BONTA AALAYA
12340530
DSL253 - Statistical Programming

1

INTRODUCTION

In this assignment, we observe the behaviour of bivariate normal distribution under the effect of various parameters

DATA

Cdf is visualised using in-built matlab function mvncdf and parameters are given

METHODOLOGY

Question 1

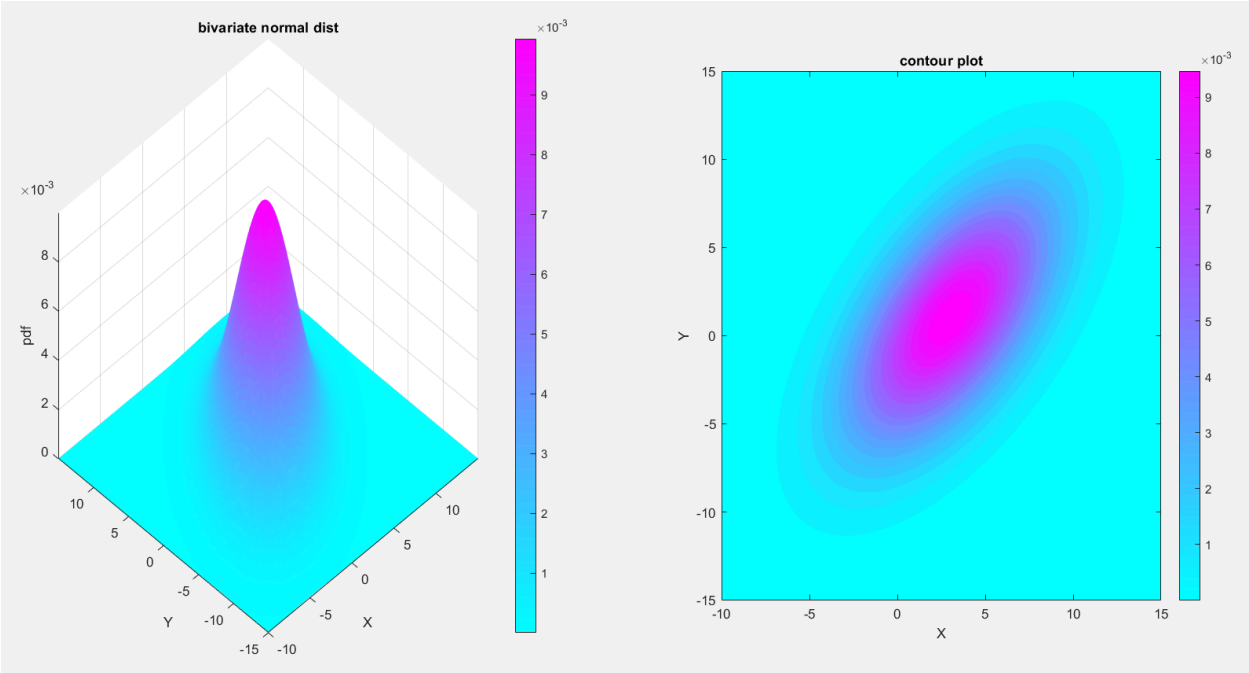
Parameters are taken from user, and standard deviation and covariance of x and y are computed. From this, we get covariance matrix. We create a new variable 'mean_vec' to store values of mean(x) and mean(y) together. Since marginal distribution of Y is normal, i.e, $Y \sim N(\mu_Y, \sigma_Y^2)$, we can use normcdf here to calculate P(3 < Y < 8).

$$E(Y|x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$$

$$Y|X = x \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(X - \mu_X), \quad \sigma_Y^2(1-\rho^2)\right)$$

RESULTS

Region	Probability
P(3 < Y < 8)	0.2638
P(3 < Y < 8 | X = 7)	0.4401
P(-3 < X < 3)	0.4332
P(-3 < X < 3 | Y = -4)	0.6431



DISCUSSION

This section covers the observations from the results table (computed for

$\mu_X = 3, \mu_Y = 1, \sigma_X^2 = 16, \sigma_Y^2 = 25, \rho_{xy} = 3/5$)

P(3 < Y < 8) is low, which is expected from the spread of y (since variance is high)

P(3 < Y < 8 | X = 7) is greater than P(3 < Y < 8), which is expected due to positive correlation between x and y. Knowing X = 7 reduces uncertainty, which makes probability higher

P(-3 < X < 3) is reasonable since $X \sim N(3, 16)$

P(-3 < X < 3 | Y = -4) > P(-3 < X < 3) again, expected due to positive correlation

CONCLUSION

We can understand how probabilities are affected by marginalization and conditioning it with the other variable in a bivariate normal distribution. In this case, we took a positive correlation between the two variables

INTRODUCTION

In this assignment, we observe the behaviour of multivariate normal distributions and observe the effect of parameters and calculate Mahalanobis distances to help visualise better

DATA

Parameters are taken from user and the pdf is generated using built-in **mvnpdf** function

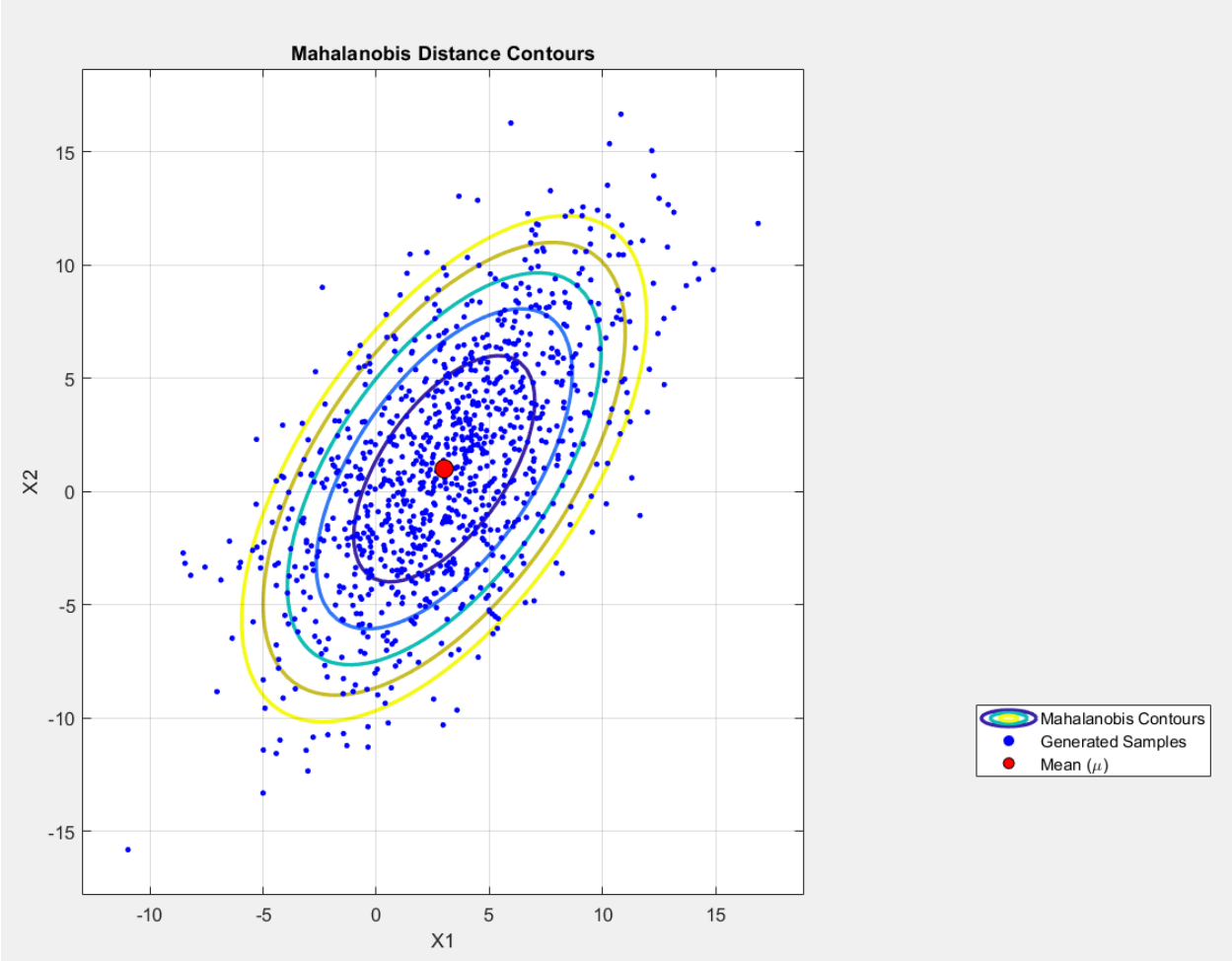
METHODOLOGY

Question 2

We take mean and variance dynamically, for simplicity, we assume correlation coeff for all variables is the same and check if covariance matrix is correctly computed.
Value of c can be adjusted. Here, I used **chi2inv** function to get a value of c for

$\mu_X = 3,\ \mu_Y = 1,\ \sigma_X^2 = 16,\ \sigma_Y^2 = 25,\ \rho_{xy} = 3/5$

RESULTS

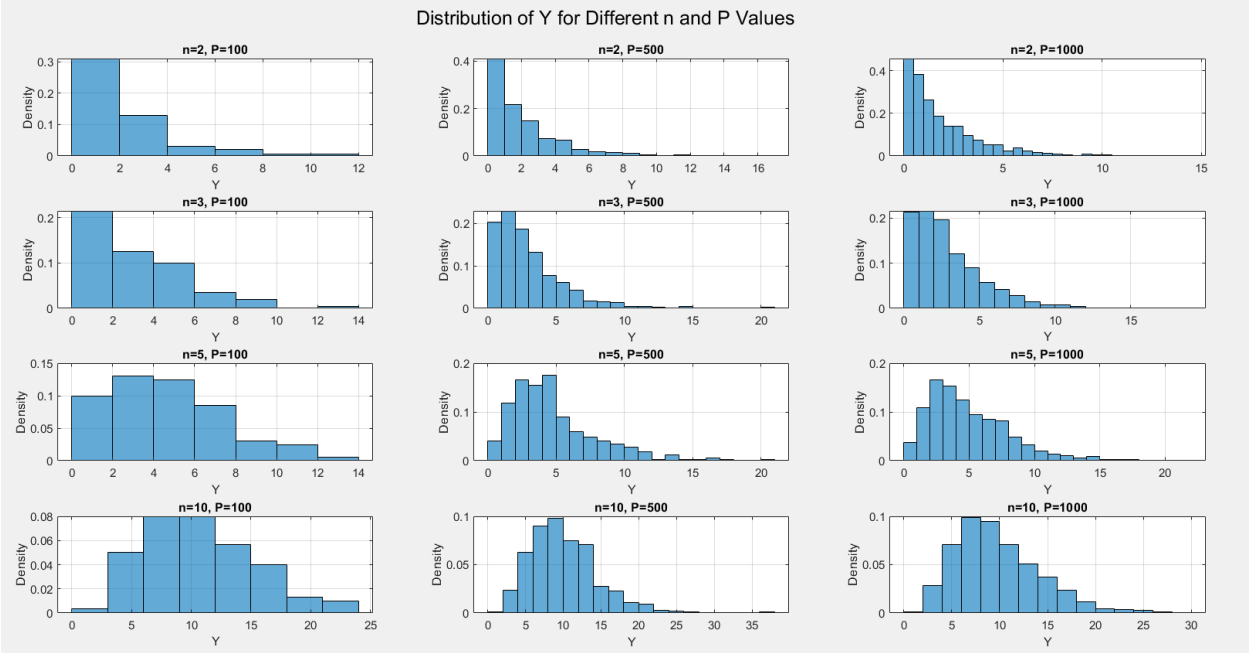


For c = 1.39

Probability P((X - mean)' * inv(cov_matrix) * (X - mean) ≤ 1.39) = 0.5130

For c = chi2inv(0.95,2)

Probability P((X - mean)' * inv(cov_matrix) * (X - mean) ≤ 5.99) = 0.9490



DISCUSSION

Mahalanobis distance follows chi square dist with dimension 2. The chosen c = 1.39 corresponds to the 50th percentile (median) of $\chi 2$. Empirical value 0.9490 is close to theoretical 0.95, indicating correct code. Higher correlation increases spread of points. These contours align with eigenvectors of Σ, helping us visualise the spread of normal distribution in different directions

CONCLUSION

We generated and validated samples from bivariate normal dist. The results confirm that the Mahalanobis distance effectively models the spread and correlation of multivariate Gaussian data. The contour plot successfully illustrates how points are distributed relative to their statistical significance

INTRODUCTION

In this assignment, we classify a list of data points into either one of two distributions

DATA

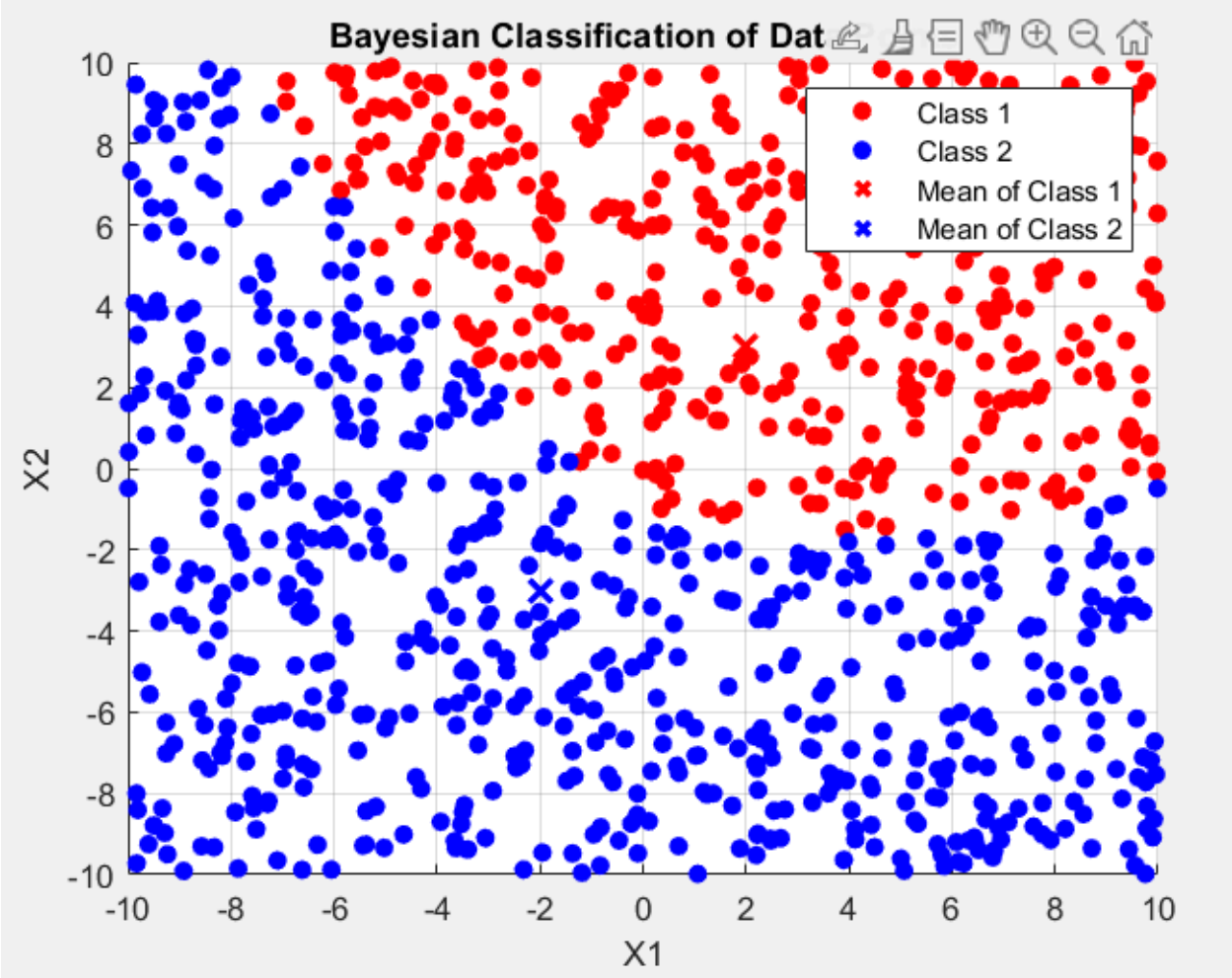
The data points were provided in a .txt file

METHODOLOGY

Question 3

We skip 1st cell because it is not a numerical value and only consider columns 2 and 3. Then, we calculate likelihood using in-built function and classify based on likelihood

RESULTS



DISCUSSION

Bayesian classifier classifies data points into 2 different classes. Classification is based on Mahalanobis distance. We can clearly observe the decision boundary from the above result.

CONCLUSION

We were successfully able to classify data points into 2 different classes, showing how Bayesian classification is implemented,

---