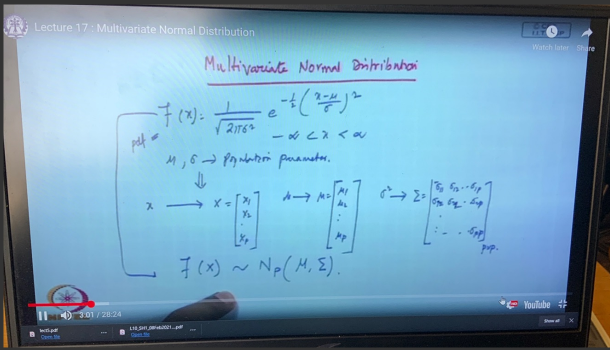
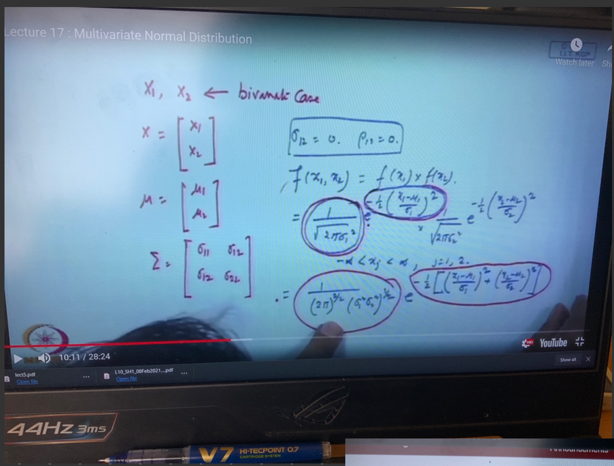
Lecture 1

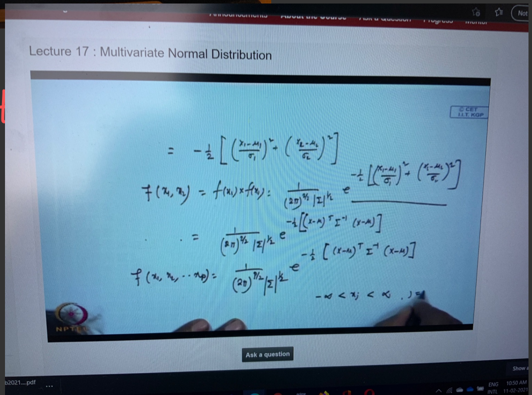


For bivariate case:

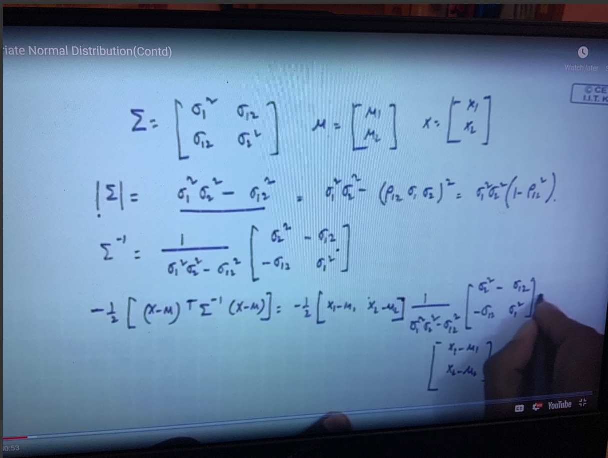
Assume they're independent

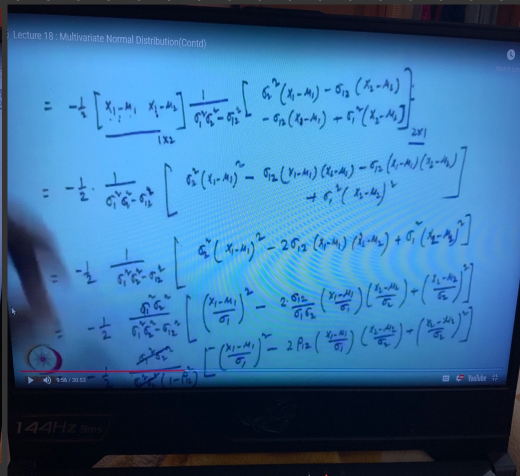


Derive den[?] exponents from populat



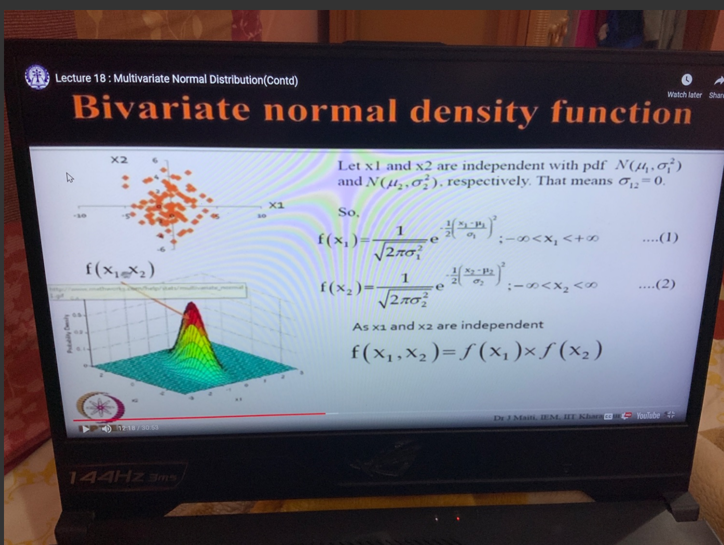
Resembles ellipse



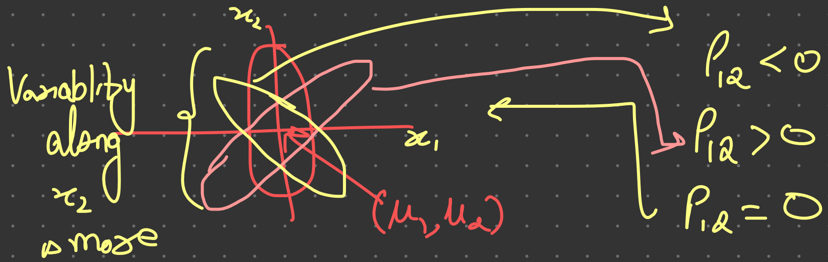


If $\rho_{12}=0$,

$$\text{Exponent} = \frac{-1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

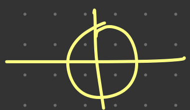


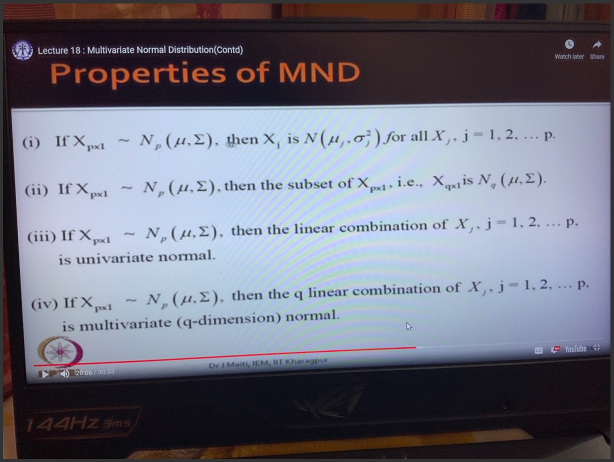
Represent ellipse:

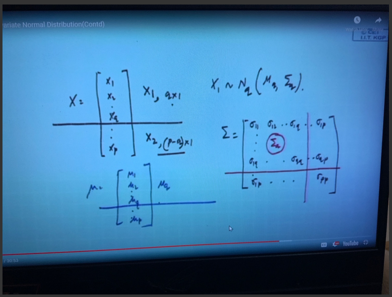


If varablity along x_1 is more

If $\sigma_1^2 = \sigma_2^2$



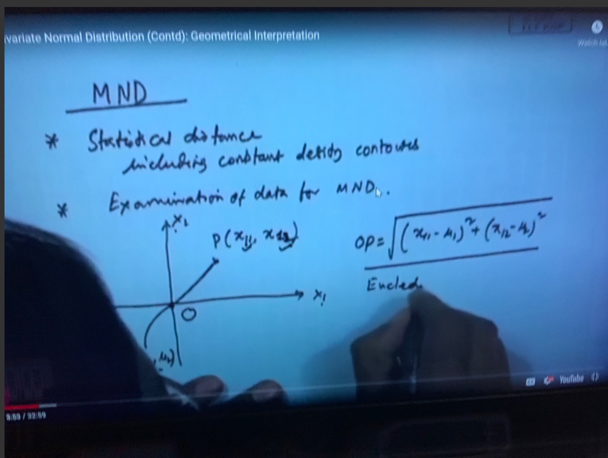




Le

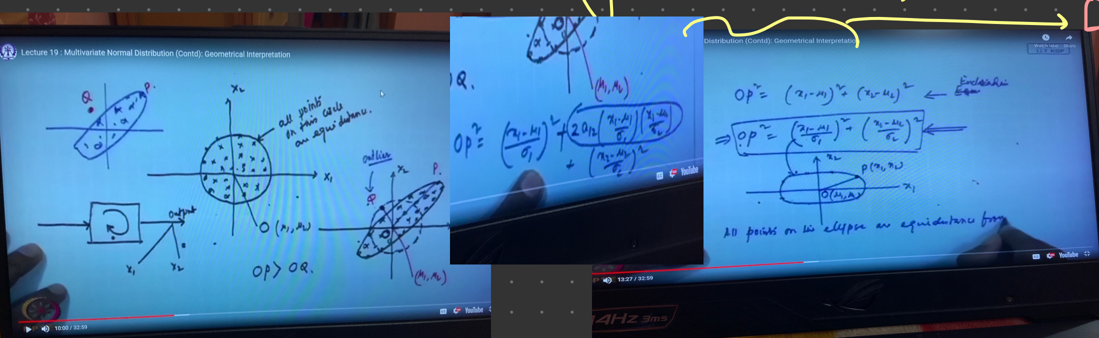
Lecture 19 geometric representation
Statistical distance including constant density contour
Examination of data for MND

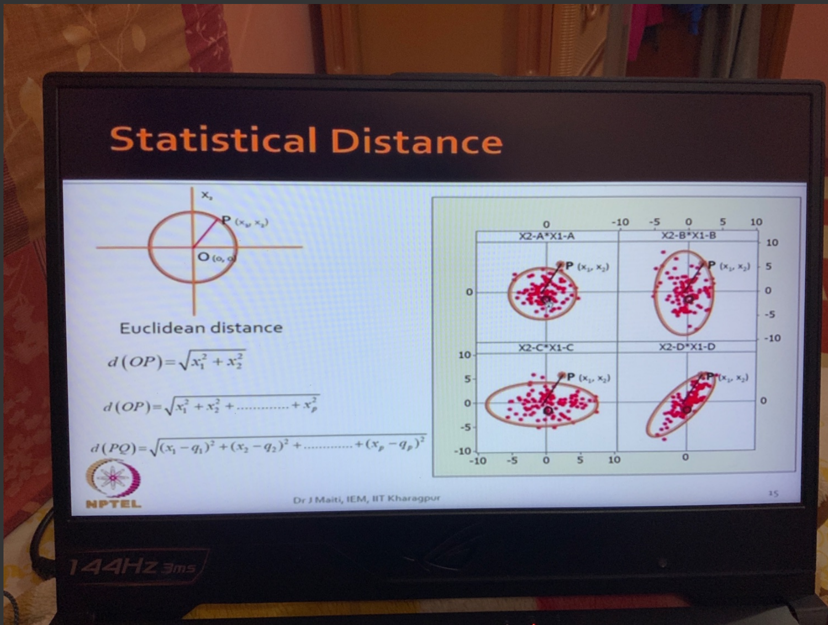
Kadian distance



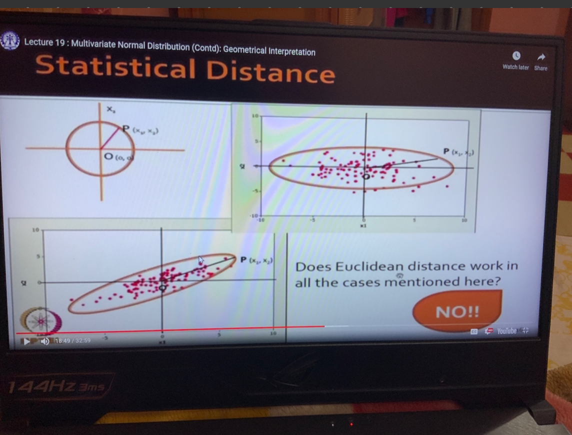
Euclean Distance → Bad as it doesn't capture variablity between x_1, x_2

Stastical Distance

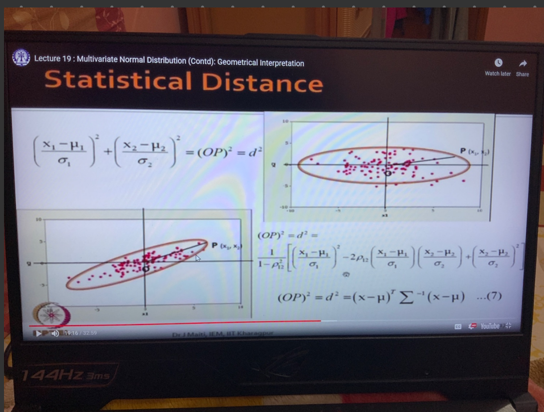




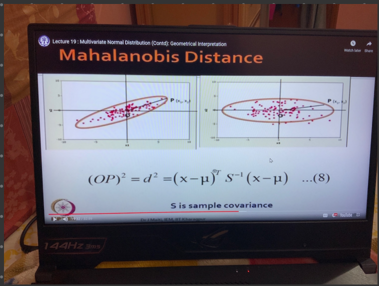
Statistical Distance
Euclidean distance
$d(OP)=\sqrt{x_1^2+x_2^2}$
$d(OP)=\sqrt{x_1^2+x_2^2+.............+x_p^2}$
$d(PQ)=\sqrt{(x_1-q_1)^2+(x_2-q_2)^2+.............+(x_p-q_p)^2}$
X2-A*X1-A
X2-B*X1-B
X2-C*X1-C
X2-D*X1-D
NPTEL
Dr J Maiti, IEM, IIT Kharagpur
144Hz 3ms



Lecture 19 : Multivariate Normal Distribution (Contd): Geometrical Interpretation
Statistical Distance
Does Euclidean distance work in all the cases mentioned here?
NO!!
144Hz 3ms

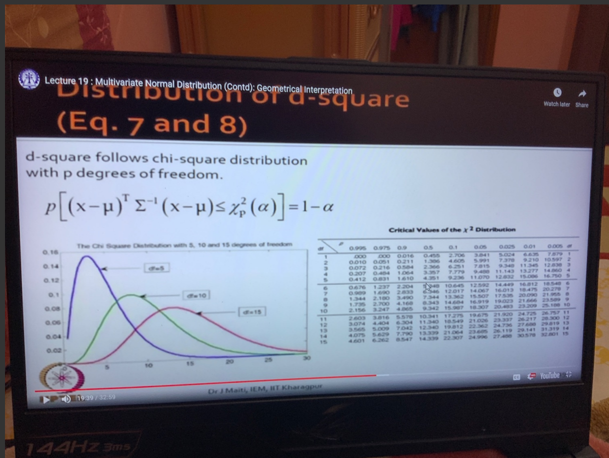


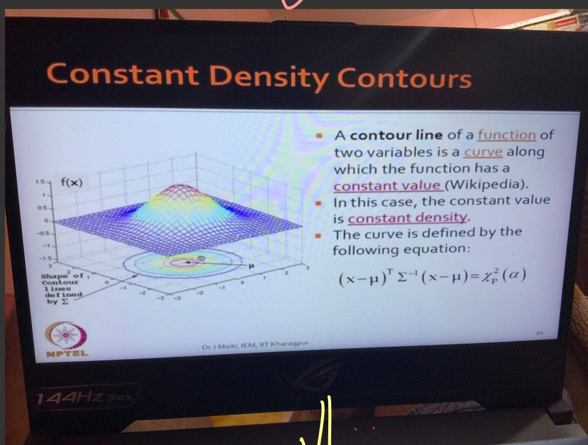
Lecture 19 : Multivariate Normal Distribution (Contd): Geometrical Interpretation
Statistical Distance
$\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2=(OP)^2=d^2$
$(OP)^2=d^2=$
$\frac{1}{1-\rho_{12}^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$
$(OP)^2=d^2=(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})$ ...(7)
144Hz 3ms

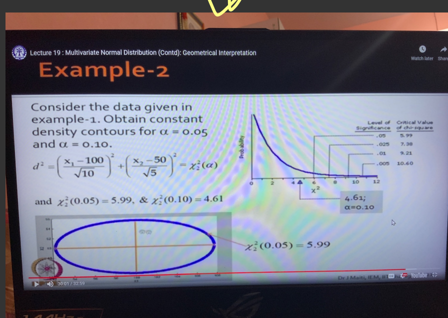


Lecture 19 : Multivariate Normal Distribution (Contd): Geometrical Interpretation
Mahalanobis Distance
$(OP)^2=d^2=(\mathbf{x}-\boldsymbol{\mu})^T S^{-1}(\mathbf{x}-\boldsymbol{\mu})$ ...(8)
S is sample covariance
144Hz 3ms

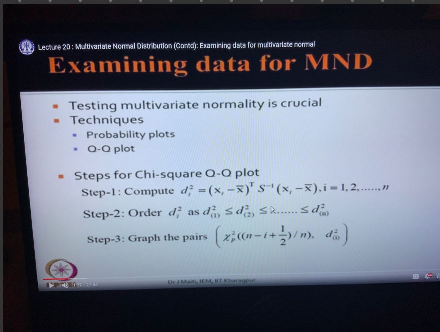
statistical distance
?
!
Exponent follows chi-square

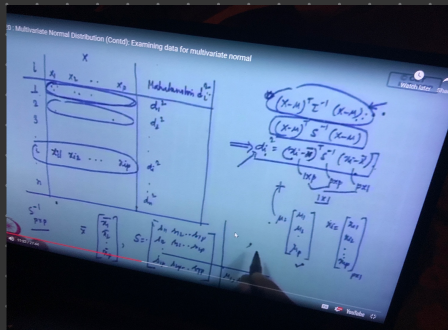


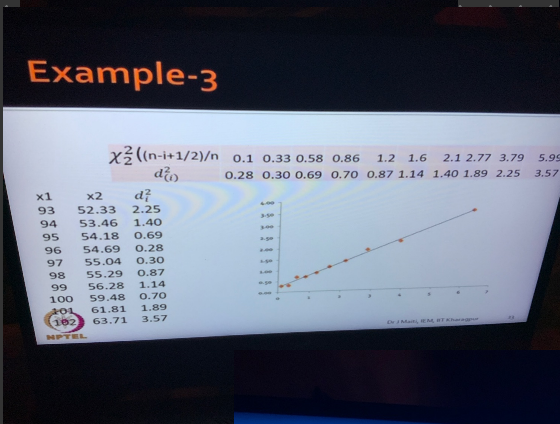


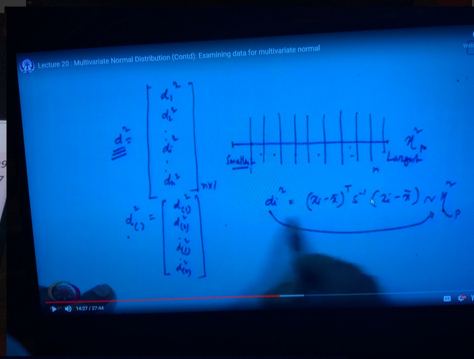


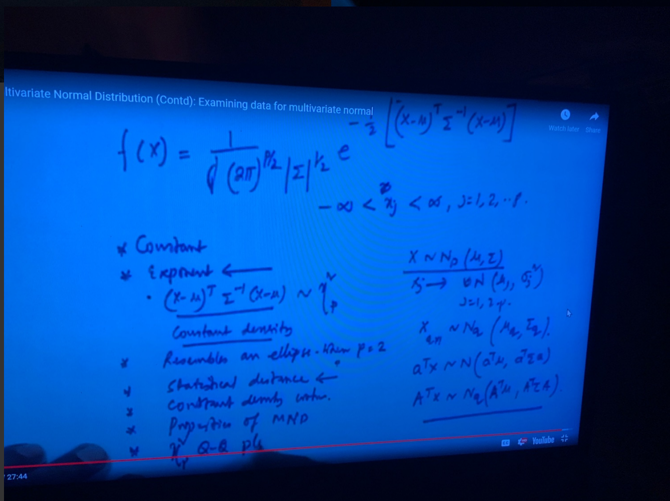
Lecture-20


Lecture 20 : Multivariate Normal Distribution (Contd): Examining data for multivariate normal
Examining data for MND
Testing multivariate normality is crucial
Techniques
Probability plots
Q-Q plot
Steps for Chi-square Q-Q plot
Step-1: Compute $d_i^2 = (x_i - \bar{x})^T S^{-1} (x_i - \bar{x}), i = 1, 2, \dots, n$
Step-2: Order d_i^2 as $d_{(1)}^2 \le d_{(2)}^2 \le \dots \le d_{(n)}^2$
Step-3: Graph the pairs $\left(\chi_p^2((n-i+\frac{1}{2})/n),\ d_{(i)}^2\right)$
Dr J Maiti, IEM, IIT Kharagpur


Multivariate Normal Distribution (Contd): Examining data for multivariate normal
Mahalanobis d_i^2
$(X-\mu)^T \Sigma^{-1} (X-\mu)$
$(X-\mu)^T S^{-1} (X-\mu)$
$d_i^2 = (x_i-\bar{x})^T S^{-1} (x_i-\bar{x})$
Watch later
Share
YouTube


Example-3
$\chi_2^2((n-i+1/2)/n)$ 0.1 0.33 0.58 0.86 1.2 1.6 2.1 2.77 3.79 5.99
$d_{(i)}^2$ 0.28 0.30 0.69 0.70 0.87 1.14 1.40 1.89 2.25 3.57
x1 x2 d_i^2
93 52.33 2.25
94 53.46 1.40
95 54.18 0.69
96 54.69 0.28
97 55.04 0.30
98 55.29 0.87
99 56.28 1.14
100 59.48 0.70
101 61.81 1.89
102 63.71 3.57
NPTEL
Dr J Maiti, IEM, IIT Kharagpur


Lecture 20 : Multivariate Normal Distribution (Contd): Examining data for multivariate normal
Smallest
Largest
$d_i^2 = (x_i-\bar{x})^T S^{-1} (x_i-\bar{x}) \sim \chi_p^2$
14:27 / 27:44


ltivariate Normal Distribution (Contd): Examining data for multivariate normal
$f(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}[(x-\mu)^T \Sigma^{-1} (x-\mu)]}$
$-\infty < x_j < \infty, j = 1, 2, \dots p$
* Constant
* Exponent
$(x-\mu)^T \Sigma^{-1} (x-\mu) \sim \chi_p^2$
Constant density
* Resembles an ellipse when p = 2
* Statistical distance
* Constant density contour
* Properties of MND
* χ^2 Q-Q plot
$X \sim N_p(\mu, \Sigma)$
$a^T X \sim N(a^T\mu, a^T \Sigma a)$
$A^T X \sim N_q(A^T\mu, A^T \Sigma A)$
Watch later
Share
YouTube
27:44