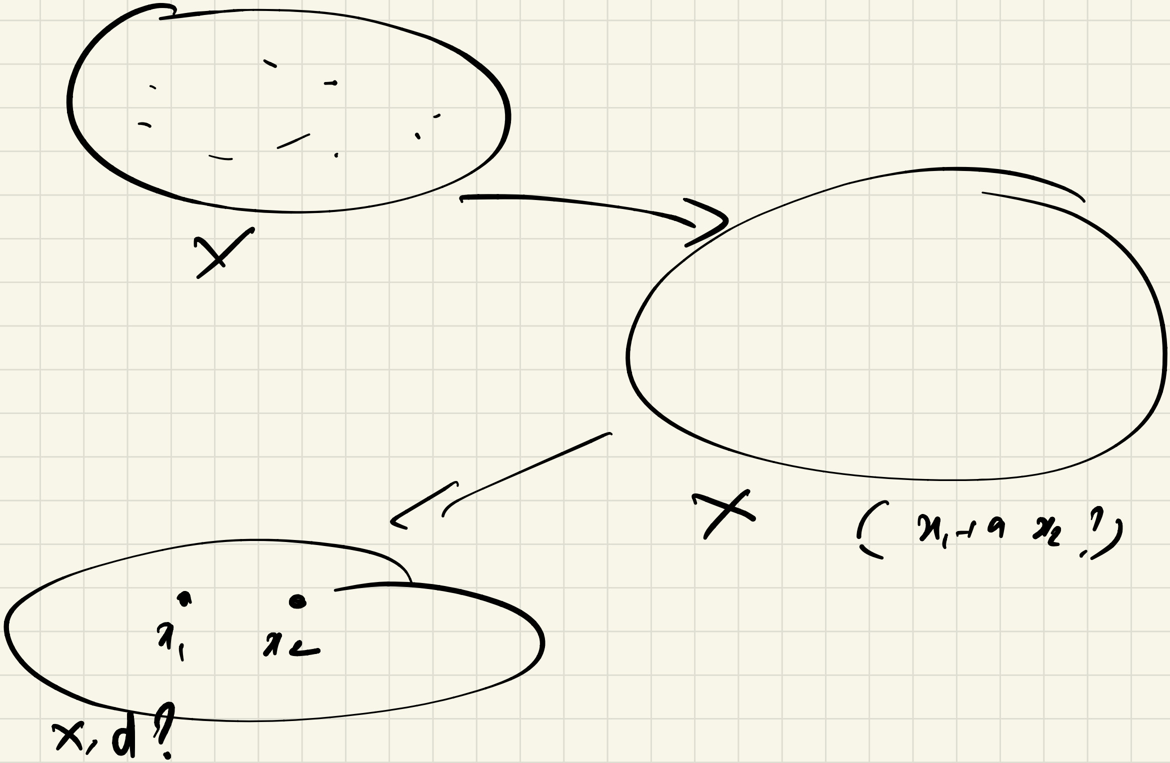
Metric space

H.w

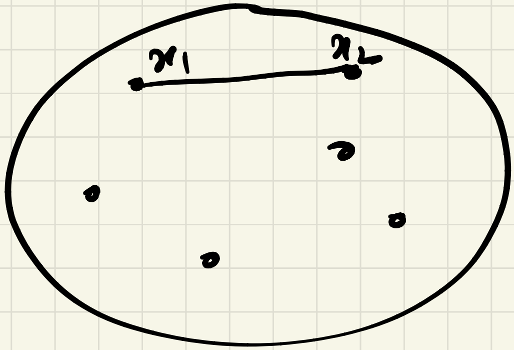
→ $\mathbb{R}, \mathbb{Q}, \mathbb{I}, \mathbb{N}$



$$x_1 + a x_2 \to Y_4$$


X
X
x1
x2
X, d ?

$x_1 = \begin{bmatrix}2\\3\end{bmatrix}$ $x_2 = \begin{bmatrix}3\\4\end{bmatrix}$



$d_2 = \sqrt{(3-2)^2 + (4-3)^2}$

$d_p = |x-y|_p$

(X, d_1) (X, d_2) (X, d_3)

$\to M_1$ $\to M_2$

(X, d)

$d_1 = |x-y|$

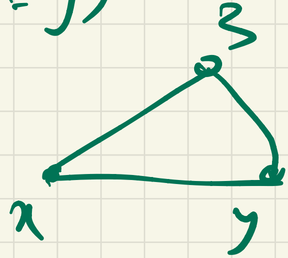
$d_2' = k|x-y|$

Def:

$X \neq \phi$, $d : X \times X \to \mathbb{R}_+$ is called a metric space if 'd' satisfies

(i) $d(x,y) \geq 0 \quad \forall x, y \in X$

$(d(x,y) = 0 \iff x = y)$



(ii) $d(x,y) = d(y,x)$

(iii) $d(x,z) \leq d(x,y) + d(y,z)$

$\forall x, y, z \in X$

$\underline{R^1}$

e.g. a) $d(x,y) = |x-y|$ (b) $d(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$

$d(x,y) = (x-y)^2$

- examples in $\mathbb{R}^n$, $x, z \in X$

$x = (x_1, \dots, x_n)$

$z = (z_1, \dots, z_n)$

$$d_1(x,z) = \sum_{k=1}^{n} |x_k - z_k|$$

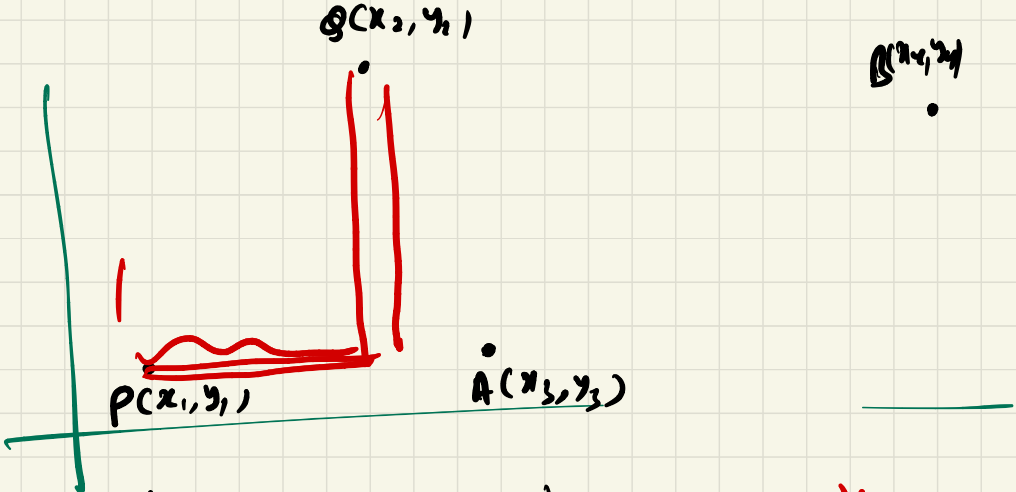
$$d_2(x,z) = \left(\sum_{k=1}^{n} (x_k - z_k)^2 \right)^{1/2}$$

$$d_\infty(x,z) = \max_{k=1,\dots,n} \{ |x_k - z_k| \}$$

e|

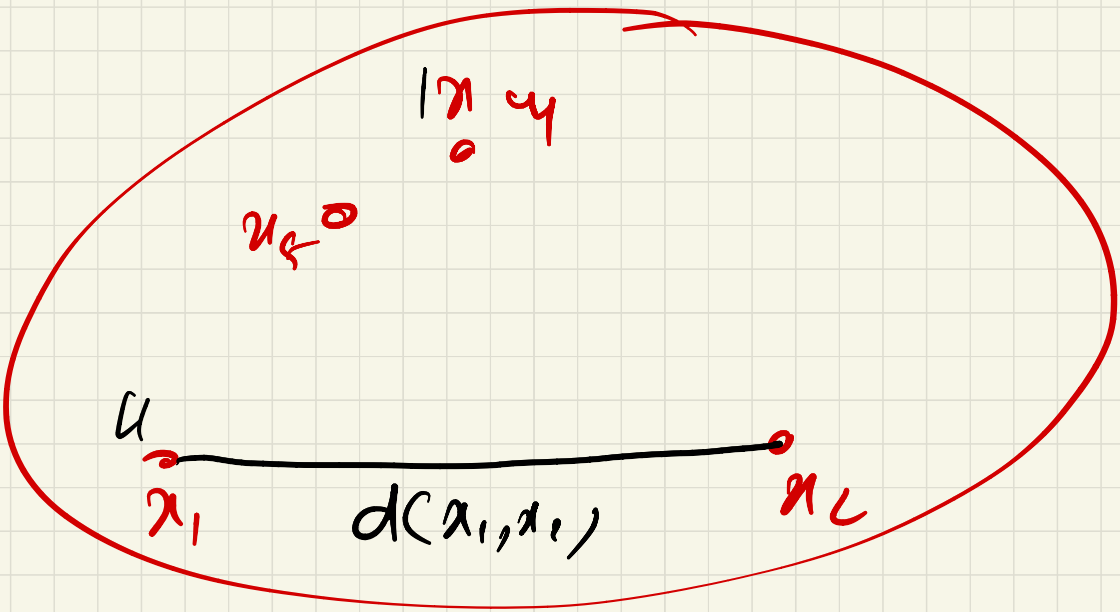
$\bullet\ P = (x_1, y_1)$ $\qquad$ $Q = (x_2, y_2)$

$A = (x_3, y_3)$ $\qquad$ $B = (x_4, y_4)$



$d_1(P, Q) = |x_2 - x_1| + |y_2 - y_1|$ $\quad$ $d_\infty(P, Q) = |y_2 - y_1|$

$d_1(A, B)$ $\qquad$ $d_\infty(A, B) =$



$x_1, \quad x_2$

$$(x_1,x_2) \xrightarrow{d} \delta_{12}$$

$5 > 4$

$R \to \quad 5 \quad 1$

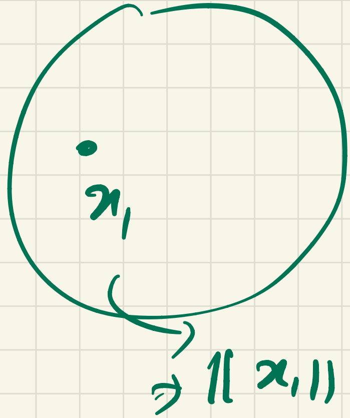
4

Norm

• Vector space V

$\|\cdot\| : V \longrightarrow R_+$

↳ a real valued function



• Prop ~~∀~~ $x \in V$, $\|\cdot\|$ should

(i) $\|x\| = 0 \iff x = 0$

$(\|x\| \geq 0 \quad \forall\, x \in V)$

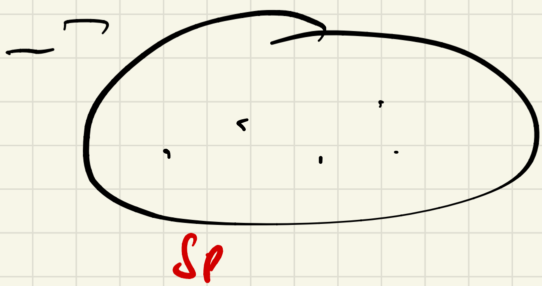
ii) $\|\lambda x\| = |\lambda| \cdot \|x\| \quad \lambda \in R$

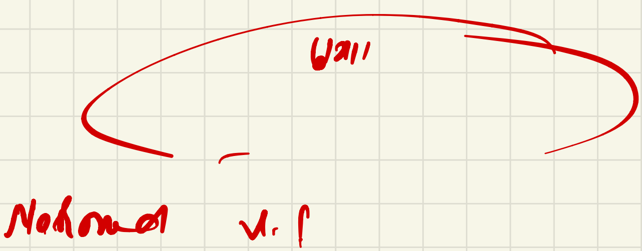
↳ homogeneity property

iii) Δ-inequality

$\|x + y\| \leq \|x\| + \|y\| \quad \forall\, x, y \in X$

- A vector space 'V' with a norm is called "Normed linear space" or "Normed vector space".

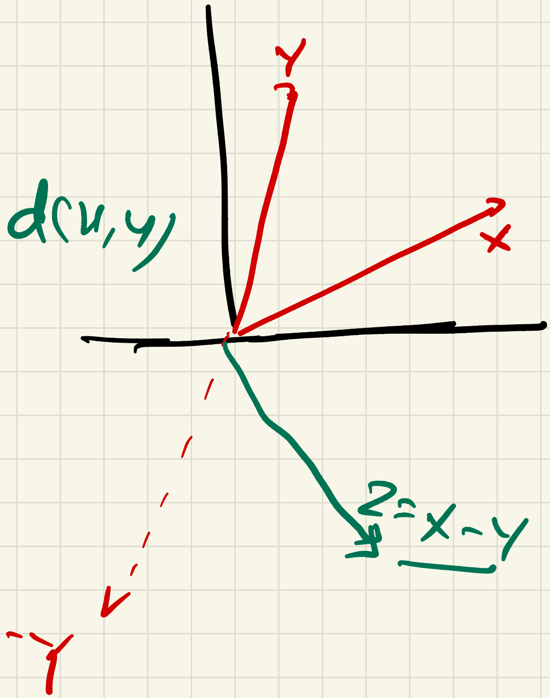




Metric Induced by norms

→ $x, y \in V$

$$\|z\| = \|x - y\| = d(x, y)$$



* Every norm on a vector space induces a metric

⇒ But all metrics are not induced by a norm.

$\Rightarrow$ Metrics induced by a norm should have translation invariance

$$d(x+a, y+a) = d(x,y)$$

*

$$\delta(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

$$\delta(x,y) = \{0, 1\}$$

$$\delta(ax, ay) \neq |a|\,\delta(x,y)$$

$\mathbb{R}$

$(\mathbb{R}, \|\cdot\|_1)$ $(\mathbb{R}, \|\cdot\|_2)$ $(\mathbb{R}, \|\cdot\|_\infty)$

* Norms, Lp. norms Lpq norm

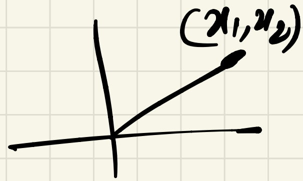
* Convexity of norm functions

* Convex sets generated norm inequalities

R^2

$x \in R^2$ $\quad$ $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\|x\|_1 = |x_1| + |x_2|$



$\|x\|_2 = (x_1^2 + x_2^2)^{1/2}$

$\|x\|_p = (|x_1|^p + |x_2|^p)^{1/p}$

$\vdots$

$\|x\|_\infty = \max(|x_1|, |x_2|)$.

$x \in R^n$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$\vdots$

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

Inner product space

$\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$

$x, y, z \in V$ and $a, b \in \mathbb{R}$

1) $\langle x, y \rangle = \langle y, x \rangle$

2) $\langle x, x \rangle = 0 \iff x = 0$

3) $\langle x, x \rangle \geq 0$

4) $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$

Linearity

$X, (X, d)$
$(X, \|\cdot\|)$

$\|\cdot\| \to d$

e.g., $x, y \in \mathcal{X} \subseteq R^n$

$x = (x_1, x_2, x_3, \ldots x_n)$

$y = (y_1, y_2, y_3, \ldots y_n)$

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T y$$

$x, y \in R^{m \times n}$

$$\langle x, y \rangle = \mathrm{tr}(x^T y)$$

$$\langle x, x\rangle = \left(\sum_{i=1}^{n} x_i^2\right)$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$

$$\|x\|_2 = \left(\langle x, x\rangle\right)^{1/2}$$

l_2- norm is induced by inner product

$X \rightarrow (X, d) \rightarrow (X, \|\cdot\|) \rightarrow \|x\| \rightarrow d$

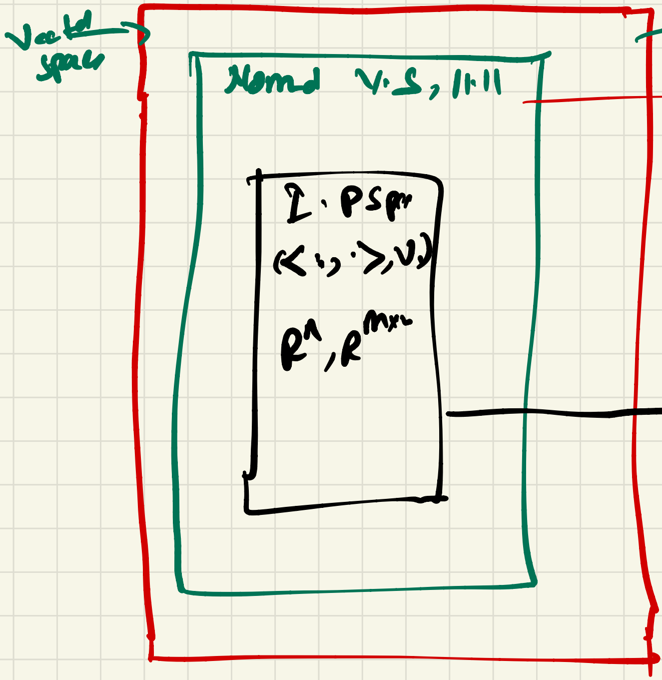
$\|\cdot\| \leftarrow \langle\cdot\rangle$

$$\|x\|_2 = \left(\sum x_i^2\right)^{1/2}$$

$$\langle\ \cdot\ \rangle \rightarrow \|\cdot\| \longrightarrow d$$

$x, y \in \mathbb{R}^p$

$$\langle x-y, x-y\rangle \rightarrow \langle x-y, x-y\rangle^{1/2}$$



$V(x_1 + a x_2)$

norm can be defined $\|\cdot\|$
$(X, \|\cdot\|)$

IP $x \in \mathbb{R}^n$
$\langle x, x\rangle = \|x\|^2$
$X \in \mathbb{R}^{m\times n}$
$X^T X = \operatorname{tr}(X^T X) = \|X\|_F$

e.g., RKHS

* Not all normed spaces are inner product spaces

* 'l_1 norm' cannot be induced by any inner product

* Norm

$x^{(i)}$

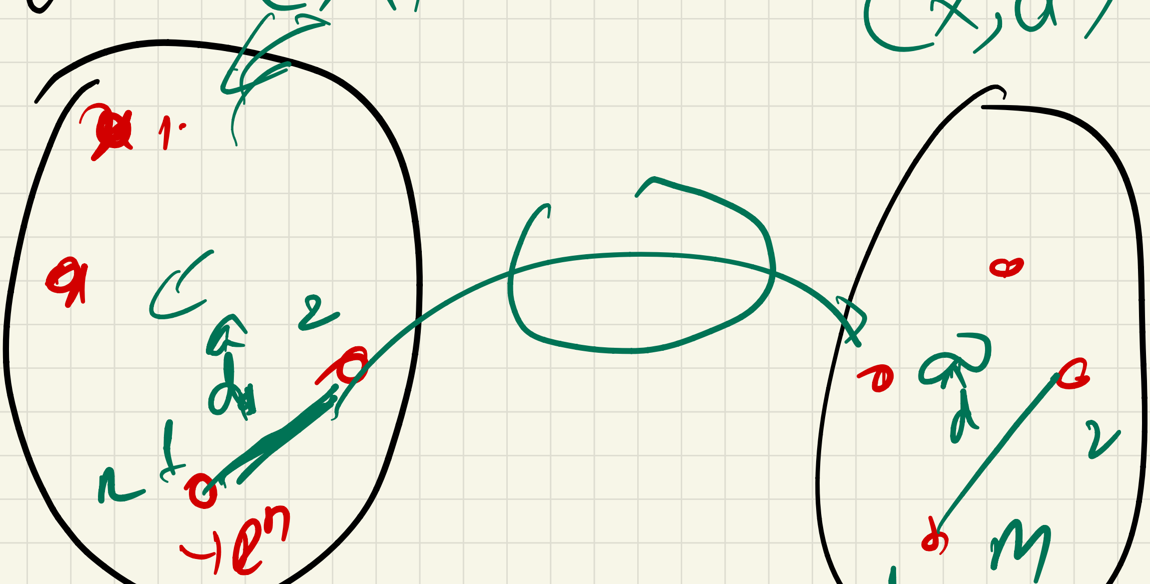
$x \in R^n \longrightarrow \tilde{x} \in R^m$

$n = 3$

$m << n$ $\quad n = 2$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \longrightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$(\hat{x}, \hat{d}) \quad (\tilde{x}, \tilde{d})$



$m < n$