Akshay
Lecture 3

Date:
Page No.

if $\rho_{12} = 0$ i.e they are independent.

$f(x_1, x_2) = f(x_1) \cdot f(x_2)$

Properties

Prop 1

$\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$

$a' = [a_1, a_2, \ldots a_p]$

$a'\underset{\sim}{x} = a_1x_1 + a_2x_2 \ldots + a_px_p$

$a'\underset{\sim}{x} \sim N(a'\underset{\sim}{\mu}, a'\Sigma a)$

⇒ univariate

if $a' = [1, 0, \ldots 0]$

$a'\underset{\sim}{x} = x_1 \sim N(\mu_1, \sigma_{11})$

Prop 2

$$A = \begin{bmatrix} a_{11} & a_{21} & & a_{q1} \\ a_{12} & & & \\ a_{1p} & & & a_{qp} \end{bmatrix}_{p \times q}$$

$$A'\underset{\sim}{X} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots a_{1p}x_p \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2p}x_p \\ \\ a_{q1}x_1 + a_{q2}x_2 + \ldots + a_{qp}x_p \end{bmatrix}_{q \times 1}$$

$$A'\underset{\sim}{X} \sim N_q(A'\mu, A'\Sigma A)$$

$$\underset{\sim}{X} + \underset{\sim}{d}_{p\times 1} \sim N_p(\underset{\sim}{\mu} + \underset{\sim}{d}, \Sigma)$$

property 3

All subsets of components of X have a (multivariate) normal distribution.

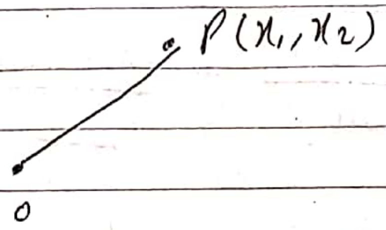
$$\underset{\sim}{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} X_{(1)\, q\times 1} \\ \hline X_{(2)\, (p-q)\times 1} \end{bmatrix}, \quad \underset{\sim}{\mu}_{p\times 1} = \begin{bmatrix} \mu_{(1)\, p\times 1} \\ \hline \mu_{(2)\, (p-q)\times 1} \end{bmatrix}$$

$$\Sigma_{p\times p} = \left[\begin{array}{c|c} \Sigma_{11\, q\times q} & \Sigma_{12\, q\times p-q} \\ \hline \Sigma_{21\, (p-q)\times q} & \Sigma_{22\, (p-q)\times(p-q)} \end{array}\right]$$

$$X_{(1)} \sim N_q(\mu_{(1)}, \Sigma_{11})$$

Date:
Page No.

Distance

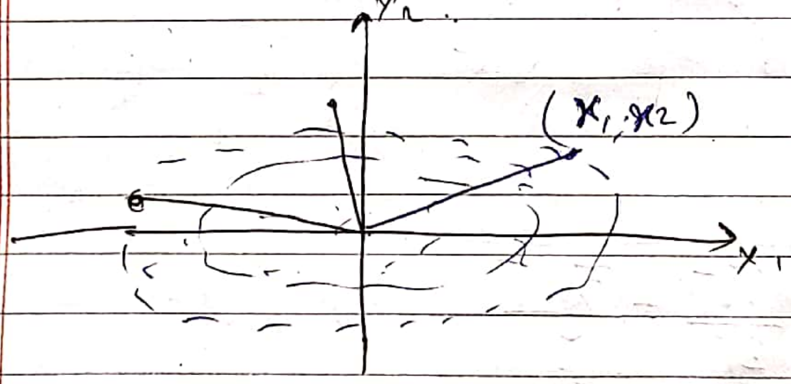


$$d(O,P) = \sqrt{x_1^2 + x_2^2}$$

if it's p dimensionale $= \sqrt{x_1^2 + x_2^2 + \cdots + x_p^2}$.

$$x_1^2 + x_2^2 + \cdots + x_p^2 = c^2$$

$$d(P,Q) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_p-y_p)^2}$$



uncorrelated

X_1 is more variable than X_2

$$x_1^* = \frac{x_1}{\sqrt{S_{11}}}, \quad x_2^* = \frac{x_2}{\sqrt{S_{22}}}$$

Statistical distance

$$d(O,P) = \sqrt{(x_1^*)^2 + (x_2^*)^2} = \sqrt{\frac{x_1^2}{S_{11}} + \frac{x_2^2}{S_{22}}}$$ → Covariance

$$\frac{x_1^2}{S_{11}} + \frac{x_2^2}{S_{22}} = c^2$$ Ellipse centered at 0,0

Scanned with CamScanner

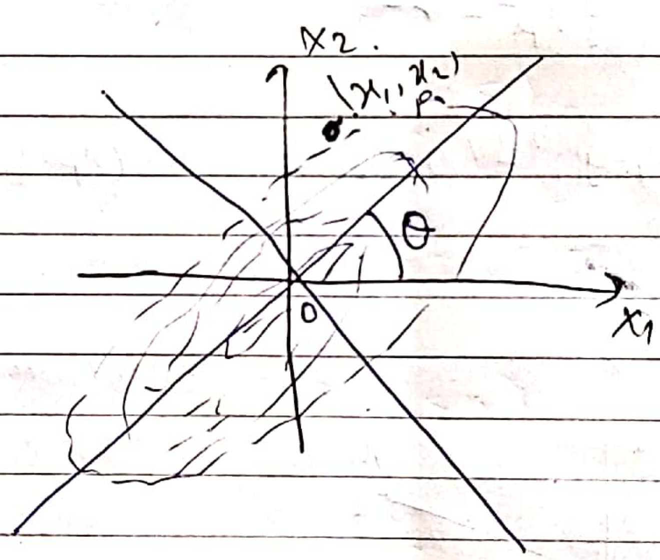
Date:
Page No.

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

$$d(P,Q) = \sqrt{\frac{(x_1-y_1)^2}{s_{11}} + \frac{(x_2-y_2)^2}{s_{22}}}$$

$Q = (y_1, y_2 \ldots y_p) \quad P = (x_1, x_2, \ldots x_p)$

$$d(P,Q) = \sqrt{\frac{(x_1-y_1)^2}{s_{11}} + \frac{(x_2-y_2)^2}{s_{22}} + \cdots + \frac{(x_p-y_p)^2}{s_{pp}}}$$



X_1 and X_2 are hugely correlated.

$$d(O,P) = \sqrt{\frac{\tilde{x}_1^2}{\tilde{s}_{11}} + \frac{\tilde{x}_2^2}{\tilde{s}_{22}}}$$

$$\tilde{x}_1 = x_1 \cos\theta + x_2 \sin\theta$$

$$\tilde{x}_2 = -x_1 \sin\theta + x_2 \cos\theta$$

$$d(O,P) = \sqrt{a_{11} x_1^2 + a_{22} x_2^2 + 2a_{12} x_1 x_2}$$

$$d(P,Q) = \sqrt{a_{11}(x_1-y_1)^2 + 2a_{12}(x_1-y_1)(x_2-y_2) + a_{22}(x_2-y_2)^2}$$

ellipse centred at Q. if we consider all points from Q at c distance.

prop 4.

$$X \sim N_p(\underset{\sim}{\mu}, \Sigma)$$

$$(\underset{\sim}{X} - \underset{\sim}{\mu})' \Sigma^{-1} (\underset{\sim}{X} - \underset{\sim}{\mu}) \sim \chi^2_p$$

Squared statistical distance from $\underset{\sim}{X}$ to $\underset{\sim}{\mu}$

⇒ Multivariate CLT

Let $X_1, X_2, X_3, \dots, X_n$ be independent observation from a pop^n with mean $\underset{\sim}{\mu}$, Covariance Σ then.

(1) $\sqrt{n}(\underset{\sim}{\bar{X}} - \underset{\sim}{\mu})$ is approximately $N_p(0, \Sigma)$

$$\bar{X} \sim N_p\left(\underset{\sim}{\mu}, \frac{\Sigma}{n}\right)$$

(2) $n(\underset{\sim}{\bar{X}} - \underset{\sim}{\mu})' S^{-1} (\underset{\sim}{\bar{X}} - \underset{\sim}{\mu})$ is approximately χ^2_p

for $n-p$ large