Question 2
Multivariate Gaussian

Atishay Jain & Gohil Megh Hiteshkumar

October 15, 2022

1 Generating Sample points from 2D Gaussian

We have,

$$Covariance\,Matrix, C = \begin{bmatrix} 1.6250 & -1.9486 \\ -1.9486 & 3.8750 \end{bmatrix}$$

$$Mean, \mu = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

- First, we used $eig()$ function in MATLAB to calculate a diagonal matrix with eigenvalues of C as its diagonal elements (let this matrix be E) and a matrix with eigenvectors of C as its columns (let this matrix be P).
- The matrix P is orthogonal as C is symmetric (therefore it can be diagonalized by an orthogonal matrix). So, we can write C in terms of P and E as -

$$C = PEP^{-1}$$

- Let X be the Multivariate Gaussian, W be the column vector of univariate i.i.d Gaussian, A be a coefficient matrix and μ be the Mean vector, then we can write X as

$$X = AW + \mu$$

- Now, Covariance matrix of $X := AW + \mu$ is AA^T. Since $P^{-1} = P^T$, we have $C = PEP^T$, and as eigenvalues (diagonal element of E) are real, we can split the matrix E into product of two identical diagonal matrices (say, M) such that their diagonal elements are square root of diagonal element of E.
- So, $E = MM$ and as M is also diagonal, we have $M = M^T$. Putting M in place of E in expression for C and from both expressions of C, we have

$$C = AA^T = PMM^TPT = (PM)(PM)^T$$

- On comparing both sides, a possible value of A is definitely $A = PM$.
- After getting a possible matrix A, we repreated the given experiment for each N. We generated N normally distributed random vectors as a column vectors of a matrix. We used $randn()$ function in MATLAB for this.
- Then, we used $X = AW + \mu$ for generating N number of Multivariate Gaussian sample points.
- At last, we also plotted the Scatter plots for each value of $N \in \{10, 100, 1000, 10000, 100000\}$

2 ML Estimate of Mean

2.1 Calculation of ML Estimate

For Maximum Likelihood Estimation of mean -

- Take the log likelihood function, which is sum of Multivariate gaussian's PDF's logarithms
- Differentiate it with respect to μ and equate to 0

The PDF of Multivariate gaussian is

$$p(y) = \frac{1}{(2\pi)^{\frac{D}{2}}|C|^{0.5}} e^{-0.5(y-\mu)^T C^{-1}(y-\mu)}$$

Taking log of the above PDF to get log likelihood function as -

$$\sum_{i=1}^{N} -0.5(y_i - \mu)^T C^{-1}(y_i - \mu) - \log((((2\pi)^D)|C|)^{0.5})$$

We use this formula for differentiation -

$$\frac{\partial}{\partial \mu}(\mu - x)^T C^{-1}(\mu - x) = 2C^{-1}(\mu - x)$$

On differentiating, we get

$$\sum_{i=1}^{N} \frac{\partial}{\partial \mu}(-0.5(y_i - \mu)^T C^{-1}(y_i - \mu)) - \sum_{i=1}^{N} C^{-1}(y_i - \mu) = 0$$

On multiplying by C on both side and solving, we get the MLE result of Mean as -

$$\mu = \frac{1}{N}\sum_{i=1}^{N} y_i$$

2.2 Boxplot of error

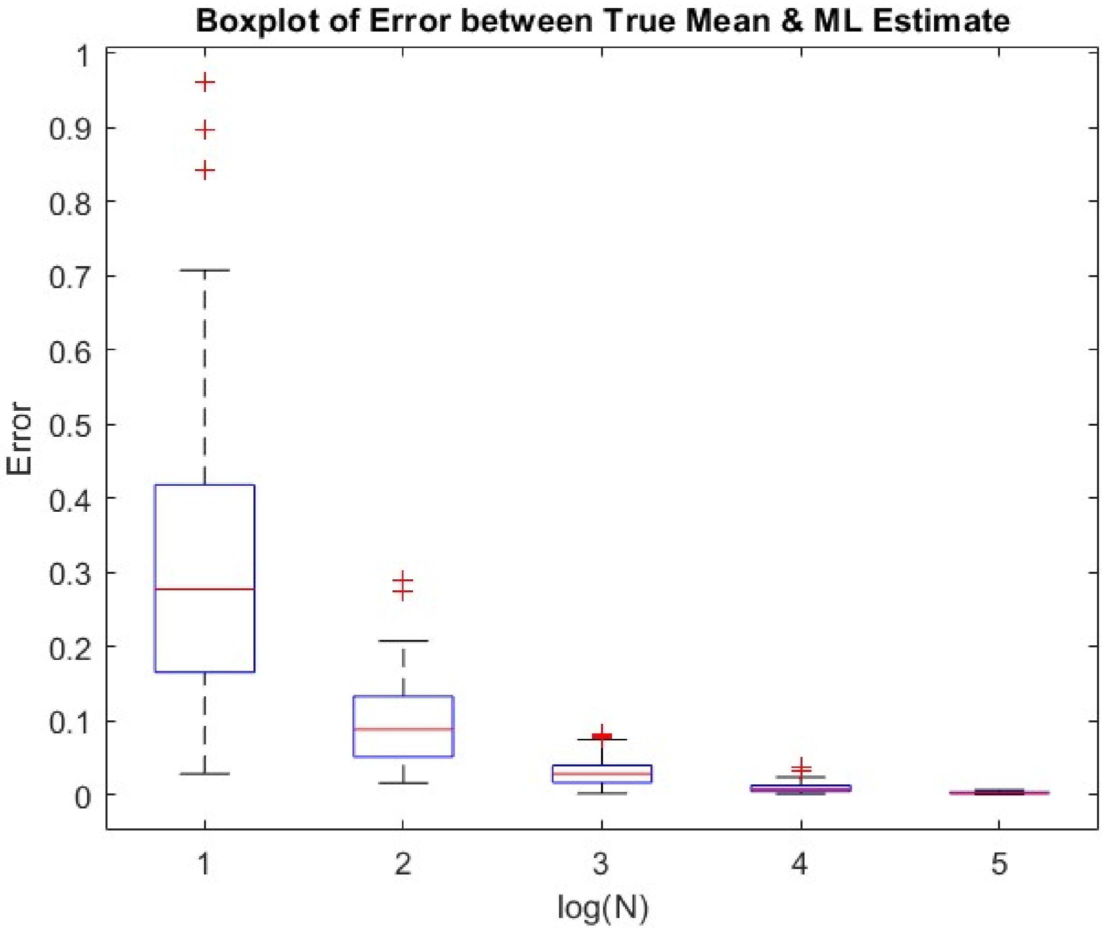

Figure 1: Boxplot of error between True Mean and ML Estimate of Mean

The above Boxplot shows errors between True and ML estimate values of Mean for different values of N and the x-label is taken as $log(N)$

3 ML Estimate of Covariance

3.1 Calculation of ML Estimate

Again as in ML estimate for Mean, we take the same log likelihood function as above, but differentiate with respect to C and equate it to 0
We use the following formulas for differentiating wrt. C

$$\frac{\partial}{\partial C}(x-\mu)^T C^{-1}(x-\mu) = -C^{-T}(x-\mu)(x-\mu)^T C^{-T}$$

$$\frac{\partial}{\partial C}\log(|C|) = C^{-T}$$

Using the above formulas and differentiating the log likelihood function -

$$\sum_{i=1}^{N} -0.5(C^{-T}(y_i-\mu)(y_i-\mu)^T C^{-T} - C^{-T}) = 0$$

$$\frac{1}{N}\sum_{i=1}^{N}(y_i-\mu)(y_i-\mu)^T = C^T$$

As we know that Covariance Matrix is symmetric, so $C = C^T$, and we get

$$\frac{1}{N}\sum_{i=1}^{N}(y_i-\mu)(y_i-\mu)^T = C$$

Thus, we get the ML Estimate for C.

3.2 Boxplot of error

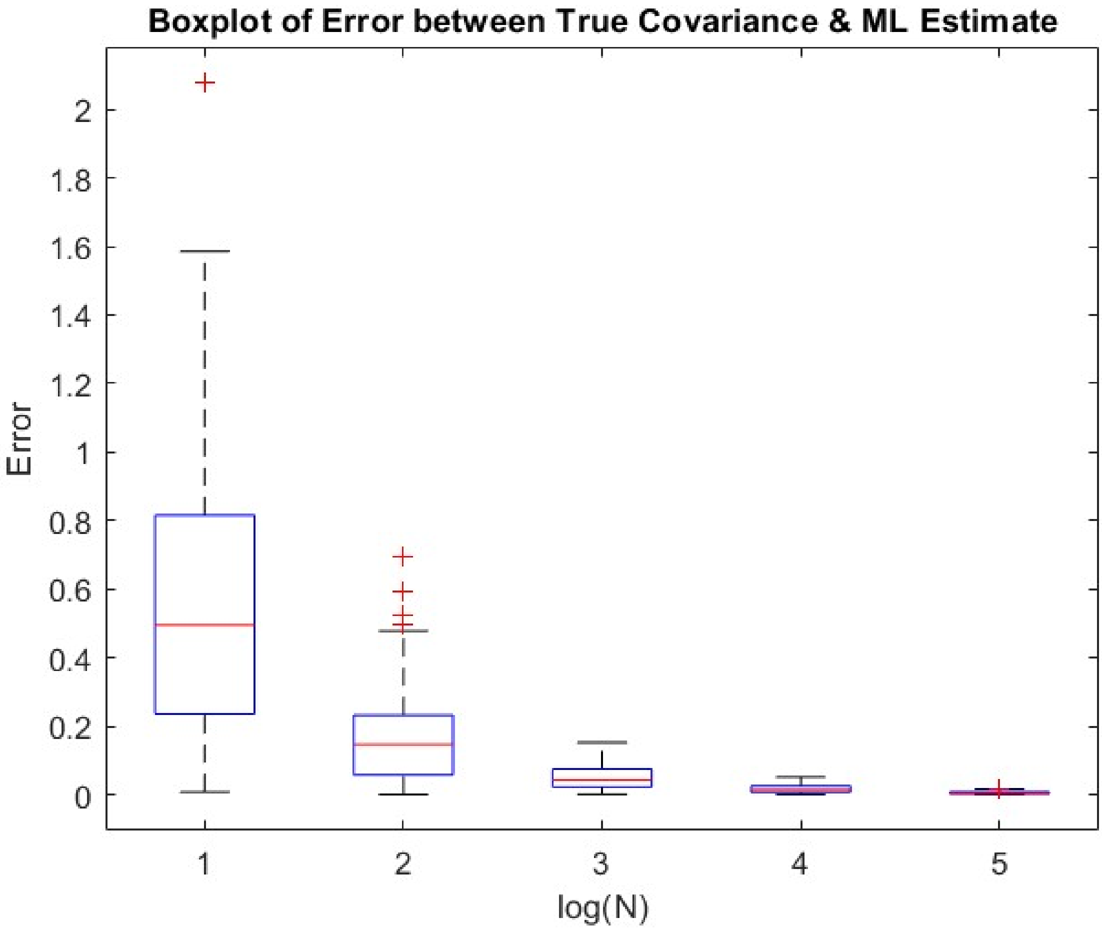

Figure 2: Boxplot of error between True Covariance and ML Estimate of Covariance

The above Boxplot shows errors between True and ML estimate values of Covariance for different values of N and the x-label is taken as $log(N)$

4 Principal modes of Variation

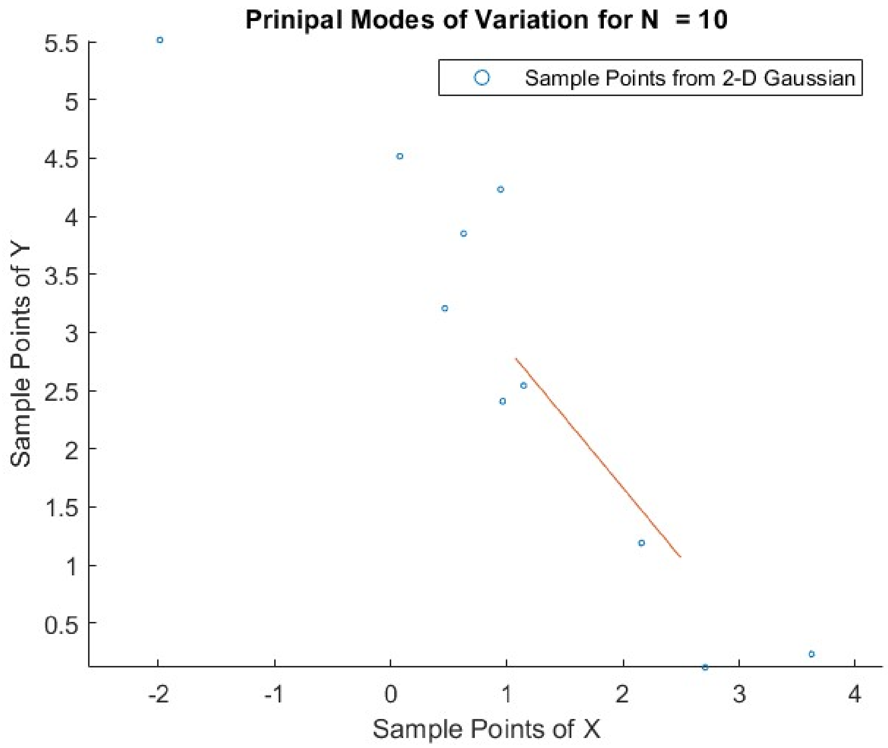

Figure 3: Sample points scatter plot & Principal Modes of Variation

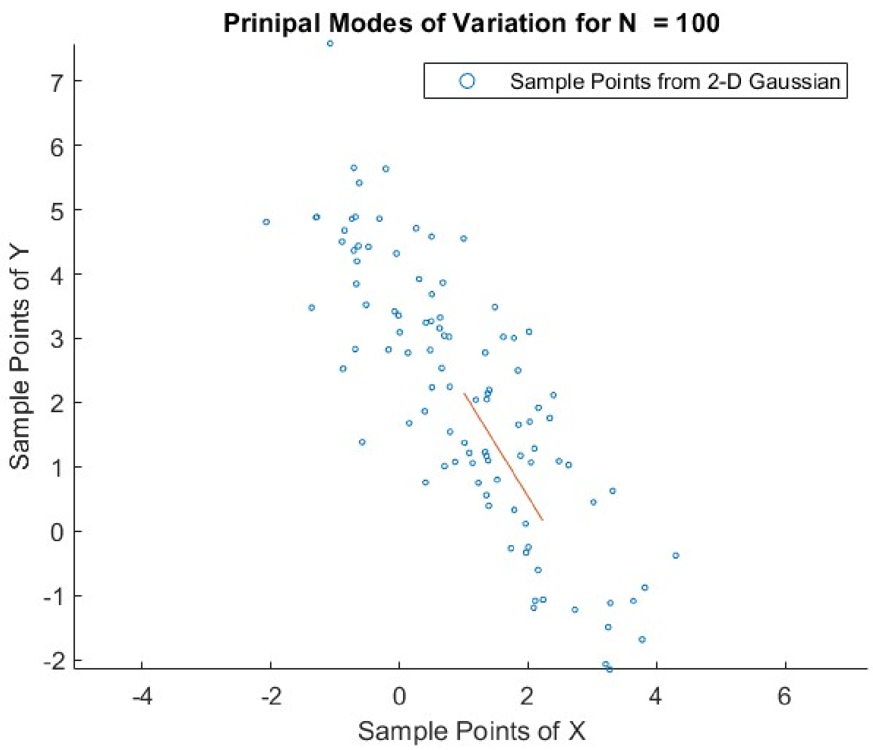

Figure 4: Sample points scatter plot & Principal Modes of Variation

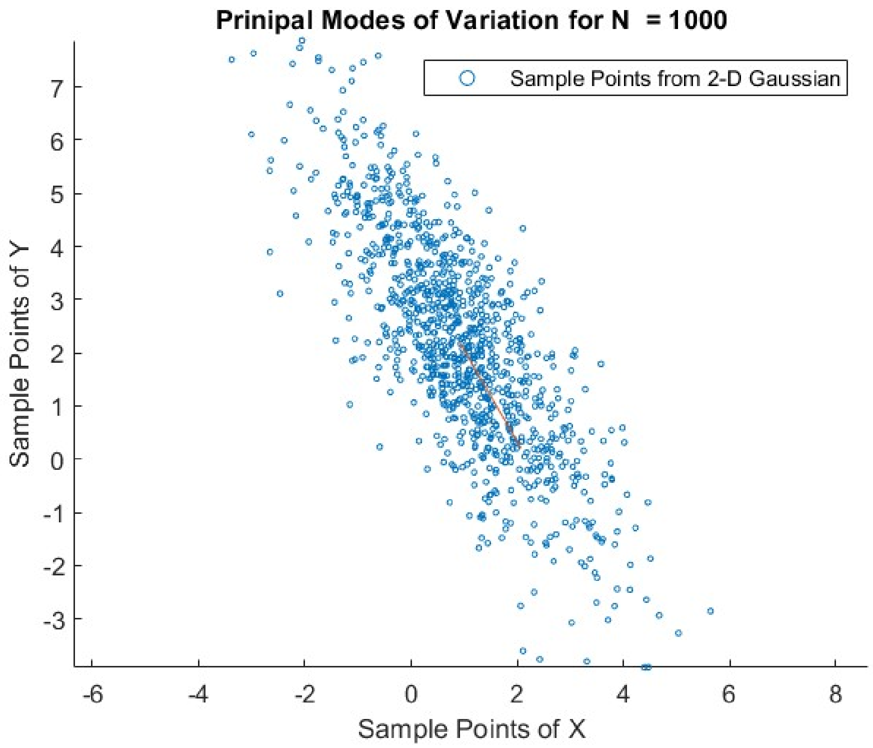

Figure 5: Sample points scatter plot & Principal Modes of Variation

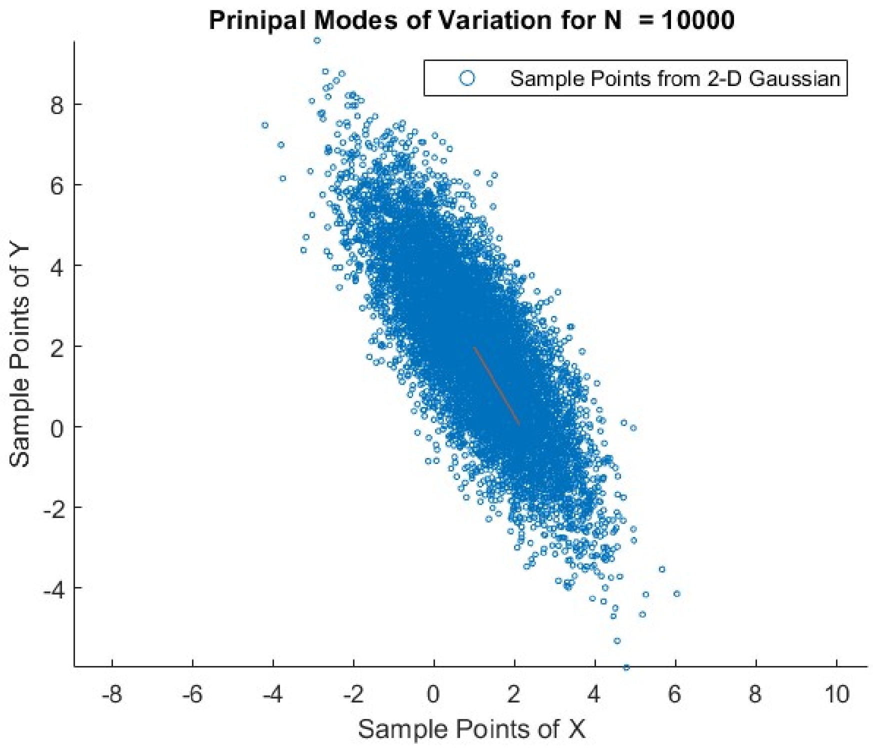

Figure 6: Sample points scatter plot & Principal Modes of Variation

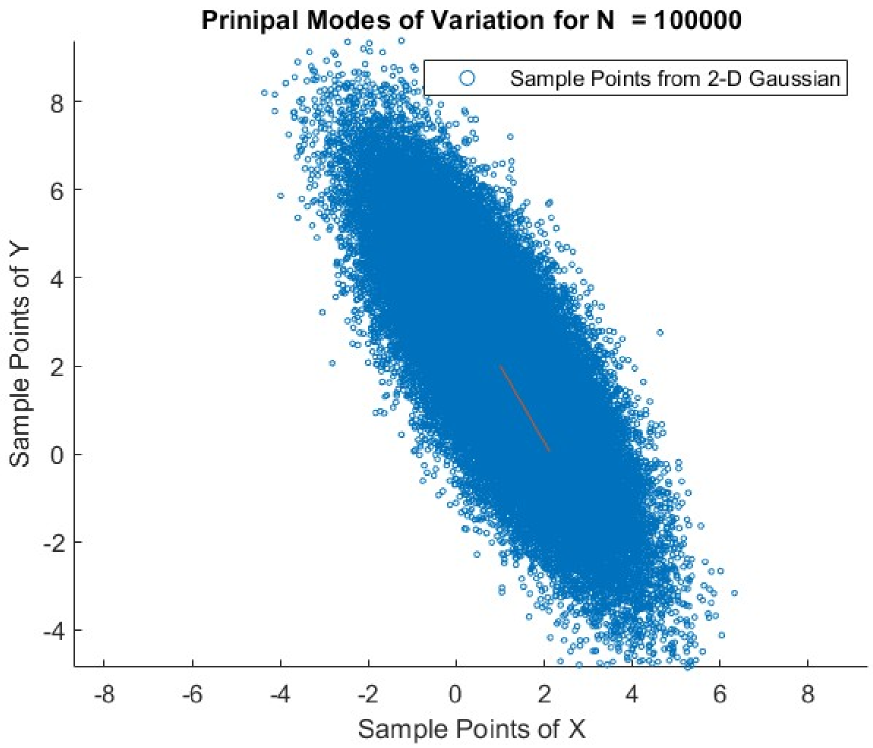

Figure 7: Sample points scatter plot & Principal Modes of Variation

The above plots show the 2D scatter plot of the generated data and the principal modes of variation of the data by plotting a line starting at the empirical mean and going a distance equal to the empirical eigenvalue's square root along a direction given by the empirical eigen-vector for each value of N.