14/12/2024

LA for Data Science, Mansoura, session 5

→ SVD (Singular Value Decomposition)

→ PCA (Principle Component Analysis)

↳ Dimensionality Reduction.

- Rank Reduction
  (Reduced Rank Matrix)

* Covariance Matrix ; (Correlation Matrix)

→ Image Compression (Lossy compression)

→ $Tr(A) \equiv Trace(A)$ $\quad 1+2j \quad 1-2j$ $\quad$ Review

$= \sum_{i=1}^{m} a_{ii}$

$$\begin{bmatrix} a_{11} & a_{12} & \\ a_{21} & a_{22} & \\ & & a_{33} \ddots \end{bmatrix}$$

$$\boxed{Tr(A) = \sum_i \lambda_i(A)} \rightarrow (1)$$

$\sum_i a_{ii} = a_{11} + a_{12} + a_{13}$

$$\boxed{det(A) = \prod_i \lambda_i(A)} \rightarrow (2)$$

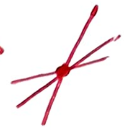
$\prod_i a_{ii} = a_{11} \times a_{22} \times a_{33}$

→ Revisiting "Solving Linear system of equations"

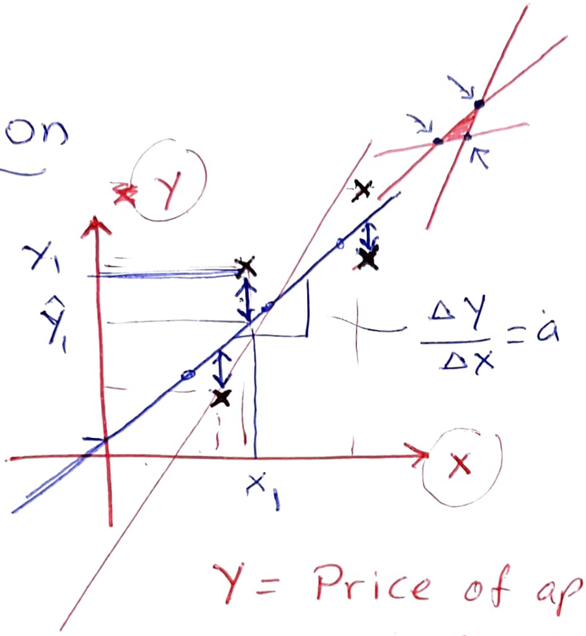
→ Overdetermined case;

$$a_1 x + b_1 y = c_1$$
$$a_2 x + b_2 y = c_2$$
$$a_3 x + b_3 y = c_3$$

determined →

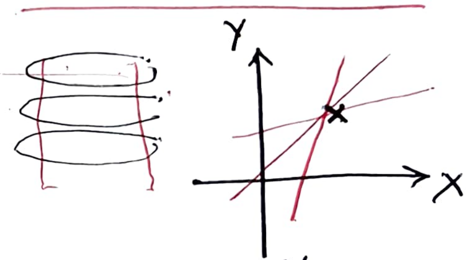


→ Regression



error$_i$ $= Y_1 - \hat{Y}_1$

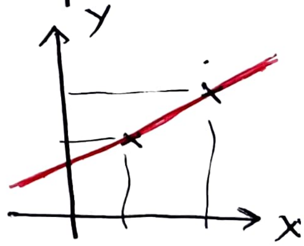
$\frac{\Delta Y}{\Delta X} = a$



Y = Price of apt.
x = area of apt.

$$\boxed{Y = ax + b}$$ — Intercept

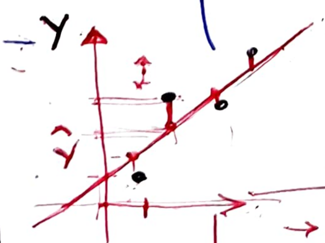
slobe   model parameters



3 data points

$$\begin{array}{cc} x_1 & \\ ① & \begin{bmatrix} 100 \\ 1000 \end{bmatrix} \\ Y_1 & \end{array} \quad \begin{array}{cc} x_2 & \\ ② & \begin{bmatrix} 120 \\ 1400 \end{bmatrix} \\ Y_2 & \end{array} \quad \begin{array}{cc} x_3 & \\ ③ & \begin{bmatrix} 130 \\ 1600 \end{bmatrix} \\ Y_3 & \end{array}$$

$$\begin{cases} 1000 = 100\,a + b \to 0 \\ 1400 = 120\,a + b \\ 1600 = 130\,a + b \end{cases}$$



$$\underbrace{\begin{bmatrix} 100 & 1 \\ 120 & 1 \\ 130 & 1 \end{bmatrix}}_{A_{3\times 2}} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{B}_{2\times 1}} = \underbrace{\begin{bmatrix} 1000 \\ 1400 \\ 1600 \end{bmatrix}}_{\vec{x}_{3\times 1}}$$

if $A_{2\times 2}$

$$\boxed{\vec{\beta} = A^{-1}_{n\times n}\vec{x} \quad a \checkmark,\ b \checkmark}$$

$$\underset{2\times3}{A^T}\;\underset{3\times2}{A}\;\vec{\beta}_{2\times1} = \underset{2\times3}{A^T}\,\vec{X}_{3\times1}$$

$$(A^TA)^{-1}(A^TA)_{2\times2}\;\vec{\beta}_{2\times1} = (A^TA)^{-1}A^T_{2\times3}\;\vec{X}_{3\times1}$$

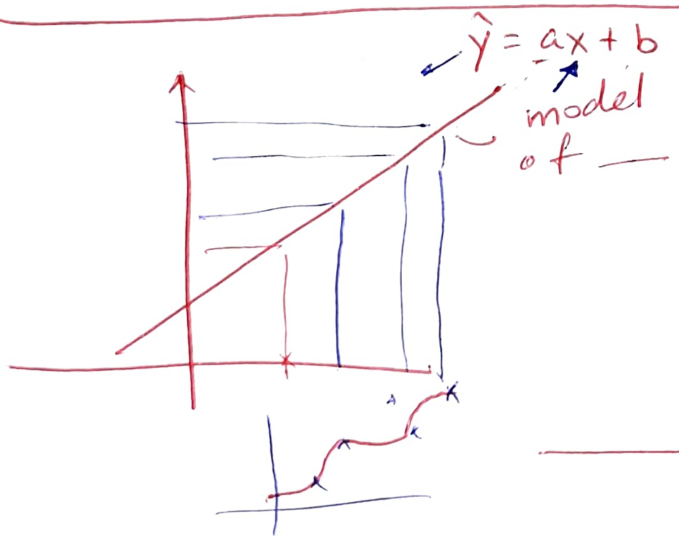
$$I_{2\times2}\;\vec{\beta}_{2\times1} = (A^TA)^{-1}A^T\,\vec{x}$$

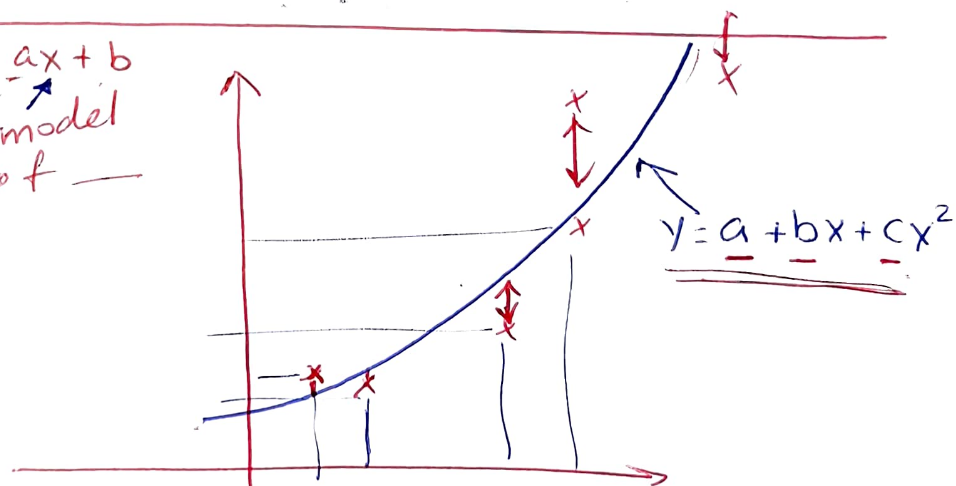
$$\vec{\beta} = \underline{(A^TA)^{-1}A^T}\;\vec{x}$$

Pseudo-inverse

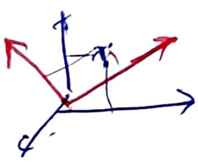
Least square solutions

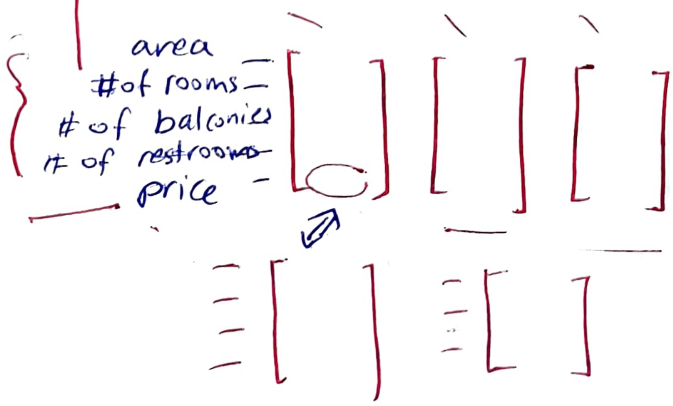
$$\boxed{\begin{aligned} A_{2\times2}\times\vec{\beta}_{2\times1} &= \vec{X}_{2\times1} \\ \vec{\beta} &= A^{-1}\,\vec{X}_{2\times1} \end{aligned}}$$





PCA ⇒





SVD

Singular Value Decomp.

$$\underset{3\times 4}{A_{m\times n}} = \boxed{U \Sigma V^T}$$

$$= U_{m\times m} \quad \Sigma_{m\times n} \quad V^T_{n\times n}$$

matrix of singular values

$s_1, s_2, \dots, s_k$

$\sigma_1, \sigma_2, \dots, \sigma_k$

$$\begin{bmatrix} s_1 & & & & \\ & s_2 & & & \\ & & s_3 & & \\ & & & s_4 & 0 \end{bmatrix}$$

$$\begin{bmatrix} s_1 & & \\ & s_2 & \\ & & s_3 \\ & & 0 \end{bmatrix}$$

U & V;

→ Columns of U are eigenvectors of AA^T

→ Columns of V are eigenvectors of A^TA

→ rows of V^T

eigen decomposition

"Diagonalization"

if

$A_{n\times n}$; Diagonalizable

$$A = \boxed{PDP^{-1}}$$

$$A_{n\times n} = X \Lambda X^{-1}$$

$$\begin{bmatrix} \uparrow & & \uparrow \\ v_1 & \cdots & v_n \\ \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} & & \\ & & \end{bmatrix}$$

$A \; B$

$\cancel{U_{m\times n}} \; \cancel{\Sigma_{n\times m}}$

"Σ";

→ singular values are square roots eigenvalues of A^TA or eigenvalues of AA^T

e.g.,
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} & \\ & \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

s_1, s_2

$$A\vec{x} = \vec{y}$$

$$(U \Sigma V^T)\vec{x} = \vec{y}$$

$\vec{\alpha}$

$\vec{\beta}$

$\vec{y}$

$$\Sigma = \begin{bmatrix} \ddots \end{bmatrix}$$

$$A\vec{x} = \vec{y}$$

$$P D P^{-1} \vec{x} = \vec{y}$$

$$D = \begin{bmatrix} \ddots \end{bmatrix}$$

→ Lossy compression

→ Rank Reduction (reduced rank matrix)

$$\begin{bmatrix} | & | & | \\ u_1 & u_2 & \cancel{u_3} \\ | & | & | \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \cancel{1} \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} - & V_1 & - \\ - & V_2 & - \\ - & V_3 & - \end{bmatrix}$$

SVD

128

128

256

A =

100 x 100

10,000

$$\underset{m\times n}{A} = \begin{bmatrix} | & | & \\ u_1 & u_2 & \cdots \\ | & | & \end{bmatrix} \underbrace{\begin{bmatrix} s_1 & & \\ & s_2 & \\ & & s_k \end{bmatrix}}_{D} \begin{bmatrix} - & V_1 & - \\ - & V_2 & - \\ & \vdots & \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

128 x 128

$$= s_1 \begin{bmatrix} | \\ u_1 \\ | \end{bmatrix} \begin{bmatrix} - & V_1 & - \end{bmatrix} + s_2 \begin{bmatrix} | \\ u_2 \\ | \end{bmatrix} \begin{bmatrix} - & V_2 & - \end{bmatrix} + \cdots + s_k \begin{bmatrix} | \\ u_k \\ | \end{bmatrix} \begin{bmatrix} - & V_k & - \end{bmatrix}$$

100

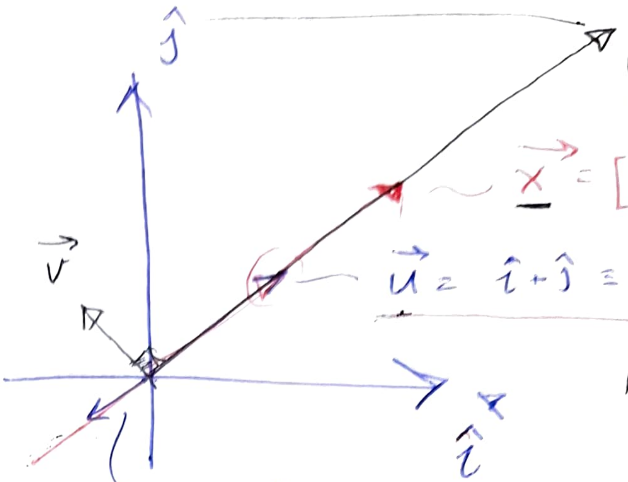
80

0.01

$$= s_1 \begin{bmatrix} \; \end{bmatrix}_{m\times n} + s_2 \begin{bmatrix} \; \end{bmatrix}_{m\times n} + \cdots + s_k \begin{bmatrix} \; \end{bmatrix}_{m\times n}$$

$$\vec{x} = a u_1 + b_1 u_2 + \cancel{c_3 u_3}$$

Review



$\vec{y} = \begin{bmatrix}5\\5\end{bmatrix}$

$\vec{x} = \begin{bmatrix}2.5\\2.5\end{bmatrix}$

$\vec{u} = \hat{i}+\hat{j} = \begin{bmatrix}1\\1\end{bmatrix}$

$\vec{z} = \begin{bmatrix}-.5\\-.5\end{bmatrix}$

$\vec{x} = 2.5\,\hat{i} + 2.5\,\hat{j}$

2 basis

$\vec{y} = 5\,\hat{i} + 5\,\hat{j}$

$\vec{z} = -0.5\,\hat{i} + -0.5\,\hat{j}$

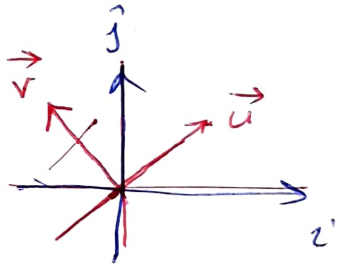
$$\boxed{\begin{aligned}\vec{x} &= 2.5\,\vec{u} \\ \vec{y} &= 5\,\vec{u} \\ \vec{z} &= -0.5\,\vec{u}\end{aligned}}\quad\begin{aligned}&\cancel{+0\vec{v}}\\&\cancel{+0\vec{v}}\\&\cancel{+0\vec{v}}\end{aligned}$$

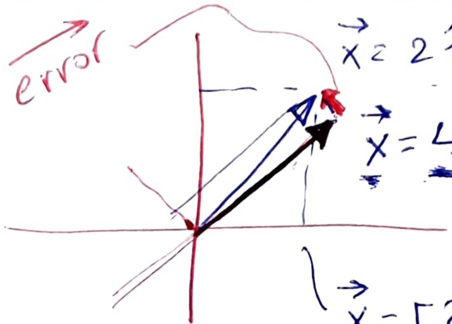
new basis $\vec{u}$, $\vec{v}$

$\vec{x} = \begin{bmatrix}2.5\\0\end{bmatrix}$

$\vec{y} = \begin{bmatrix}5\\0\end{bmatrix}$

$\vec{z} = \begin{bmatrix}-0.5\\0\end{bmatrix}$



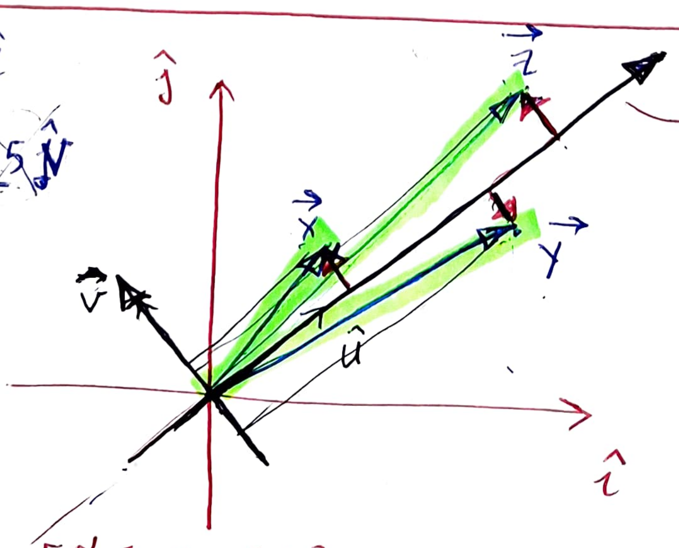


$\vec{x} = 2\hat{i} + 3\hat{j}$

$\vec{x} = 4\,\hat{u} + 0.5\,\hat{v}$

$\vec{x} = \begin{bmatrix}2\hat{i}\\3\hat{j}\end{bmatrix}$

$\vec{x} \approx 4\,\hat{u}$

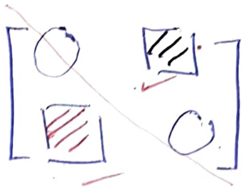


$\begin{bmatrix}\alpha\\\beta\end{bmatrix} \Rightarrow [\alpha]$

Principle components.

$$B = (AA^T) = \begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix}\begin{bmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{bmatrix}$$

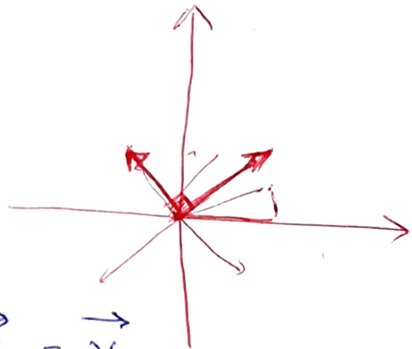
$$B = B^T$$



→ AA^T is symmetric

→ eigenvalues of a symmetric matrix are real numbers.

→ eigenvectors of a sym. matrix~~es~~ are orthogonal.



→ S.V.D $A\vec{x} = \vec{y}$

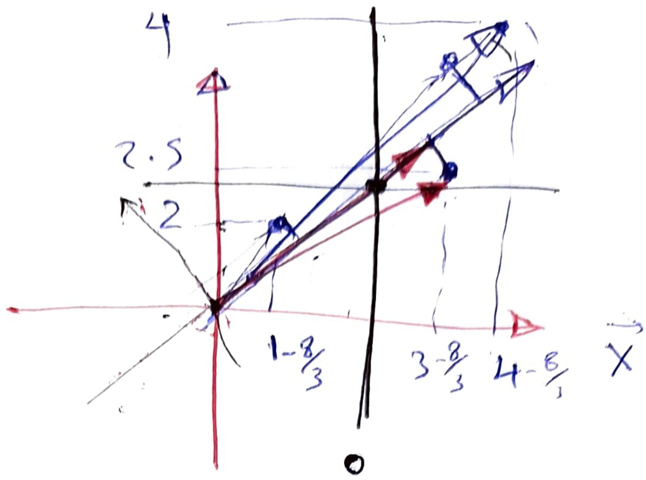
$$U \Sigma V^T \vec{x} = \vec{y}$$

Standardization → shifting → zero mean

→ scaling (variability; standard deviation = 1)

to be discussed in detail in "Prob. & Stat. sessions" ← σ ←

Covariance matrix



$\begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_k \end{bmatrix}$

$\mu_x = \frac{8}{3} = \frac{1+3+4}{3}$

$\underline{SV}\left(\begin{bmatrix} x_1 - \mu_x & y_1 - \mu_y \\ x_2 - \mu_x & y_2 - \mu_y \\ x_3 - \mu_x & y_3 - \mu_y \end{bmatrix} \right)$

s_1 ✓

$\cancel{s_2}$ ↓

prob.

$\underline{eig} = \left(\begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix} \right)$ — covariance matrix

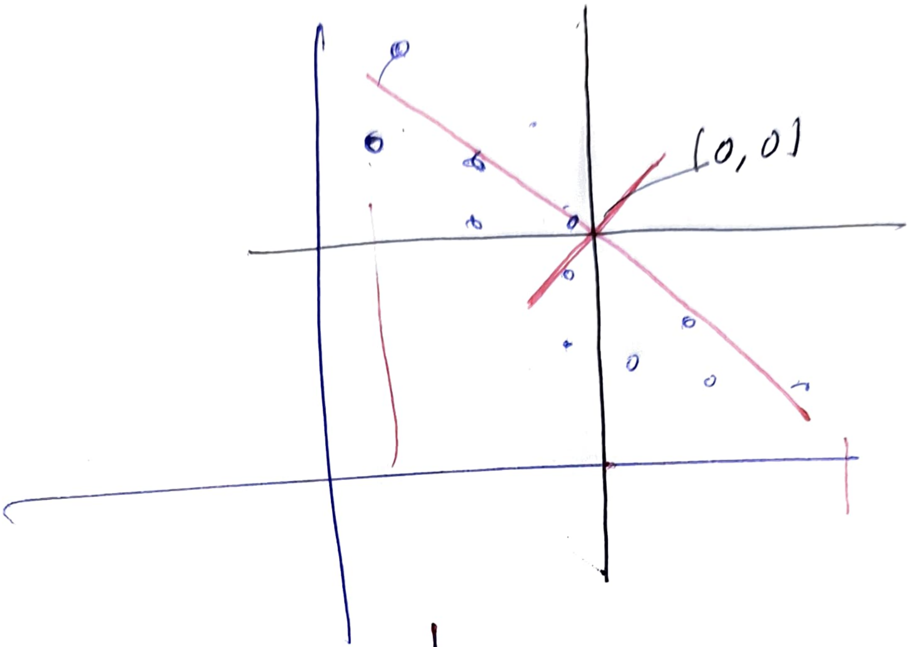
$\begin{bmatrix} \lambda_1 & \\ & \cancel{\lambda_2} \end{bmatrix}$

PCA

eigen decomposition

$\begin{bmatrix} s_1 & & & & \\ & s_2 & & & \\ & & s_3 & & \\ & & & \cancel{s_4} & \\ & & & & \cancel{s_5} \end{bmatrix} \begin{bmatrix} x_{11} \\ \vdots \\ x_{15} \end{bmatrix} \approx \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$

$k \ll n$


(0,0)