LC #1 : No work   30-Jan-2025

LC #2 :   4-Feb-2025

Multi Variate Analysis

پہلے جو پڑھا وہ uni variate تھا۔

1 Variable

(mean, median, mode, sd, variation)

Bivariate → 2 Variable

↳ correlation, regression.

No of observations (n) → object

→ Daily life examples (5) → Applications of multivariate Analysis in IT. (Search) } Assignment

matrix of one column → Vectors → uni variate

Examples: Educational system, Environmental sciences.

Definition:

It deals with the study of variables determined on more than one characteristics.

The no's of person under consideration is called Object.

Example:

Personality

Eye Color, height, Dressing sense → variable

Let's consider the example of multi Variate Analysis ~~1 variable~~ 'Personality' having the characteristics like eye color, height and dressing sense. The number of persons considered are the objects.

'Applications In Different Fields'

1. Marketing & market Research

Custom Profiling
- Demographics → e.g (Age, Gender, education)
- Geographic locations (The people who approach us)
- Brand loyality ([illegible])
- Product Preferences (کتنی بار product use کرنا پسند کرے گا)
  ↳ (finding alternatives)

(2) Education :- → learning Analysis
- → Acedemic Performance
- → class environment
- → Socio Economics stats
- → parental environment.

Organisation Of MultiVariate Data : Matrices & Vectors

* Multivariate is represented in rectangular form.

$$\underbrace{\text{rectangular}}_{p \;\; n}$$

p: no of variables — n: no. of observation (objects)

Vectors: matrix having 1 column.

Univariate :

- ↳ single variable
- ↳ no relationship study
- ↳ Mean, mode, median, σ^2, σ
- ↳ Bar plot, Histogram, frequency table

e.g Suppose a univariate is

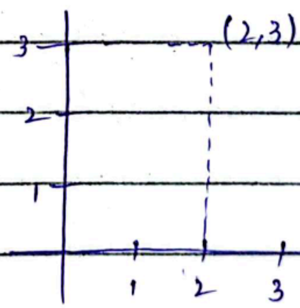
X
x_1
x_2
⋮
x_n

then it's vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$

n → no. of object
1 → no of col → variable

Bivariate : Two Variables involved

- ↳ Scatter plot
- ↳ line plot
- ↳ Relationship study
- ↳ Regression, Correlation



Multivariate :

↳ multiple Variables Involved.

1 Observation ki 1st entry

1 Observation ki p^{th} entry

object — variable

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & & \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

variables

$\therefore \; n > p$

no. of variables (object)

LC #3 : 6-Feb-2025

Vector: Matrix with one column is a vector.

It is represented by $\vec{x}$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ Vertically.}$$

If we want to write it as horizontally.

$$\vec{x}' = \vec{x}^T = \left[x_1 \ x_2 \ x_3\right]^T_{1 \times n}$$

1 — Variable

n — objects (Observations)

Determinant of 2X2

e.g $A = \begin{bmatrix} 6 & 2 \\ 1 & 9 \end{bmatrix}$

$\det A = |A| = (6)(9) - (2)(1) = 8$

Determinant of 3X3

$$A = \begin{array}{ccccc} \boxed{1} & \boxed{5} & \boxed{9} & \boxed{1} & \boxed{5} \\ 2 & 0 & 1 & 2 & 0 \\ 3 & 8 & 6 & 3 & 8 \\ \uparrow & \uparrow & & \text{Repeat} & \text{Repeat} \\ C_1 & C_2 & & C_1 & C_2 \end{array}$$

∴ Determinant can be Determined only for **Square** matrix

$= (0 + 15 + 144) - (60 + 8 + 0)$

Properties Of Determinant :

1- Determinant only exist for square matrices.

Square matrix ⇒ no. of rows = no. of Col

2- ~~Diagonal~~ matrix

Determinant of Diagonal matrix = product of Diagonal matrix.

e.g
$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 5 \end{bmatrix}$$

$|A| = 2 \times 4 \times 5 = 40$

3- If element of single row or column of a matrix are multiplied by the Scaler (K), The determinant of new matrix $= K|A|$

e.g $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$ $|A| = 2-0 = 2$ Now multiply R_1 with 2

$B = \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}$ $|A| = 4-0 = 4$ 2×2

$2|A|$

Scale Power (multiplied rows)

Scaler $|A|$

If every element of a matrix $(n \times n)$ is multiplied by scaler c then $|cA| = c^n |A|$

4- If Two rows or Two columns are interchanged, signs of determinant is reversed.

e.g $\begin{vmatrix} 2 & 4 \\ 1 & 6 \end{vmatrix} = 8$

$C_1 \longrightarrow C_2$

$\begin{vmatrix} 4 & 2 \\ 6 & 1 \end{vmatrix} = -8$

5- If any whole row or column is 0 then the determinant of that matrix will be zero.

6- If two row or two column are exactly same to each other, then the determinant of that matrix will be zero.

7- If zero matrix / null matrix is given the determinant will be zero.

8- $|AB| = |A||B|$

Mean Vector and Variance and Covariance:

Summary Statistics and Descriptive statistics:

are measures that describe the data in multivariate analysis the most commonly used measures are mean, variance, covariance & correlation

Q: The dataset with 3 variables (X_1 = expenditures, X_2 = expenses X_3 = house hold income) and five observations are:

X_1	X_2	X_3
1.2	3.4	5.6
2.3	4.5	6.7
3.4	5.6	7.8
4.5	6.7	8.9
5.6	7.8	9.0

Calculate Mean Vector & variance - Covariance matrix.

Mean :

Mean of x_1 = $\frac{1.2 + 2.3 + 3.4 + 4.5 + 5.6}{5} = 3.4$

Mean of x_2 = $\frac{3.4 + 4.5 + 5.6 + 6.7 + 7.8}{5} = 5.6$

Mean of x_3 = $\frac{5.6 + 6.7 + 7.8 + 8.9 + 9.0}{5} = 7.6$

Mean Vector = $\vec{x} = \begin{bmatrix} 3.4 \\ 5.6 \\ 7.6 \end{bmatrix}$ → If we want to write it horizontally it will be written as:-

Mean Vector = $\vec{x}^T = \begin{bmatrix} 3.4 & 5.6 & 7.6 \end{bmatrix}^T$

$\overrightarrow{x_i} = \sum_{i=1}^{n} \frac{x_i}{n}$ (General form)

∴ To compute Variance and Covariance matrix, we need to calculate variance of each variable and covariance b/w pairs of variables.

Variance :

Variance of x_1 = $\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x_1})^2$

Variance of x_1 = $\frac{(1.2-3.4)^2 + (2.3-3.4)^2 + (3.4-3.4)^2 + (4.5-3.4)^2 + (5.6-3.4)^2}{4}$

$= \frac{4.84 + 1.21 + 0 + 1.21 + 4.84}{4} = 3.025$

Variance of x_2 = $\frac{(3.4-5.6)^2 + (4.5-5.6)^2 + (5.6-5.6)^2 + (6.7-5.6)^2 + (7.8-5.6)^2}{4}$

$= \frac{4.84 + 1.21 + 0 + 1.21 + 4.84}{4} = 3.025$

$$\text{Variance of } x_3 = \frac{(5.6-7.6)^2+(6.7-7.6)^2+(7.8-7.6)^2+(8.9-7.6)^2+(9.0-7.6)^2}{4}$$

$$= \frac{4+0.81+0.04+1.69+1.96}{4} = 2.125$$

Co-variance:

Between X_1 and X_2 :-

$$= \frac{(1.2-3.4)(3.4-5.6)+(2.3-3.4)(4.5-5.6)+(3.4-3.4)(5.6-5.6)+(4.5-3.4)(6.7-5.6)+(5.6-3.4)(7.8-5.6)}{4}$$

$$= \frac{(-2.2)(-2.2)+(-1.1)(-1.1)+0+(0.6)(1.1)+(2.2)(2.2)}{4} = 2.08875$$

Between X_2 and X_3 :-

$$= \frac{(3.4-5.6)(5.6-7.6)+(4.5-5.6)(6.7-7.6)+(5.6-5.6)(7.8-7.6)+(6.7-5.6)(8.9-7.6)+(7.8-5.6)(9-7.6)}{4}$$

$$= (-2.2)(-2)+(-1.1)(-0.9)+0+(1.1)(1.3)+(2.2)(1.4)/4 = 2.475$$

Between X_1 and X_3 :-

$$= \frac{(1.2-3.4)(5.6-7.6)+(2.3-3.4)(6.7-7.6)+(3.4-3.4)(7.8-7.6)+(4.5-3.4)(8.9-7.6)+(5.6-3.4)(9-7.6)}{4}$$

$$= (2.2)(-2)+(-1.1)(-0.9)+0+(1.1)(1.3)+(2.2)(1.4)/4 = 2.475$$

$$\Sigma = \begin{bmatrix} \text{Variance of } x_1 & \text{covariance of } x_1x_2 & \text{covariance of } x_1x_3 \\ \text{Covariance of } x_2x_1 & \sigma^2 \text{ of } x_2 & \text{Covariance of } x_2x_3 \\ \text{Covariance of } x_3x_1 & \text{Covariance } x_3x_2 & \sigma^2 \text{ of } x_3 \end{bmatrix}$$

Variance Covariance matrix is:

$$\begin{bmatrix} 3.4 & 2.8875 & 2.475 \\ 2.8875 & 5.6 & 2.475 \\ 2.475 & 2.475 & 7.6 \end{bmatrix}$$

LC # 4 01-Feb-2025

Find mean vector, covariance matrix and correlation matrix for two random variables x_1 and x_2 while joint prob func $f(x_1, x_2)$ is given by: $\therefore E(x) = \mu$ mean ↲ exclude

x_1/x_2	0	1	$f(x_1)$	$x_1 f(x_1)$	$x_2 f(x_2)$
-1	0.24	0.06	0.30	-0.30	0
0	0.16	0.14	0.30	0	0.2
+1	0.40	0.00	0.40	0.40	
$F(x_2)$	0.80	0.2	1	$\Sigma x_1 f(x_1) =$ 0.10	$\Sigma x_2 f(x_2)$ 0.2

Summation of $f(x_1)$ and $f(x_2)$

Mean Vector $\vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0.10 \\ 0.20 \end{bmatrix}$

$x_1 \rightarrow$ column

$x_2 \rightarrow$ row

$\mu_1 = \Sigma x_1 f(x_1) = 0.10$

$\mu_2 = \Sigma x_2 f(x_2) = 0.20$

Variance-Covariance Matrix:-

$$\vec{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

$\sigma_{11} = \cancel{E(x_1)^2 - [E(x_1)]^2 = 0.69}$

$= \Sigma x_1^2 f(x_1) - \mu_1^2$

$= [(-1)^2(0.30) + (0)^2(0.30) + (1)^2(0.40)] - (0.10)^2 = 0.69$

$$\sigma_{22} = E(x_2)^2 - [E(x_2)]^2$$

$$= \sum x_2^2 f(x_2) - \mu_2^2 = 0.16$$

$$\sigma_{12} = E(x_1, x_2) - E(x_1)E(x_2)$$

$$= \sum_{j=1}^{n} x_1 x_2 P(x_1, x_2) - \mu_1 \mu_2$$

$$= [(-1)(0)(0.24) + (0)(0)(0.16) + (1)(0)(0.40) + (-1)(1)(0.06) + (0)(1)(0.14) + \cancel{(-1)(1)(0)}$$
$$(+1)(1)(0)] - (0.1)(0.2)$$

$$= -0.080$$

$$\vec{\Sigma} = \begin{bmatrix} 0.69 & -0.080 \\ -0.080 & 0.16 \end{bmatrix}$$

LC #5 Continue: 13-Feb-2025

Correlation Matrix:

$$\vec{P} = D^{-1} \vec{\Sigma} D^{-1}$$

D = Diagonal matrix

$$D = \mathrm{diag} \sqrt{\sigma_{ii}}$$

So,

$$\vec{D} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 \\ 0 & \sqrt{\sigma_{22}} \end{bmatrix}$$

* Variance $\longrightarrow$ represents in diagonal

* Inverse of Diagonal matrix = $\frac{1}{Num}$

$$\vec{D} = \begin{bmatrix} \sqrt{0.69} & 0 \\ 0 & \sqrt{0.16} \end{bmatrix}$$

$$\vec{D} = \begin{bmatrix} 0.8307 & 0 \\ 0 & 0.4 \end{bmatrix}$$

$$D^{-1} = \begin{bmatrix} \frac{1}{0.8307} & 0 \\ 0 & \frac{1}{0.4} \end{bmatrix}$$

D کے اوپر vector لگاؤ یا نہیں تو کوئی مسئلہ نہیں-

$$D^{-1} = \begin{bmatrix} 1.2038 & 0 \\ 0 & 2.5 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 1.2038 & 0 \\ 0 & 2.5 \end{bmatrix} \begin{bmatrix} 0.69 & -0.08 \\ -0.08 & 0.16 \end{bmatrix} \begin{bmatrix} 1.2038 & 0 \\ 0 & 2.5 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} (1.2038)(0.69)+0(-0.08) & 1.2038(-0.08)+0(0.16) \\ 0\ (0.69)+2.5(-0.08) & 0\ (-0.08)+2.5(0.16) \end{bmatrix} \begin{bmatrix} 1.2038 & 0 \\ 0 & 2.5 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 0.8306 & -0.0963 \\ -0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 1.2038 & 0 \\ 0 & 2.5 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 0.9997+0 & 0+(-0.24075) \\ \begin{matrix} -0.24076 \\ \cancel{-0.04076}+0 \end{matrix} & 0+1 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 0.9997 & -0.24075 \\ \begin{matrix} -0.24076 \\ \cancel{-0.04076} \end{matrix} & 1 \end{bmatrix}$$

Question: from $\vec{\Sigma} = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix}$ obtain $\vec{P}$

$$\vec{P} = \vec{D}^{-1} \vec{\Sigma} \vec{D}^{-1}$$

$$D = \text{diag}\sqrt{\sigma_{ii}}$$

$$D = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & 0 \\ 0 & \sqrt{\sigma_{22}} & 0 \\ 0 & 0 & \sqrt{\sigma_{33}} \end{bmatrix}$$

upto [illegible] mal

$$D = \begin{bmatrix} \sqrt{4} & 0 & 0 \\ 0 & \sqrt{9} & 0 \\ 0 & 0 & \sqrt{25} \end{bmatrix}$$

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$D^{-1} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.2 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 2+0+0 & 0.5+0+0 & 1+0+0 \\ 0+0.3+0 & 0+2.7+0 & 0+\cancel{0.6}-0.9+0 \\ 0+0+0.4 & 0+0+-0.6 & 0+0+5 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 2 & 0.5 & 1 \\ 0.3 & 2.7 & -0.9 \\ 0.4 & -0.6 & 5 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 1+0+0 & 0+0.15+0 & 0+0+0.2 \\ 0.15+0+0 & 0+0.81+0 & 0+0+-0.18 \\ 0.2+0+0 & 0+-0.18+0 & 0+0+1 \end{bmatrix}$$

$$\vec{P} = \begin{bmatrix} 1 & 0.15 & 0.2 \\ 0.15 & 0.81 & -0.18 \\ 0.2 & -0.18 & 1 \end{bmatrix}$$

Quiz on ~~[illegible]~~