Week 3

Univariate	Multivariate

Central tendency
- Mean → Mean Vector
- Median
- Mode

Dispersion
- Range
- IQR
- Std Deviation (variance) → Covariance matrix
  - → Correlation matrix

Simultaneous Occurrence

X_1 = i^{th} Multivariate observation

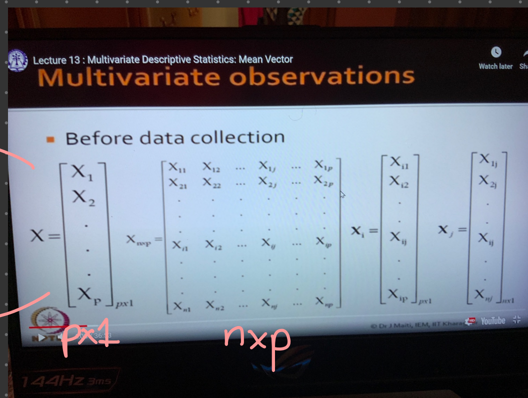
$p \times 1$

Rows are no of observations

$i = 1 \quad n$

Columns are number of variables

$j = 1 \quad p$



px1 nxp

x_{ij} = i^{th} observation on j^{th} variable to be collected

→ Random Observation

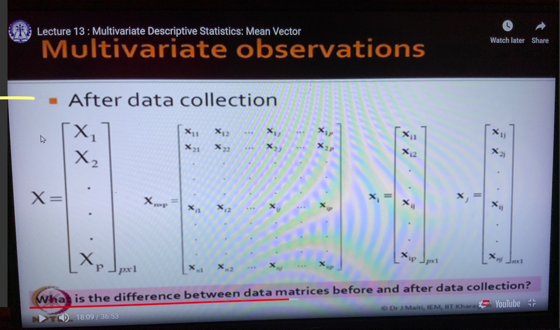
→ Random Variable

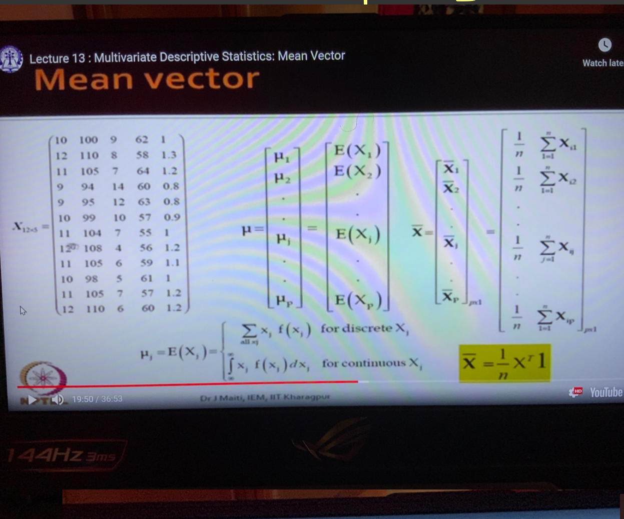
$$\boxed{\begin{array}{l} x \longrightarrow \\ E(x) \longrightarrow \mu \end{array}}$$

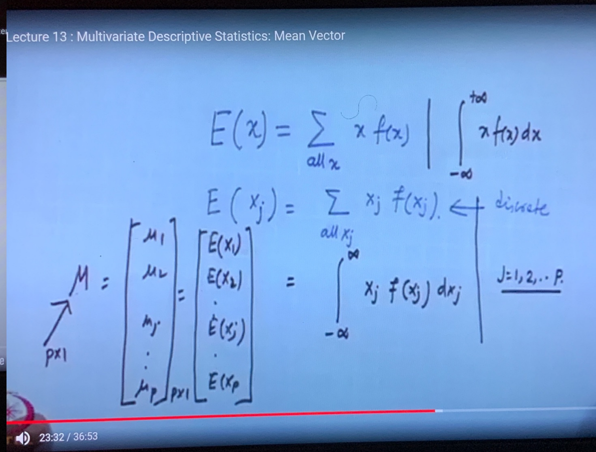
$$x_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{pj} \end{bmatrix}_{n \times 1}$$

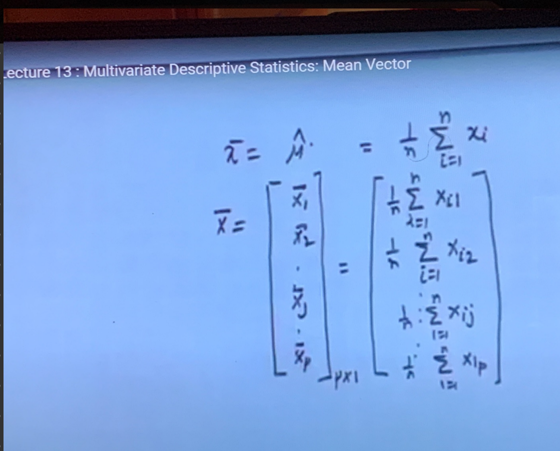
D

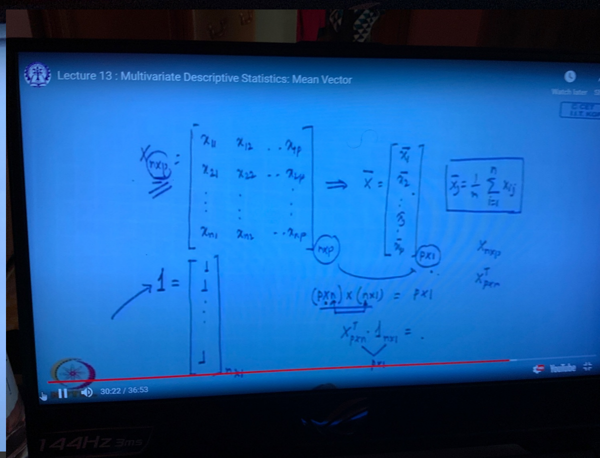
Looks like fixed value unlike the previous case











Covariance Matrix [lec 2]

From population point of view,

Population Covariance Matrix

$$x \to V(x) = E(x-\mu)^2 = \sum (x-\mu)^2 f(x)$$

$$= \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx = \sigma^2$$

$$\sigma_j^2 = V(x_1) = E[X_j - \mu_j]^2 = \sum_{\text{all } x_1} (x_1 - \mu_j)^2 f(x_1) \leftarrow \text{Dis}$$

Continuous ✓

$$= \int_{-\infty}^{\infty} (x_j - \mu_j)^2 f(x_j)\,dx_j$$

For $j = 1 \to \sigma_1^2$

$j = 2 \to \sigma_2^2$

$\vdots$

$j = p = \sigma_p^2$

If x_1, x_2 are not independent,

When x_1 varies, x_2 covaries

$$V(x_j) = E(x_j - \mu_j)^2$$

$$\Downarrow$$

$$\sigma_{jk} = Cov(X_j, X_k) = E(X_j - \mu_j)(x_k - \mu_k)$$

$$= \sum_{\text{all } x_j, x_k} (x_j - \mu_j)(x_k - \mu_k)\, f(x_j, x_k)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_j-\mu_j)\,(x_k-\mu_k)\,f(x_j, x_k)\,dx_j\,dx_k$$

$\sigma_j^2 = \sigma_{jj} \rightarrow$ Variance of x_j

σ_{jk}^2 : Covariance between x_j, x_k

Population Covariance Matrix

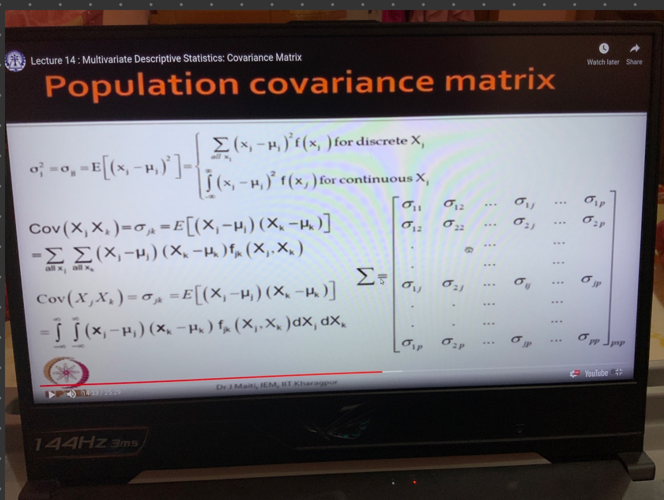
$$\Sigma = \begin{array}{c|c} & x_1 \quad x_2 \quad \cdots \quad x_p \\ \hline \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{array} & \begin{bmatrix} \sigma_{11} & & & \\ & \sigma_{22} & & \\ & & \sigma_{33} & \\ & & & \ddots \\ x_{1p} & x_{2p} & & \sigma_{pp} \end{bmatrix}_{p\times p} \end{array}$$

Diagonal : Describe variance

Rest : Describe Covariance

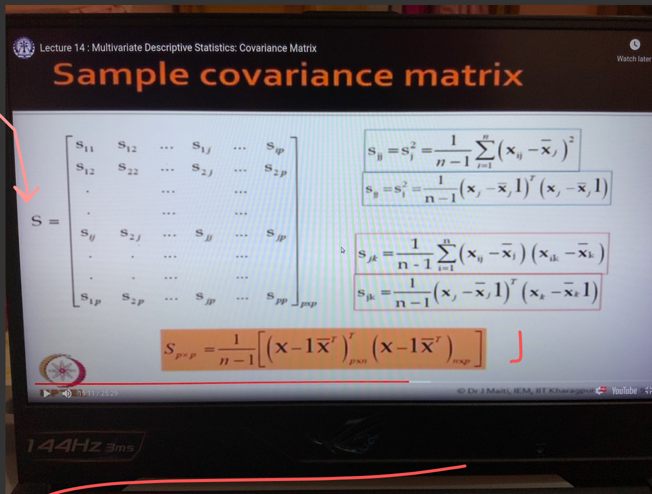
$\sigma_{jk} = \sigma_{kj}$

$\Sigma \downarrow$ population covariance matrix



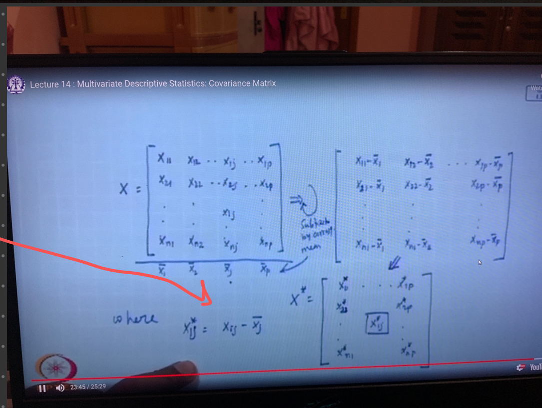
Sample Covariance Matrix:

Size: $p \times p$



→ Mentors

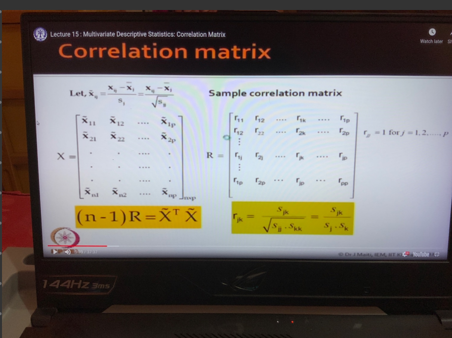
→ Ways to appro

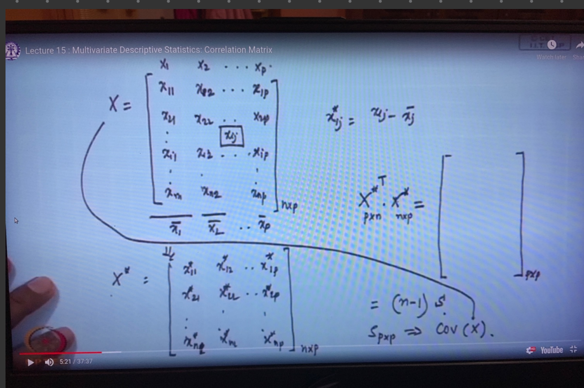


Can be modified and written as

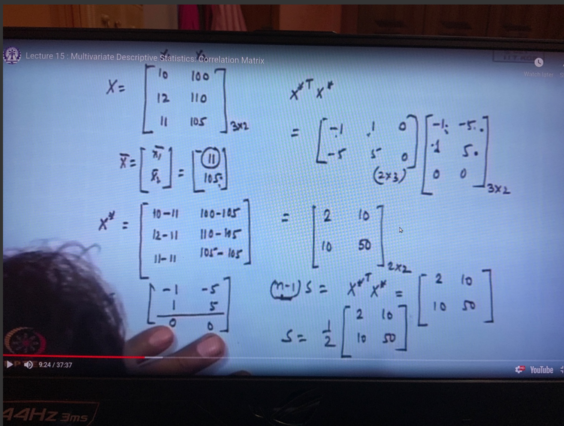
$$S_{ij} = \frac{1}{n-1} \sum_{2}^{n} X^*_{ij} \, X^*_{iR}$$

Lec 3

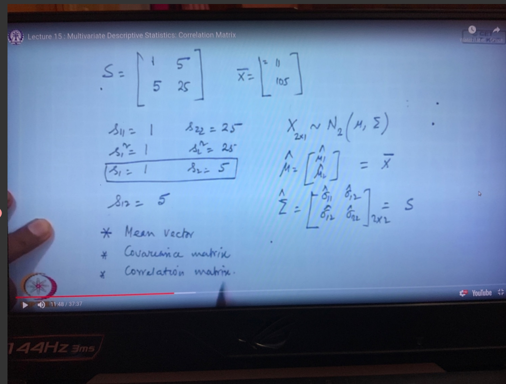




Example:



$\Longrightarrow$



Correlation Matrix

$$\rho_{p\times p} = \begin{bmatrix} 1 & - & - & - & \rho_{1p} \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \\ \rho_{1p} & & & & 1 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_{11} & - & \cdots & \sigma_{1p} \\ & \sigma_{22} & & \vdots \\ \sigma_{1p} & - & - & \sigma_{pp} \end{bmatrix}$$

Covariance

Population Correlation Matrix

$$\text{Correlation}(x_j, x_k) = \frac{\text{Covariance}(x_j, x_k)}{\text{Std}(x_j, x_k)}$$

$$\text{Correlation}(x_j, x_k) = \frac{\text{Cov}(x_j, x_k)}{\sigma_j \cdot \sigma_k}$$

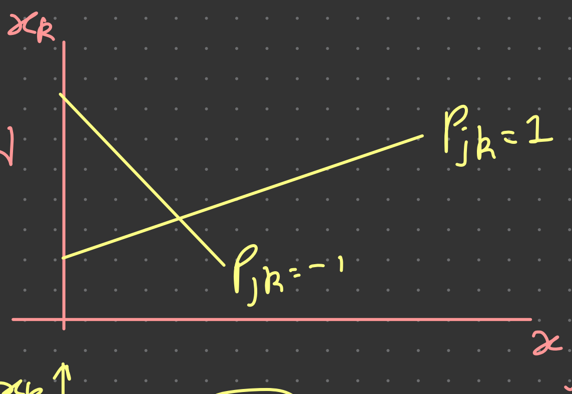
$\rho_{jk} = \frac{\sigma_{jk}}{\sigma_j \sigma_k} \Leftarrow$ For $j=k$,

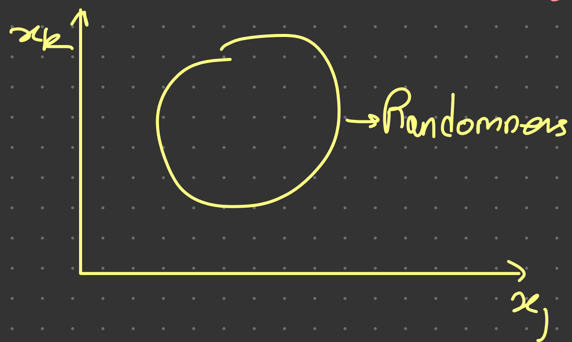
$$\rho_{jj} = \frac{\sigma_{jj}}{\sigma_j \sigma_j} = 1$$

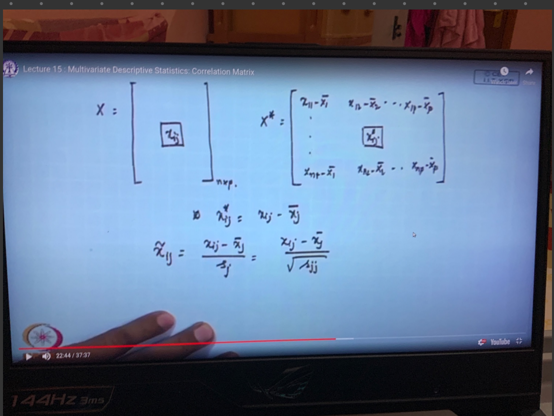
$\rho_{jk}=1 \rightarrow$ Positively Corelated

$\rho_{jk}=-1 \rightarrow$ Negatively corelated

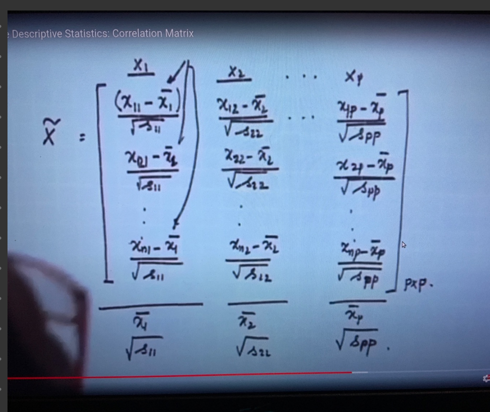
$\rho_{jk}=0 \rightarrow$ No correlation



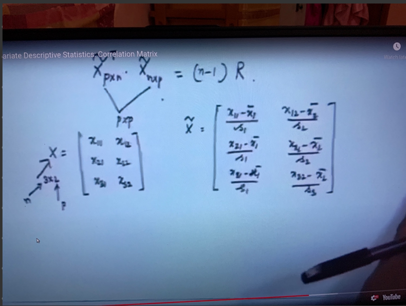


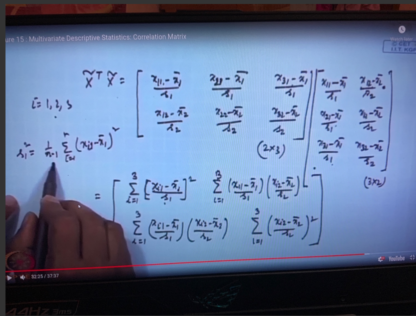


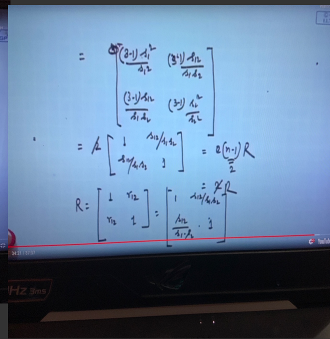
$\Rightarrow$



$$\underbrace{\tilde{X}^T_{p\times n}\cdot\tilde{X}_{n\times p}}_{p\times p} = (n-1)R$$







$$\boxed{r_{ij} = \frac{s_{ij}}{s_i s_j}} \longrightarrow r_{11} = 1$$

Lecture- 4

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ \vdots & & & \vdots \\ \vdots & & \_\_ & \vdots \end{bmatrix} \qquad R = \begin{bmatrix} 1 & r_{12} & - & r_{1p} \\ & & & \\ r_{1p} & - & - & 1 \end{bmatrix}$$

$$D_s = \begin{bmatrix} s_{11} & 0 & 0 & 0 \\ \vdots & s_{22} & & 0 \\ \vdots & & & \\ 0 & - & & s_{pp} \end{bmatrix}_{p\times p}$$

$$\boxed{R = D_s^{-1/2} S D^{-1/2} \qquad S = D_s^{1/2} R D_s^{1/2}}$$

Sum Square & Cross Product Matrix : [SSCP]

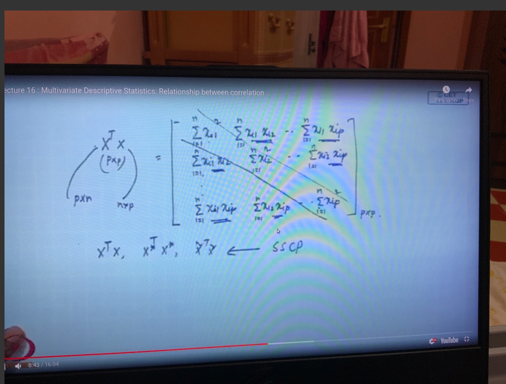
$$(n-1)S = X^{*T}X^{*}$$

$$(n-1)R = \tilde{X}^{T}\tilde{X}$$

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}$$

$$X^{T}X = \begin{bmatrix} \sum_{i=1}^{3} x_{11}^{2} & \sum_{i=1}^{3} x_{12}x_{12} \\ \sum_{i=1}^{3} x_{11}x_{12} & \sum_{i=1}^{3} x_{12}^{2} \end{bmatrix}_{p\times p}$$

Can be expended for pxp



Sum products's for variance

Cross products for covariance