CS M146 605796227

Homework 0 KRISH PATEL

Problem 1) $y = x\sin(z)e^{-x}$

$$\frac{dy}{dx} = \sin(z)e^{-x} - x\sin(z)e^{-x}$$

Problem 2) $X = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$ $y = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ $z = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

i) $y^T z = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = [2+9] = [11]$

ii) $Xy = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2+12 \\ 1+9 \end{bmatrix} = \begin{bmatrix} 14 \\ 10 \end{bmatrix}$

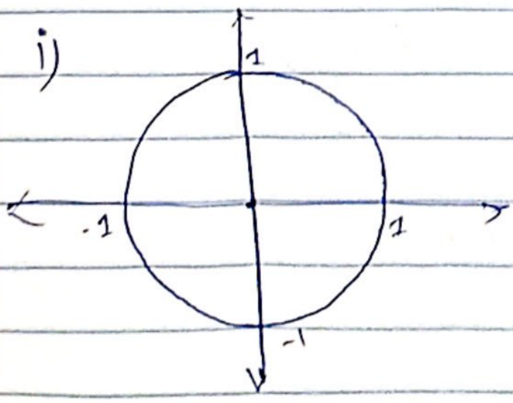
iii) determinant $= 6 - 4 = 2$ Non zero hence invertible.

inverse $= \frac{1}{2}\begin{bmatrix} 3 & -4 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1.5 & -2 \\ -0.5 & 1 \end{bmatrix}$

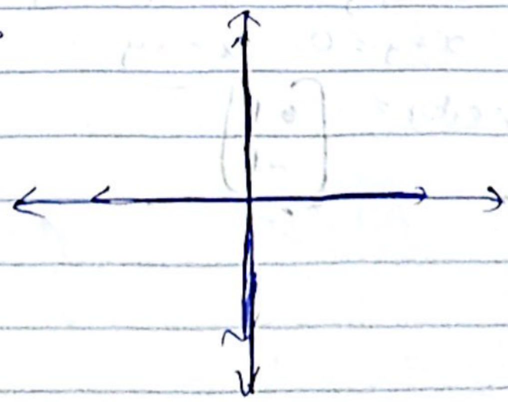
iv) The rank of X is : $\begin{vmatrix} 2 & 4 \\ 1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix}_{-R_1} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix}^{-R_2}$

$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ = Rank = 2

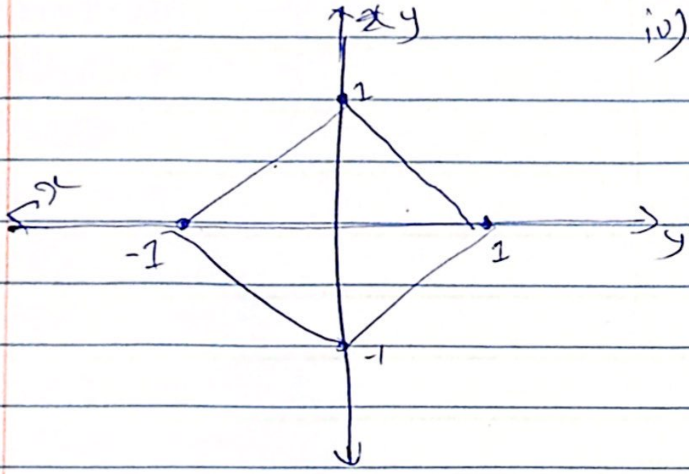
b) i)



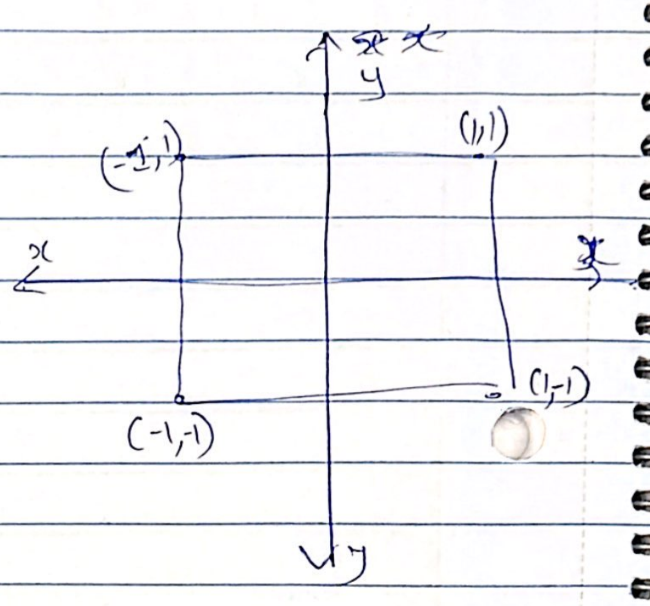
ii)



b) iii)



iv)



c) i) ~~Eigenvalues: For a g. square matrix, there exist~~

Eigenvalues: For a square matrix, there exists a vector/s ~~eigenvec~~ such that when multiplied with the matrix, the ~~produces~~ a ~~resul~~ product ~~gives~~ a is a multiple of the vector. The multiple is known as the eigenvalue, and the vector/s associated with it is known as a eigenvector.

ii) $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $\quad \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$

$(2-\lambda)^2 - 1 = 0$

$\lambda^2 - 4\lambda + 3 = 0$

$\lambda^2 - 3\lambda - \lambda + 3 = 0$

$(\lambda-3)(\lambda-1) = 0 \quad \lambda = 3, \lambda = 1$

plugging in ~~valu~~ values

$\lambda = 1$ $\begin{bmatrix} 2-1 & 1 \\ 1 & 2-1 \end{bmatrix}$

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \times 1$

$\begin{bmatrix} 2x+y \\ x+2y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

$2x + y = x \quad x + y = 0$

$x + y = 0 \quad x = -y$

eigenvector = $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix}$

$\begin{bmatrix} 2x+y \\ 2y+x \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix}$

$x = y$ $\begin{bmatrix} x = y \\ y = x \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ → eigenvector

Scanned with CamScanner

iii) ~~A^k~~ $Av = \lambda v$ → where λ is eigenvalue, v is eigenvector

$A^k(v) = (\lambda v)^k$ ~~$\neq$~~ $\Lambda^k v = \lambda^k v$ → Associative property

Thus, v_i is still an ~~of~~ eigenvector of A_k with eigenvalue λ^k

$\det(Ak - \lambda I) = 0$ $A^k = PDP'$ (SVD)

$\det(D - \lambda I) = 0$. Solving this gives the polynomial $(\lambda_1^k - \lambda)(\lambda_2^k - \lambda)(\lambda_3^k - \lambda) \ldots = 0$

~~thus~~ Thus, eigenvalues are $\lambda_1^k, \lambda_2^k, \lambda_3^k$

d) i) ~~$a^T x$~~ derivative of $a^T x = a^T$

ii) derivative ~~of~~ of $x^T Ax$ $\frac{d}{dx}(x^T Ax) = Ax + Ax^T$
$= 2Ax$

Second derivative $= 2A$

Two vectors are orthogonal if their dot product is 0

e) i) ~~p~~, Lets take 2 points on the line, x_1 and x_2.

~~w^T~~ Thus $w^T(x_1) + b = 0$ and $w^T(x_2) + b = 0$

Subtracting → $w^T(x_1) - w^T(x_2) + \cancel{b} = 0$

$w^T(x_1 - x_2) = 0$

hence, as the dot product of ~~w^T~~ w^T and directional vector $(x_1 - x_2)$ is zero, thus ~~the~~ w is orthogonal to the line

ii) Distance from origin to line is $\frac{w^T x + b}{\|w\|_2}$ when x is origin, $w^T x = 0$

$= \frac{b}{\|w\|_2}$ hence distance is $\frac{b}{\|w\|_2}$

Scanned with CamScanner

Q3) a) i) The sample mean for the data = $\frac{3}{5} = \underline{0.6}$

ii) unbiased sample variance

$$= \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

$$= \{3(1-0.6)^2 + 2(0-0.6)^2\}/4$$

$$= (0.6)^2 \times 2 + 3 \times 0.4^2$$

$$= \frac{0.72 + 0.48}{4} = \frac{1.20}{4} = 0.3$$

iii) $0.5 \times 0.5 \times 0.5 \times 0.5 \times 0.5$

$$= 0.03125$$

iv) $x \times x \times (1-x) \times (x) \times (1-x)$

$$= x^3 \times (x-1)^2$$

Need to find ~~minima~~ maxima.

$$\frac{d}{dx} = 3x^2(x-1)^2 + 2(x-1)x^3 = 0$$

$$= x = 0,\ 0.6,\ 1$$

at 0 and 1, the probability will become zero.
hence, best probability is 0.6

$P(x=1) = 0.6$, $P(X=0) = \underline{0.4}$

Scanned with CamScanner

v) $P(X = T \text{ when } Y = b)$

$$= \frac{P(X=T) \cancel{\times P(Y=b)}}{P(Y=b)} = \frac{0.1}{0.1+0.15} = \frac{0.1}{0.25} = 0.4$$

b) a) Guassian $\rightarrow \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)}$

b) Binomial $= \binom{n}{x} p^x (1-p)^{n-x}$

c) Exponential $= \lambda e^{-\lambda x}$ when $x \geq 0$, 0 otherwise.

d) Bernouilli $= p^x (1-p)^{1-x}$, when $x \in \{0,1\}$; 0, otherwise

~~b)~~ e) Uniform $= \frac{1}{b-a}$ when $a \leq x \leq b$; 0 otherwise.

~~e) Un.~~

c) Mean ~~of~~ of Bernouilli $= E[x] = p$

Variance $=$ $= p(1-p)$

d) ~~of~~ If the variance of $x = \sigma^2$, the variance of $2x$ is $(2)^2 \times \cancel{2}\sigma^2 = 4\sigma^2$

Variance of $x+2$ is the same at x (adding a constant doesn't change S.D. thus σ^2

Scanned with CamScanner

Problem 4) a) $f(n) = \ln(n)$ $\quad g(n) = \lg(n)$

$\ln n = \frac{\log_2 n}{\log_2 e}$ $\quad f(n) = \frac{g(n)}{\log_2 e}$

$f(n) = O(g(n))$

$g(n) = O(f(n))$

b) $f(n) = 3^n$, $g(n) = n^{10}$

as n increases, the gradient of 3^n grows much rapidly than n^{10}

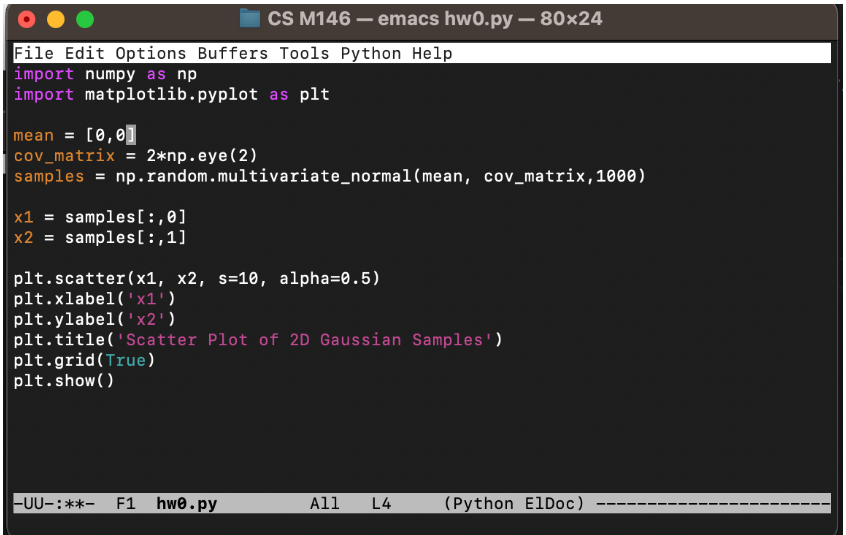
therefore $g(n) = O(f(n))$

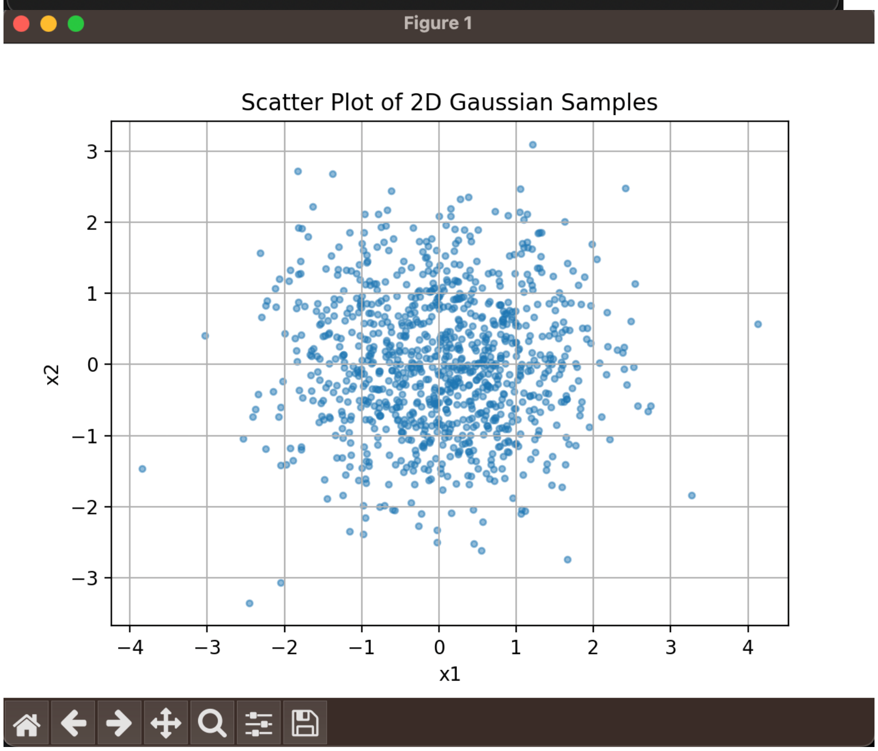
c) ~~$3^n \geq 2^n$~~ $f(n) = 3^n$ $\quad g(n) = 2^n$

$3^n \geq 2^n$ for all $n \geq 1$.

Thus, $g(n) = O(f(n))$

Question 5 a)



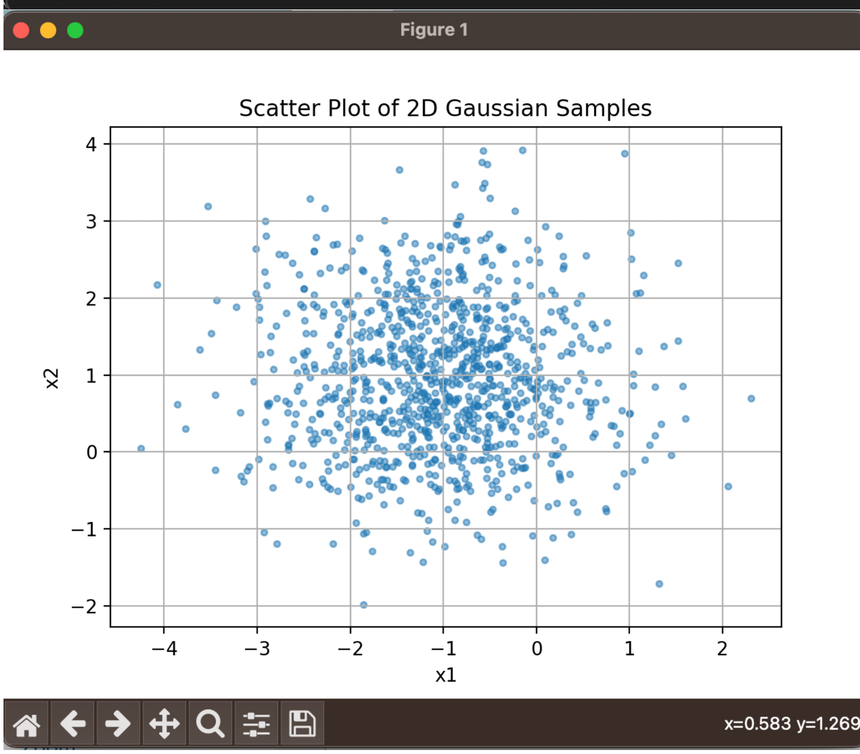


In this, it is centered around the origin(which is the mean)

Question 5 a)

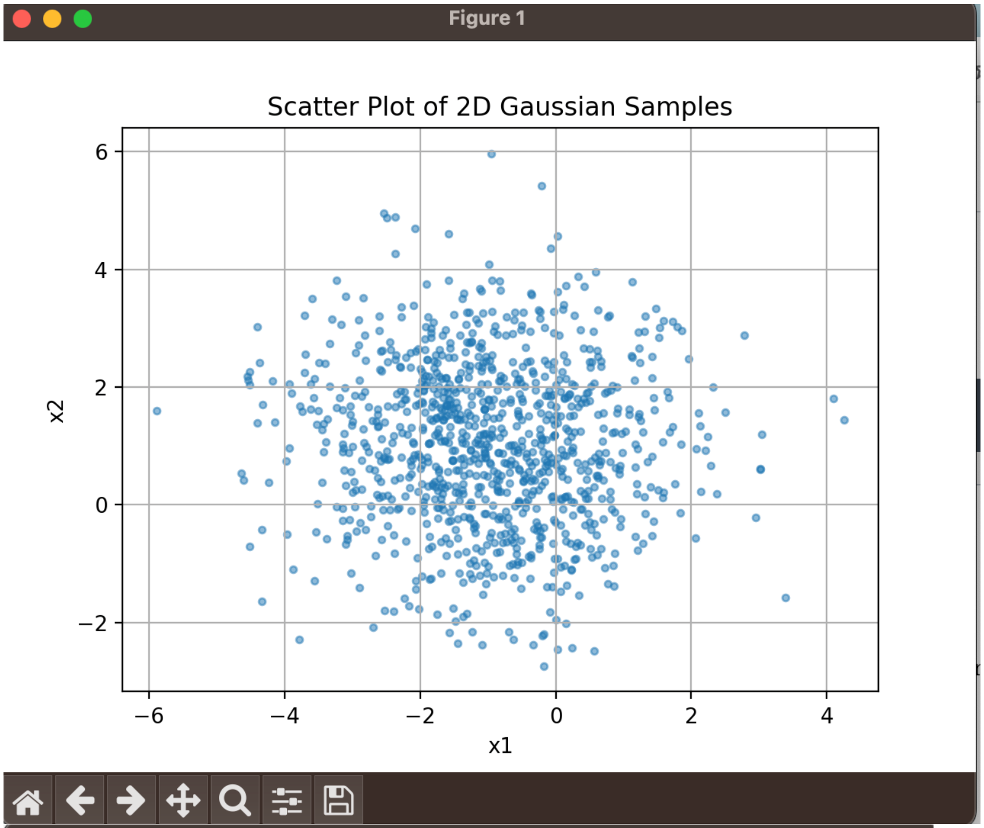
CS M146 — emacs hw0.py — 80×24

```
File Edit Options Buffers Tools Python Help
import numpy as np
import matplotlib.pyplot as plt

mean = [-1,1]
cov_matrix = np.eye(2)
samples = np.random.multivariate_normal(mean, cov_matrix,1000)

x1 = samples[:,0]
x2 = samples[:,1]

plt.scatter(x1, x2, s=10, alpha=0.5)
plt.xlabel('x1')
plt.ylabel('x2')
plt.title('Scatter Plot of 2D Gaussian Samples')
plt.grid(True)
plt.show()

-UU-:---  F1  hw0.py          All   L5     (Python ElDoc) ----------------------
```



In this graph, the points are centered around [-1,1]

Question 5 a)

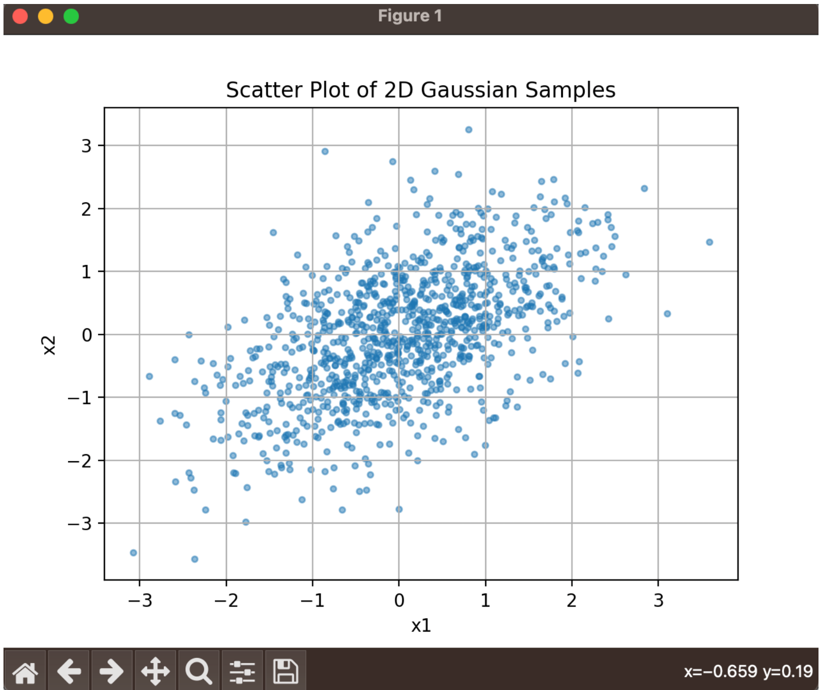


CS M146 — emacs hw0.py — 80×24

```
File Edit Options Buffers Tools Python Help
import numpy as np
import matplotlib.pyplot as plt

mean = [-1,1]
cov_matrix = 2*np.eye(2)
samples = np.random.multivariate_normal(mean, cov_matrix,1000)

x1 = samples[:,0]
x2 = samples[:,1]

plt.scatter(x1, x2, s=10, alpha=0.5)
plt.xlabel('x1')
plt.ylabel('x2')
plt.title('Scatter Plot of 2D Gaussian Samples')
plt.grid(True)
plt.show()




-UU-:---  F1  hw0.py        All    L1      (Python ElDoc) ----------------------
Can't guess python-indent-offset, using defaults: 4
```

According to this, the scatter is increased(the distance from the center)

Question 5 a)

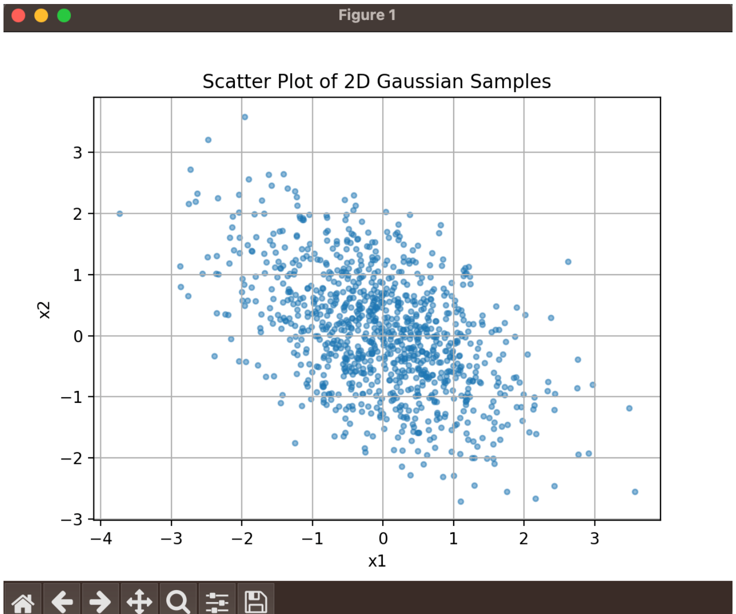


CS M146 — emacs hw0.py — 80×24

```
File Edit Options Buffers Tools Python Help
import numpy as np
import matplotlib.pyplot as plt

mean = [0,0]
cov_matrix = np.array([[1,0.5],[0.5,1]])
samples = np.random.multivariate_normal(mean, cov_matrix,1000)

x1 = samples[:,0]
x2 = samples[:,1]

plt.scatter(x1, x2, s=10, alpha=0.5)
plt.xlabel('x1')
plt.ylabel('x2')
plt.title('Scatter Plot of 2D Gaussian Samples')
plt.grid(True)
plt.show()




-UU-:---  F1  hw0.py        All   L1      (Python ElDoc) ----------------------
Can't guess python-indent-offset, using defaults: 4
```

According to this, the spread is more around the plane x1= x2 and reduced around the plane x1= -x2

Question 5 a)



CS M146 — emacs hw0.py — 80×24

```
File Edit Options Buffers Tools Python Help
import numpy as np
import matplotlib.pyplot as plt

mean = [0,0]
cov_matrix = np.array([[1,-0.5],[-0.5,1]])
samples = np.random.multivariate_normal(mean, cov_matrix,1000)

x1 = samples[:,0]
x2 = samples[:,1]

plt.scatter(x1, x2, s=10, alpha=0.5)
plt.xlabel('x1')
plt.ylabel('x2')
plt.title('Scatter Plot of 2D Gaussian Samples')
plt.grid(True)
plt.show()

-UU-:---  F1  hw0.py          All   L1     (Python ElDoc) ----------------------
Can't guess python-indent-offset, using defaults: 4
```

The direction of the spread changes(almost like a 90 degree rotation from the previous question.

b)
i)Name of the dataset: Electric Vehicle Population Data
link for the data: https://catalog.data.gov/dataset/electric-vehicle-population-data
ii)The features of the dataset include
The electric vehicle type(specifies the type of electric vehicle, clean alternative fuel vehicle eligibility which has information regarding the vehicles' CAFV eligibility, the electric range in terms of miles that it can drive. Other features involve: Base MSRP, Legislative District, DOL Vehicle ID, Vehicle Location, Electric Utility, 2020 Census Tract.

The number of examples in the dataset are: 492 million cars

The number of features in the dataset is the number of columns which is 10 features.