CS512– Artificial Intelligence

Lab Assignment 2

PART(i)

(A) DERIVATION

Q1> Let us assume

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad \text{where} \quad x \sim N(\mu, \Sigma).$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$

Assuming $\Sigma_{21}^T = \Sigma_{12}$.

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

To derive $\rightarrow$ $P(x_2 \mid x_1) = \dfrac{P(x_1, x_2)}{P(x_1)}$.

As per the multivariate Gaussian :

$$P(x) = P(x_1, x_2)$$

$$= \frac{1}{\sqrt{(2\pi)^n \Sigma}} \exp\left(\frac{-1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu).\right)$$

Let $F(x_1, x_2) = (x-\mu)^T \Sigma^{-1} (x-\mu)$.

solving $F(x_1, x_2). = \left[(x_1-\mu_1)^T, (x_2-\mu_2)^T\right] *$

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} * \begin{bmatrix} (x_1-\mu_1) \\ (x_2-\mu_2) \end{bmatrix}.$$

we assume

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}. \rightarrow X_{12} = (X^{21})^T$$

$$X_{11} = \Sigma_{11}^{-1} + \Sigma_{11}^{-1} \Sigma_{12} \left(\Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1} \Sigma_{12}^T \Sigma_{11}^{-1}$$

$$= 1 + 1\,a \left(1 - a \times 1 \times a \right)^{-1} a$$

$$= 1 + a\,(1 - a^2)\,a = 1 + \frac{a^2}{1-a^2}.$$

$$X_{12} = -\Sigma_{11}^{-1} \Sigma_{12} \left(\Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1}$$

$$= -1 * a \left(1 - 1 * \frac{1}{1} * a \right)^{-1}$$

$$= -a\,(1-a)^{-1} = \frac{-a}{1-a}$$

$$X_{22} = \Sigma_{22}^{-1} + \Sigma_{22}^{-1} \Sigma_{12}^T \left(\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^T \right)^{-1} \Sigma_{12} \Sigma_{22}^{-1}$$

$$= 1 + 1 * a \left(1 - a * \frac{1}{1} * a \right)^{-1} a\,1^{-1}$$

$$= 1 + a\,(1-a^2)^{-1} a = 1 + \frac{a^2}{1-a^2}.$$

$$X_{21} = \frac{-a}{1-a}$$

$$F(x_1, x_2) = \left[(x_1-\mu_1)^T \cancel{X}_{11} + (x_2-\mu_2)^T X_{21} \ , \ (x_1-\mu_1)^T X_{12} + (x_2-\mu_2)^T X^{22} \right] \begin{bmatrix} (x_1-\mu_1) \\ (x_2-\mu_2) \end{bmatrix}$$

$$= (x_1-\mu_1)^T \cancel{X}_{11} (x_1-\mu_1) + (x_2-\mu_2)^T X_{21} (x_1-x_1) +$$

$$(x_1-x_1)^T X_{12} (x_2-\mu_2) + (x_2-\mu_2)^T X_{22} (x_2-\mu_2)$$

$$= (x_1-\mu_1)^T X_{11} (x_1-\mu_1) + 2(x_1-\mu_1)^T X_{12} (x_2-\mu_2) + (x_2-\mu_2)^T \cancel{X}_{22} (x_2-\mu_2)$$

$$= \underbrace{(x_1-\mu_1)^T \Sigma_{11}^{-1} (x_1-\mu_1)} + [x_2-\mu_2 - \Sigma_{12}^T \Sigma_{11}^{-1} (x_1-\mu_1)]^T$$

$$(\Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12})^{-1} [x_2-\mu_2 - \Sigma_{12}^T \Sigma_{11}^{-1} (x_1-\mu_1)]$$

$$\underline{F(x_1, x_2) = F(x_1) + F(x_1, x_2)}$$

on dividing with $P(x_1) \rightarrow$ we will get $F(x_1, x_2)$ So,

$$x_2 | x_1 \sim N\left(\mu_2 + \Sigma_{12}^T \Sigma_{11}^{-1} (x_1-\mu_1) , \Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12} \right)$$

$$\sim N\left(2 + \frac{a}{1}(x_1 - 1) , 1 - \frac{a}{1}(a) \right)$$

$$\sim N\left(2 + a(x_1 - 1) , 1 - a^2 \right)$$

To derive $P(x_1 | x_2) = \dfrac{P(x_1, x_2)}{P(x_1)}$.

$$P(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^n \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

assume $F(x_1, x_2) = (x-\mu)^T \Sigma^{-1} (x_1 - \mu)$.

$$F(x_1, x_2) = \left[(x_1-\mu_1)^T, (x_2-\mu_2)^T\right] * \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} (x_1-\mu_1) \\ (x_2-\mu_2) \end{bmatrix}$$

let $\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$.

$$F(x_1, x_2) = \left[(x_1-\mu_1)^T X_{11} + (x_2-\mu_2)^T X_{21}, (x_1-\mu_1)^T X_{12} + (x_2-\mu_2)^T x_{22}\right] * \begin{bmatrix} (x_1-\mu_1) \\ (x_2-\mu_2) \end{bmatrix}.$$

$$= (x_1-\mu_1)^T X_{11} (x_1-\mu_1) + (x_2-\mu_2)^T X_{21} (x_1-\mu_1) + (x_1-\mu_1)^T X_{12} (x_2-\mu_2) + (x_2-\mu_2)^T X_{22} (x_2-\mu_2)$$

$$= (x_1-\mu_1)^T X_{11} (x_1-\mu_1) + 2(x_1-\mu_1)^T X_{12} (x_2-\mu_2) + (x_2-\mu_2)^T X_{22} (x_2-\mu_2).$$

$$\Rightarrow \underbrace{(x_1-\mu_1)^T \Sigma_{11}^{-1} (x_1-\mu_1)} + \left[x_2-\mu_2-\Sigma_{12}^T \Sigma_{11}^{-1}(x_1-\mu_1)\right]^T$$

$$\left(\Sigma_{22}-\Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12}\right)^{-1}\left[x_2-\mu_2-\Sigma_{12}^T \Sigma_{11}^{-1}(x_1-\mu_1)\right].$$

$\Rightarrow$.

$$F(x_1, x_2) = F(x_1) + F(x_1, x_2)$$

On dividing the $p(x)$ expression with $p(x_2)$. we get

$$P(x_1|x_2) \sim N\left(\mu_1 + \Sigma_{12}^T \Sigma_{22}^{-1}(x_2-\mu_2),\ \Sigma_{11}-\Sigma_{12}^T \Sigma_{22}^{-1} \Sigma_{12}\right)$$

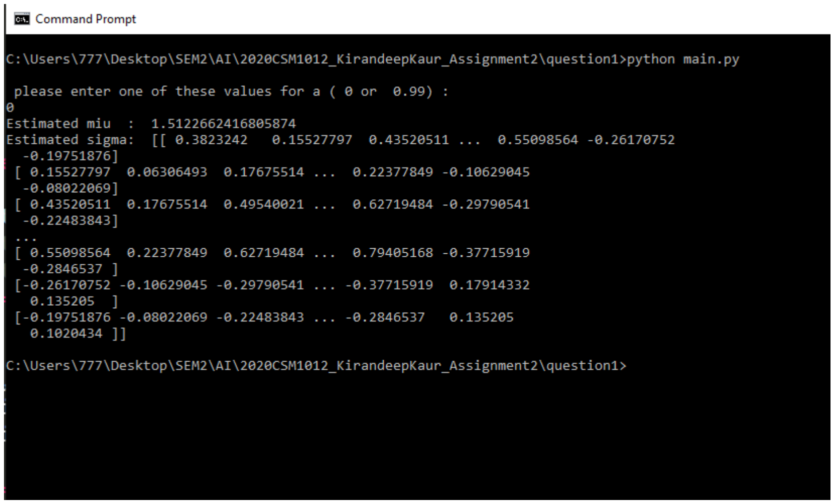
$$\sim N\left(1+\frac{a}{1}(x_2-2),\ 1-\frac{a}{1}*a\right).$$

$$\sim N\left(1+a(x_2-2),\ 1-a^2\right)$$

(.B) a = 0

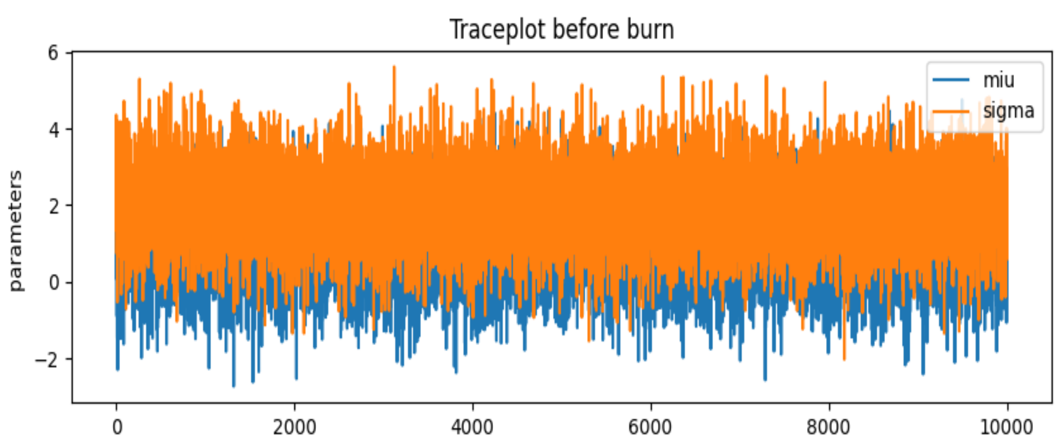
n(iterations) = 10000

The following were the estimated values for Σ and μ



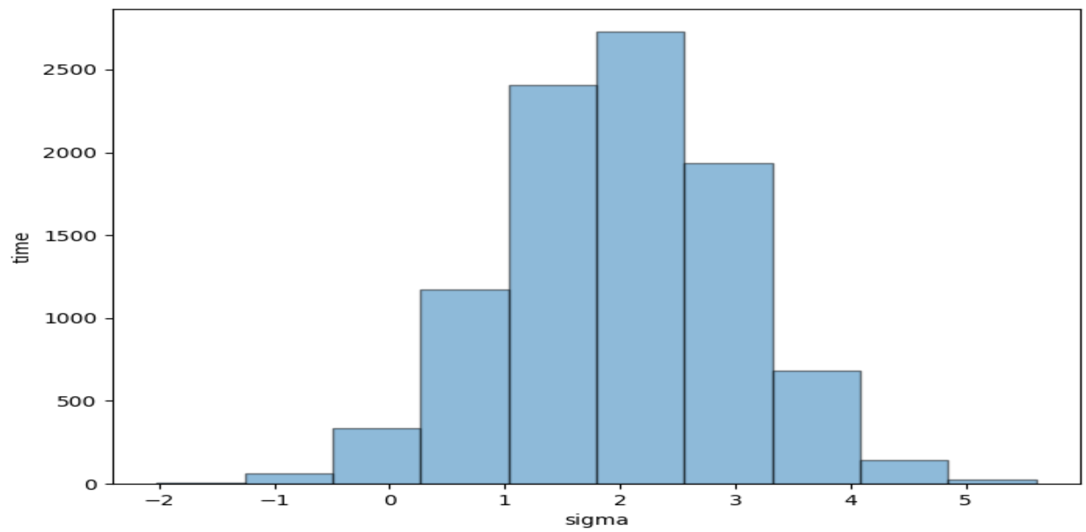
```
Command Prompt

C:\Users\777\Desktop\SEM2\AI\2020CSM1012_KirandeepKaur_Assignment2\question1>python main.py

 please enter one of these values for a ( 0 or  0.99) :
0
Estimated miu  :  1.5122662416805874
Estimated sigma:  [[ 0.3823242   0.15527797  0.43520511 ...  0.55098564 -0.26170752
  -0.19751876]
 [ 0.15527797  0.06306493  0.17675514 ...  0.22377849 -0.10629045
  -0.08022069]
 [ 0.43520511  0.17675514  0.49540021 ...  0.62719484 -0.29790541
  -0.22483843]
 ...
 [ 0.55098564  0.22377849  0.62719484 ...  0.79405168 -0.37715919
  -0.2846537 ]
 [-0.26170752 -0.10629045 -0.29790541 ... -0.37715919  0.17914332
   0.135205  ]
 [-0.19751876 -0.08022069 -0.22483843 ... -0.2846537   0.135205
   0.1020434 ]]

C:\Users\777\Desktop\SEM2\AI\2020CSM1012_KirandeepKaur_Assignment2\question1>
```

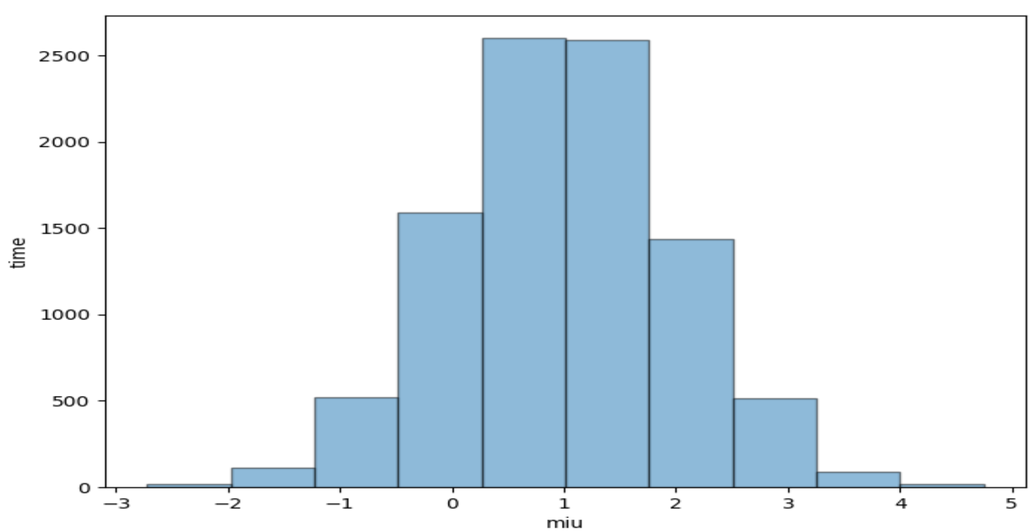
The following traceplot gives the trace of μ and Σ on the y axis with time on the x axis .



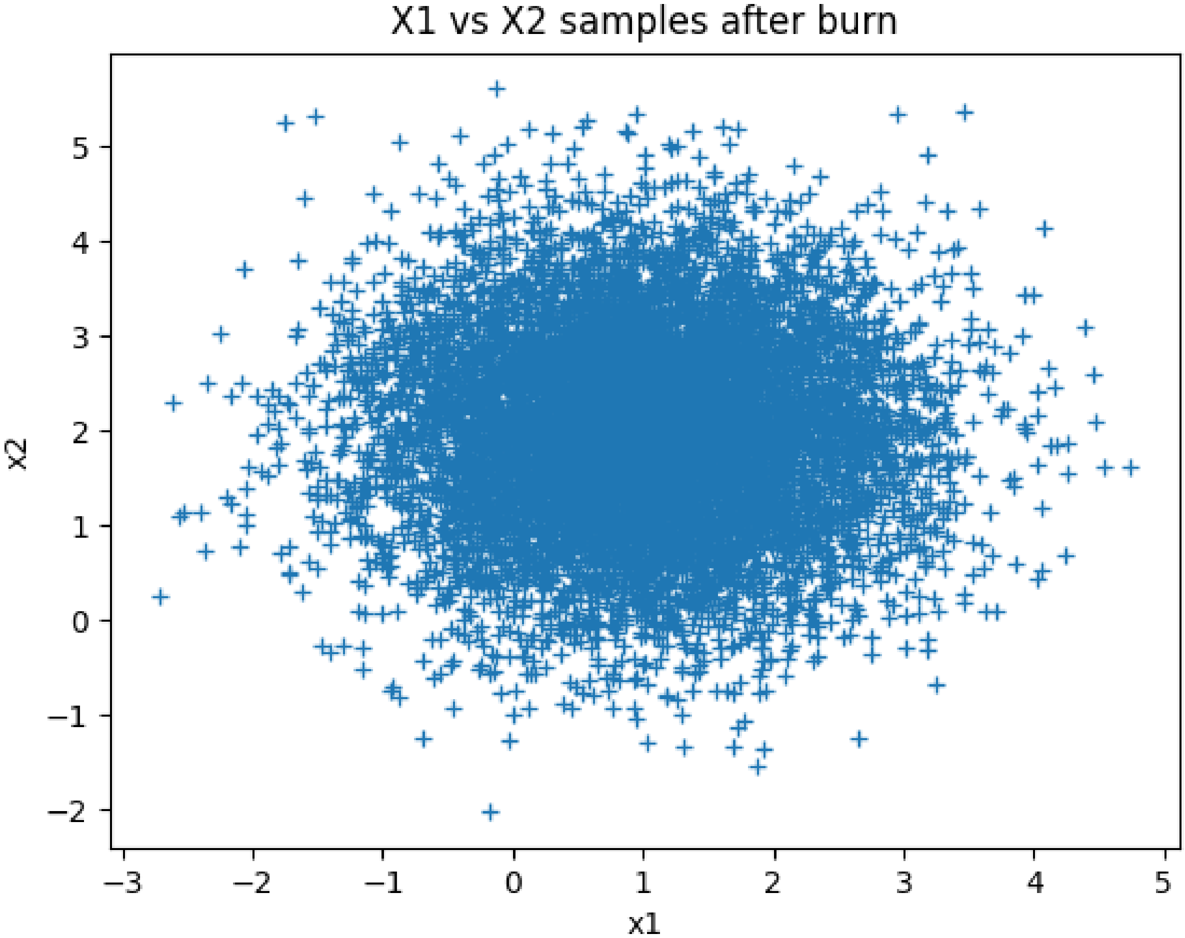
For the first 200 (appx) iterations or seconds, the graph shows a lot of fluctuations. And after that, it seems to converge. Hence we can discard the first 200 readings

The plot below shows the values of





sigma and miu after burning or discarding first 200 values.

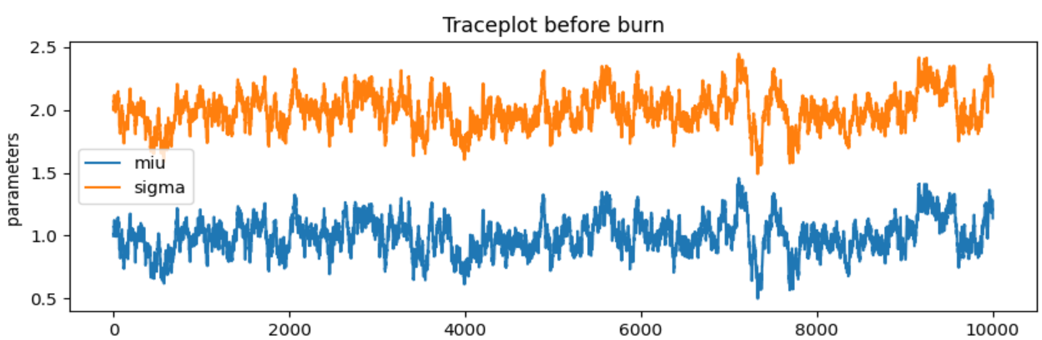


The graph above shows the values of X1 and X2 sampled resulting in a Gaussian distribution.

D)
a = 0.99
n(iterations) = 10000
The following were the first 5 estimated values (out of 10,000)for Σ and μ.

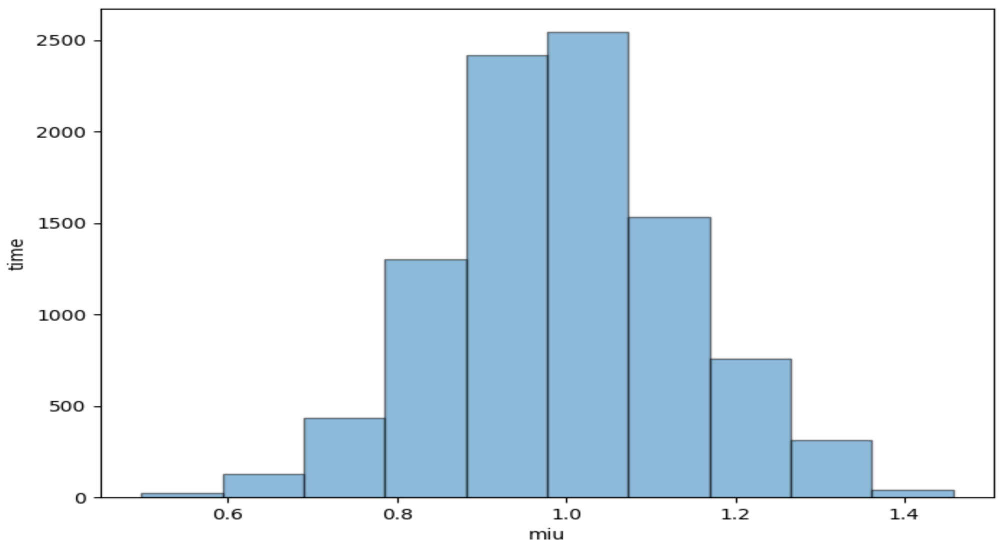
```
Command Prompt - python  main.py

C:\Users\777\Desktop\SEM2\AI\2020CSM1012_KirandeepKaur_Assignment2\question1>python main.py

 please enter one of these values for a ( 0 or  0.99) :
0.99
Estimated miu  :  1.510642586865771
Estimated sigma:  [[0.480147   0.49487466 0.5005549   ... 0.47770831 0.48855756 0.49626382]
 [0.49487466 0.51005406 0.51590854 ... 0.49236116 0.5035432  0.51148583]
 [0.5005549  0.51590854 0.52183021 ... 0.49801256 0.50932295 0.51735675]
 ...
 [0.47770831 0.49236116 0.49801256 ... 0.475282   0.48607615 0.49374327]
 [0.48855756 0.5035432  0.50932295 ... 0.48607615 0.49711545 0.5049567 ]
 [0.49626382 0.51148583 0.51735675 ... 0.49374327 0.5049567  0.51292163]]

C:\Users\777\Desktop\SEM2\AI\2020CSM1012_KirandeepKaur_Assignment2\question1>python main.py

 please enter one of these values for a ( 0 or  0.99) :
```

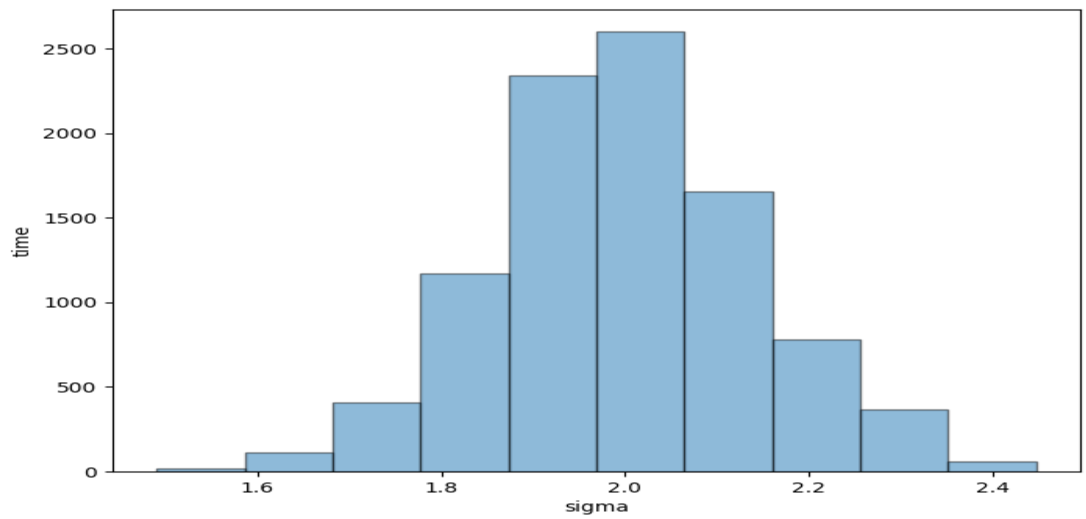
The traceplot below gives the trace of μ and Σ on the y axis with time on the x axis .



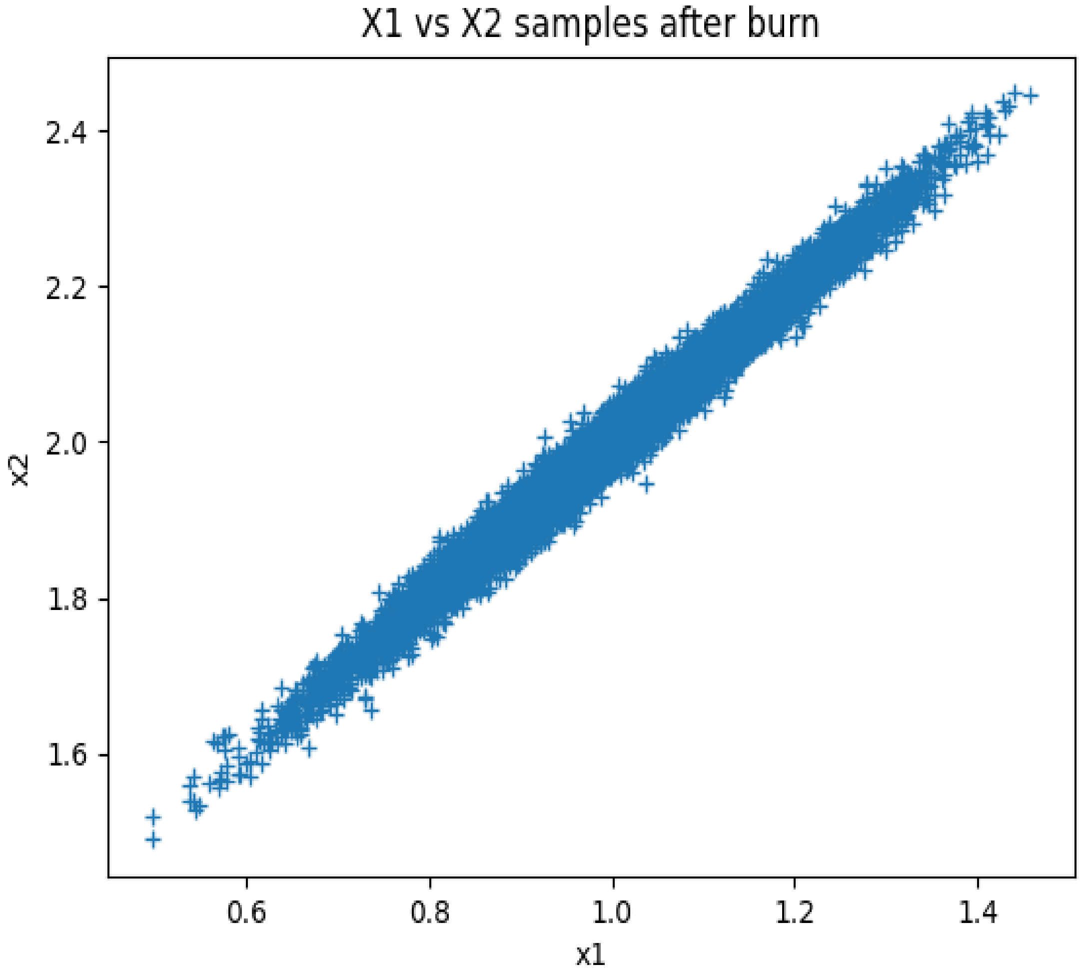
For the first 500 iterations or seconds, the graph shows a lot of fluctuations. And after that, it seems to converge. Hence we can discard the first 500 readings

This plot shows the values of sigma and miu after burning or discarding first 200 values.





This plot shows the samples taken directly from the bivariate Gaussian distribution $N_2(\mu, \Sigma)$.



After burning the discarded values, the values in the result seem to converge upto an extent. The number of discarded values reduced on increasing the value of a in Σ.