ECE 595: Homework 1

Rahul Deshmukh, Class ID: 58

(Spring 2019)

Exercise 2

2a

2(a)

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad x \in (-\infty, \infty)$$

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x)\, dx = \int_{-\infty}^{+\infty} \frac{(x-\mu+\mu)}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{\sigma^2}{\sqrt{2\pi\sigma^2}} \left[\int_{-\infty}^{+\infty} \frac{(x-\mu)}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \int_{-\infty}^{+\infty} \frac{\mu}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right]$$

As $f_X(x)$ is PDF $\Rightarrow \int_{-\infty}^{+\infty} f_X(x) = 1 \Rightarrow \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \sqrt{2\pi\sigma^2}$

$$\Rightarrow E[X] = \frac{\sigma^2}{\sqrt{2\pi\sigma^2}} \left[\int_{+\infty}^{+\infty} e^{-t} dt + \frac{\mu}{\sigma^2} \times \sqrt{2\pi\sigma^2} \right], \text{ where } t = \frac{(x-\mu)^2}{2\sigma^2}$$

$$= \frac{\sigma^2}{\sqrt{2\pi\sigma^2}} \cancel{-e^{-t}\Big|_{+\infty}^{+\infty}} + \mu \qquad \Rightarrow E[X] = \mu$$

$$Var[X] = \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)\, dx = \int_{-\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $Z = \frac{x-\mu}{\sigma}$ $\Rightarrow dz = \frac{dx}{\sigma}$

$$\Rightarrow Var[X] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \sigma^2 z^2 e^{-\frac{z^2}{2}} (\sigma dz)$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z \cdot z e^{-z^2/2} dz$$

using Integration By parts

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left[\left[-z \cdot e^{-z^2/2} \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -e^{-z^2/2} dz \right]$$

For First term, the exponential term will go to zero faster than the z term

$$\Rightarrow Var[X] = \frac{\sigma^2}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{+\infty} e^{-z^2/2} dz}_{I}$$

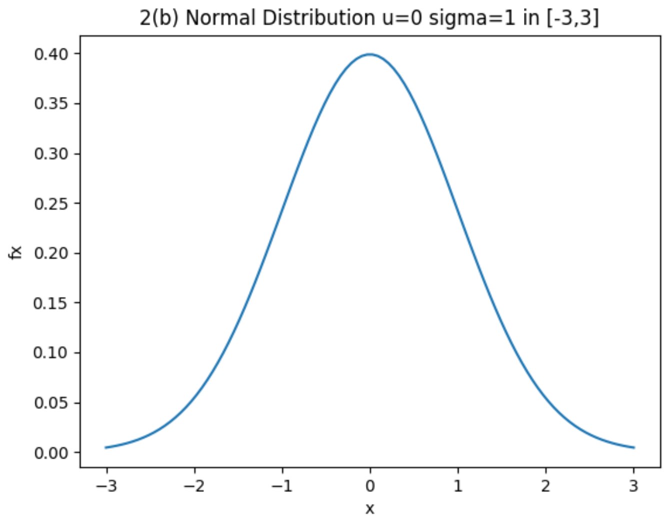
$$I^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-z^2/2} e^{-t^2/2} dz\,dt = \lim_{r \to \infty} \int_0^{2\pi}\int_0^{r} e^{-r^2/2} r\,dr\,d\theta$$

with $r^2 = z^2 + t^2$

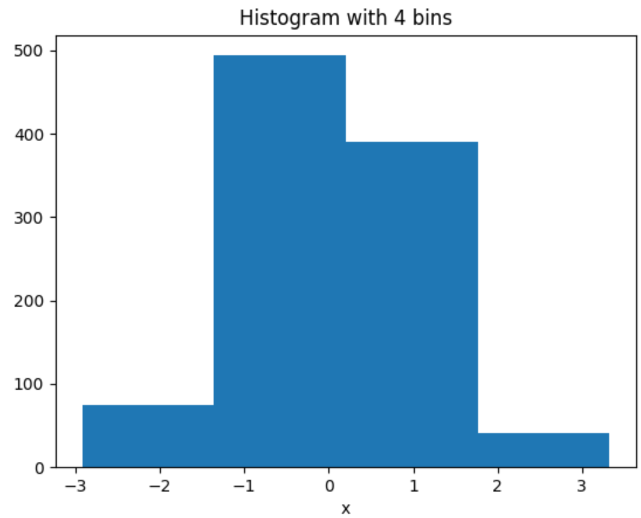
$$\Rightarrow I^2 = 2\pi \; -e^{-s}\Big|_0^{\infty} = 2\pi \quad \Rightarrow I = \sqrt{2\pi}$$

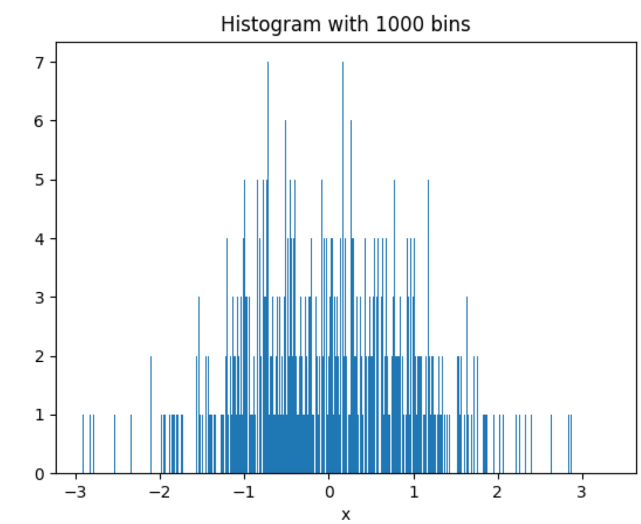
$$\Rightarrow Var[X] = \frac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = \sigma^2$$

2b Given: $\mu = 0$ and $\sigma = 1$



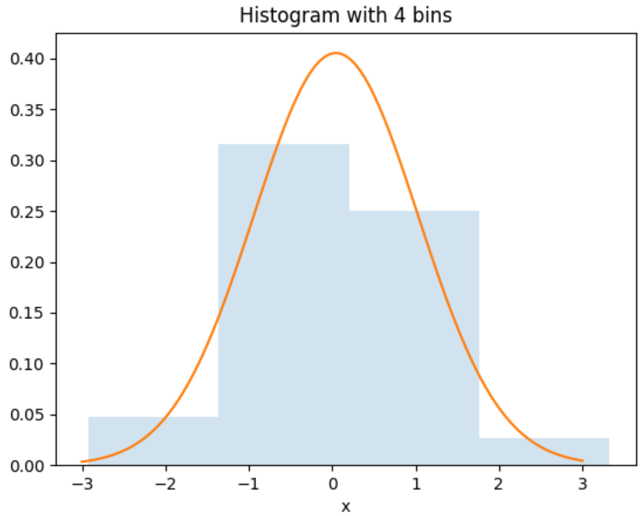
2c
ii

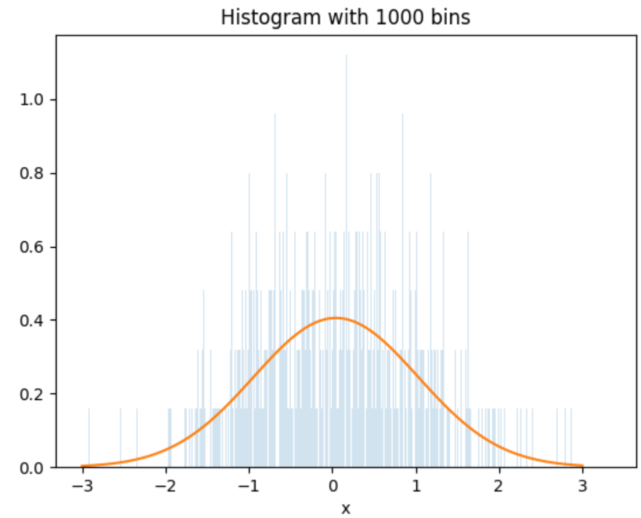




iii The estimated mean and standard deviation from the samples came out to be: $\mu = 0.044773922737663491$ and $\sigma = 0.983857216790902$

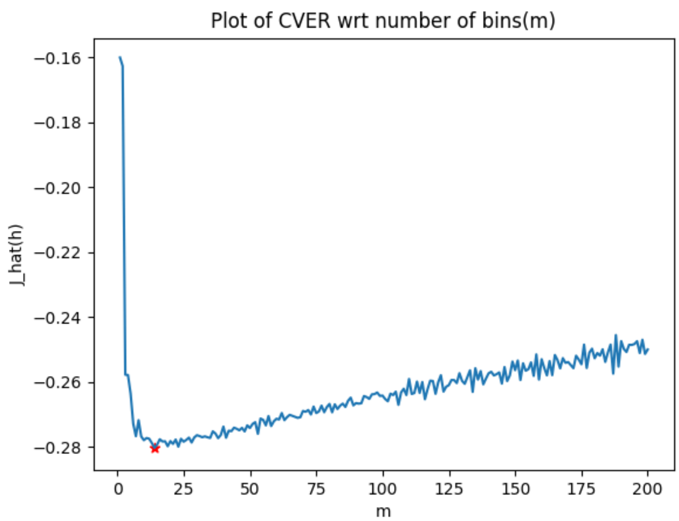
iv

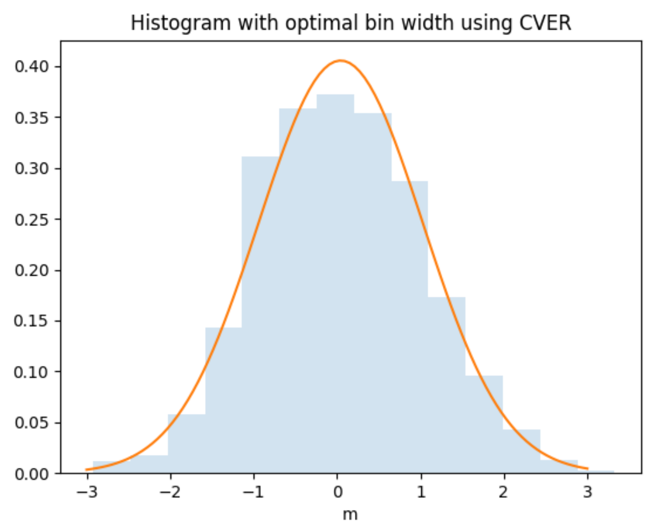




v From the above plots we can see that neither of the the histograms match the fitted Gaussian accurately. Therefore we need to find optimal bin width to obtain a better histogram.

2d From Computation m^* came out as **14**





On Comparing the above histogram with the previous histograms we can clearly observe that when using the optimal bin width we have a better histogram which represents the fitted Gaussian more accurately.

The code for this Problem is:

```
"""
ECE 595: Machine Learning-I
HW-1: Ex 2
@author: rahul
"""
# import libraries
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

#%% Exercise 2
#%% 2b
u = 0
sigma= 1
numpts = 100
lb=-3;ub=3;
x = np.linspace(lb,ub,numpts)
fx = (1/np.sqrt(2*np.pi*sigma**2))*np.exp(-(x-u)**2/(2*sigma**2))
plt.plot(x,fx)
plt.xlabel('x')
plt.ylabel('fx')
plt.title('2(b) Normal Distribution u=0 sigma=1 in [-3,3]')
plt.savefig('2b_fig')
#%% 2c
# i
n=1000
s = np.random.normal(0,1,n)
# ii
bins = 4
plt.figure()
plt.hist(s,bins)
plt.xlabel('x')
plt.title('Histogram with '+str(bins)+' bins')
plt.savefig('2c1_fig')
bins=1000
plt.figure()
plt.hist(s,bins)
plt.xlabel('x')
plt.title('Histogram with '+str(bins)+' bins')
plt.savefig('2c2_fig')
# iii
u_est, sig_est = norm.fit(s)
print('estimated u:'+str(u_est)+' estimated sigma:'+str(sig_est))
# iv
bins = 4
plt.figure()
plt.hist(s,bins,normed=True,alpha=0.2)
plt.plot(x,norm.pdf(x,u_est,sig_est))
plt.xlabel('x')
plt.title('Histogram with '+str(bins)+' bins')
plt.savefig('2c3_fig')
bins = 1000
plt.figure()
plt.hist(s,bins,normed=True,alpha=0.2)
plt.plot(x,norm.pdf(x,u_est,sig_est))
plt.xlabel('x')
```

```python
plt.title('Histogram with '+str(bins)+' bins')
plt.savefig('2c4_fig')
#%% 2d
Jh = []
range_s = max(s)-min(s)
for m in range(1,200+1):
    h = range_s/m
    temp = 2/(h*(n-1))-((n+1)/(h*(n-1)))*\
    np.sum(np.square(np.histogram(s,bins=m)[0]/n))
    Jh.append([temp])

m_star = np.argmin(Jh)+1
print(m_star)
plt.figure()
plt.plot(np.arange(1,200+1),Jh)
plt.scatter(m_star,Jh[m_star-1],c='r',marker='*')
plt.xlabel('m')
plt.ylabel('J_hat(h)')
plt.title('Plot of CVER wrt number of bins(m)')
plt.savefig('2c5_fig')
plt.figure()
plt.hist(s,m_star,normed=True,alpha=0.2)
plt.plot(x,norm.pdf(x,u_est,sig_est))
plt.xlabel('m')
plt.title('Histogram with optimal bin width using CVER')
plt.savefig('2c6_fig')
```

Exercise 3

3a
i

3(a) $\mu = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$$\Rightarrow \Sigma^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad |\Sigma| = 3$$

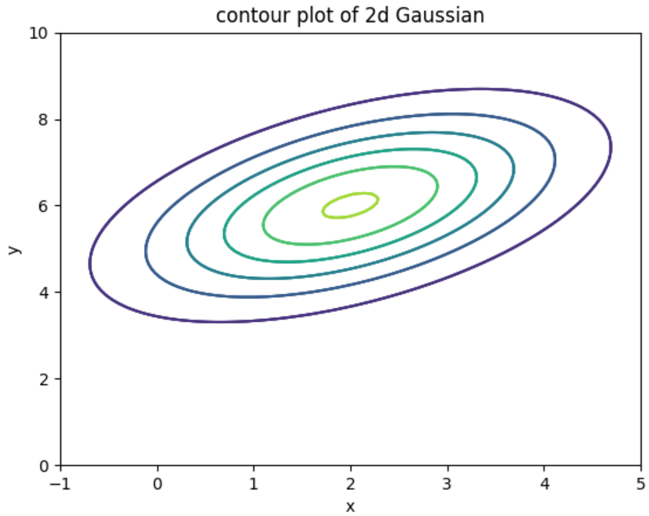
$$f_X(\underline{x}) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp\left\{ \frac{-1}{2} (\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu}) \right\}$$

$$= \frac{1}{\sqrt{4\pi^2 \times 3}} \exp\left\{ -\frac{1}{2}\cdot\frac{1}{3} [x_1-2, x_2-6] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1-2 \\ x_2-6 \end{bmatrix} \right\}$$

$$= \frac{1}{\sqrt{12\pi^2}} \exp\left\{ -\frac{1}{6} \left(2(x_1-2)^2 + 2(x_2-6)^2 - 2(x_1-2)(x_2-6) \right) \right\}$$

$$f_X(\underline{x}) = \frac{1}{\sqrt{12\pi^2}} \exp\left\{ \frac{-1}{6} \left(2x_1^2 + 2x_2^2 - 2x_1x_2 + 4x_1 - 20x_2 + 56 \right) \right\}$$

ii



3b
i

$\boxed{3(b)}$ given $X \sim N(\underline{0}, I) \Rightarrow E[X] = \underline{0}$ $E[XX^T] = I$

(i) given: $A \in R^{d \times d}$, $b \in R^d$

$Y = AX + b$

$\mu_y = E[Y] = E[AX + b] = E[AX] + E[b]$

$= A\,E[X] + b$

$= A \cdot \underline{0} + b = b$

$\Rightarrow \boxed{\mu_y = b}$

$\Sigma_y = E[(Y-\mu_y)(Y-\mu_y)^T] = E[(AX+b-b)(AX+b-b)^T]$

$= E[(AX)(AX)^T] = E[AXX^TA^T] = A\,E[XX^T]\,A^T$

$= A I A^T = A A^T$

$\Rightarrow \boxed{\Sigma_y = A A^T}$

(ii),(iii),(iv)

(ii) given: $\Sigma_y = A A^T$

To prove: Σ_y is symmetric & semi-pos-def

$$\Sigma_y^T = (AA^T)^T = (A^T)^T A^T = AA^T = \Sigma_y$$

$\Rightarrow \Sigma_y$ is symmetric

Now let $X \in R^2$

$$X^T \Sigma_y X = X^T A A^T X = (A^T X)^T (A^T X)$$

$$= \|A^T X\|_2^2 \geq 0$$

$\Rightarrow X^T \Sigma_y X \geq 0 \quad \forall X \in R^2$

$\therefore \Sigma_y$ is semi-pos-def

(iii) Let $A = U \Lambda U^T$ where $UU^T = I$ $\Rightarrow \Sigma_y = (U\Lambda U^T)(U\Lambda U^T)^T$

$$= U \Lambda U^T U \Lambda U^T$$

$$= U \Lambda^2 U^T$$

for Σ_y to be strictly pos. def by def $\lambda_i^2 > 0 \quad \forall i$

(iv) $\mu_y = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \quad \Sigma_y = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$b = \mu_y = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \qquad \Sigma_y = A A^T \quad$ Let $\Sigma_y = U \Lambda U^T$

$$\Rightarrow \Sigma_y = U \Lambda^{1/2} \Lambda^{1/2} U^T$$

$$= U \Lambda^{1/2} U^T U \Lambda^{1/2} U^T \quad \text{as } U U^T = I$$

$$= (U \Lambda^{1/2} U^T)(U \Lambda^{1/2} U^T)^{TT}$$

$$\Rightarrow \Sigma_y = (U \Lambda^{1/2} U^T)(U \Lambda^{1/2} U^T)^T = A A^T$$

$$\Rightarrow A = U \Lambda^{1/2} U^T$$

Finding Eigenvalues of Σ_y, $\det(\Sigma_y - \lambda I) = 0$

$$\Rightarrow (2-\lambda)^2 - 1 = 0 \Rightarrow \lambda = 1, 3$$

for $\lambda = 1$, $(A - I) v_1 = 0$ $\qquad$ for $\lambda = 3$, $(A - 3I) v_2 = 0$

$\Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} v = 0 \qquad \Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} v_2 = 0$

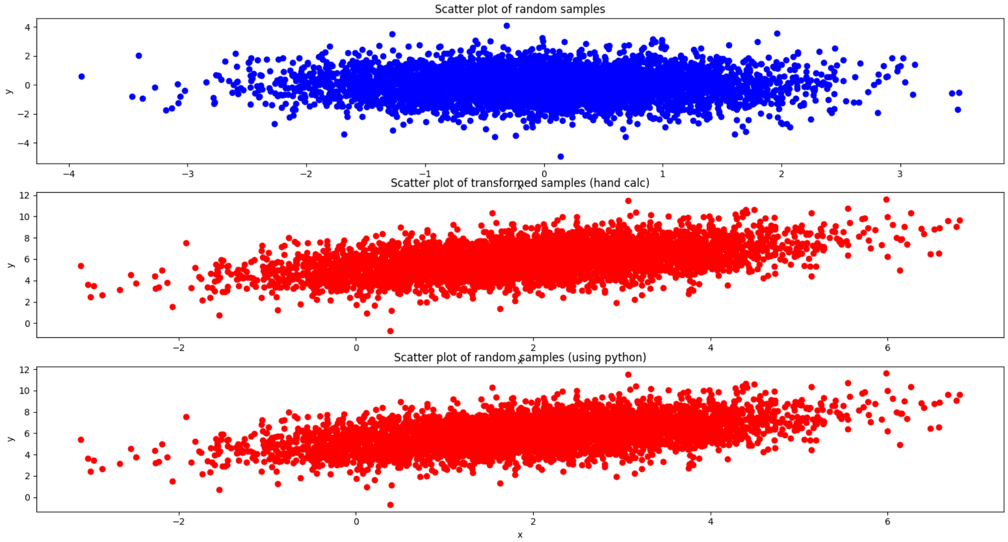
$\Rightarrow v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \Rightarrow v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\Rightarrow \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \quad U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\Rightarrow \quad A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & \sqrt{3} \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{bmatrix}$$

3c
(i),(ii)



iii We can observe that the scatter plots in the above sub-plots 2 and 3 are exactly similar. This indicates that the theoretical findings from part(b) match with the computation.

Also, if we compute the mean and co variance from the transformed samples they come out as:
$\mu_Y = \begin{bmatrix} 2.0019569 \\ 6.02267004 \end{bmatrix}$ and $\Sigma_Y = \begin{bmatrix} 2.00022568 & 0.99505734 \\ 0.99505734 & 2.00444455 \end{bmatrix}$

which is very close what is given in the problem statement.

The code for this problem is:

```
"""
ECE 595: Machine Learning-I
HW-1: Ex 3
@author: rahul
"""
import numpy as np
import matplotlib.pyplot as plt
#%% 3a
def gauss_2d(x,u,s):
    f = (1/np.sqrt((2*np.pi)**2*np.linalg.det(s)))*np.exp((-1/2)*(x-u).T@np.linalg.
        inv(s)@(x-u))
    return(f)

u=np.array([2,6])
sigma = np.array([[2,1],[1,2]])
N=100
x = np.linspace(-1,5,N)
y = np.linspace(0,10,N)
X,Y=np.meshgrid(x,y)

F = np.zeros((N,N))
for i in range(N):
    for j in range(N):
        x = np.array([X[i,j],Y[i,j]])
        F[i,j]= gauss_2d(x,u,sigma)

#plot
plt.figure(1)
plt.contour(X,Y,F)
plt.xlabel('x')
plt.ylabel('y')
plt.axis([-1,5,0,10])
plt.title('contour plot of 2d Gaussian')
plt.show()
plt.savefig('3a2')
#%% 3c
# i
N_sample = 5000
samples = np.random.multivariate_normal(np.zeros(2),np.eye(2),N_sample)
#plot
plt.figure(2)
plt.subplot(3,1,1)
plt.scatter(samples[:,0],samples[:,1],c='b',marker='o')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Scatter plot of random samples')
# ii
sqrt_3 = np.sqrt(3)
A =(1/2)*np.array([[sqrt_3+1,sqrt_3 -1],[sqrt_3 -1,sqrt_3 +1]])
b = np.array([2,6])
Y = ((A@samples.T).T+b)
plt.subplot(3,1,2)
plt.scatter(Y[:,0],Y[:,1],c='r',marker='o')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Scatter plot of transformed samples (hand calc)')
```

```
# iii
D,U = np.linalg.eig(sigma)
A_py = U@np.diag(np.sqrt(D))@U.T
Y_py = ((A_py@samples.T).T+b)
plt.subplot(3,1,3)
plt.scatter(Y_py[:,0],Y_py[:,1],c='r',marker='o')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Scatter plot of random samples (using python)')
# check
print('Sample observed mean:'+str(np.mean(Y,axis=0)))
print('Sample observed covariance:'+str(np.cov(Y.T)))
```

Exercise 4

(a),(b),(c)

4(a) given: $A \in \mathbb{R}^{m \times n}$ $x \in \mathbb{R}^m$ $y \in \mathbb{R}^n$

$$R = \max_j \sum_{k=1}^{n} |a_{jk}| \quad , \quad C = \max_k \sum_{j=1}^{m} |a_{jk}|$$

To prove: $|x^T A y| \leq \sqrt{RC}\, \|x\|_2 \, \|y\|_2$

Proof: $|x^T A y| = \left| \sum_{i=1}^{m} \sum_{j=1}^{n} x_i a_{ij} y_j \right|$

Let us flatten the double sum to a single sum using $k = (i-1)n + j$

$$\Rightarrow \text{L.H.S.} = \left| \sum_{k=1}^{mn} (x_i a_{ij} y_j)_k \right| = \left| \sum_{k=1} (\underbrace{x_i \sqrt{a_{ij}}}_{X_k} \, \underbrace{\sqrt{a_{ij}}\, y_j}_{Y_k})_k \right|$$

Now using Cauchy-Schwarz we get

$$\text{L.H.S.} \leq \left(\sqrt{\sum_k X_k^2} \right) \left(\sqrt{\sum_k Y_k^2} \right)$$

$$= \sqrt{\begin{pmatrix} x_1^2 (a_{11} + a_{12} + \cdots a_{1n}) + \\ x_2^2 (a_{21} + a_{22} + \cdots a_{2n}) + \\ \vdots \\ x_m^2 (a_{m1} + \cdots a_{mn}) \end{pmatrix}} \cdot \sqrt{\begin{pmatrix} (a_{11} + a_{21} + \cdots a_{m1})\, y_1^2 + \\ (a_{12} + a_{22} + \cdots a_{m2})\, y_2^2 + \\ \vdots \\ (a_{1n} + a_{12} + \cdots a_{mn})\, y_n^2 \end{pmatrix}}$$

Now by def, $\sum_{j=1}^{n} a_{ij} \leq R \quad \forall i \in \{1, 2 \cdots m\}$

& $\sum_{i=1}^{m} a_{ij} \leq C \quad \forall j \in \{1, \cdots n\}$

$$\Rightarrow L.H.S.^2 \leq \sqrt{R(x_1^2+x_2^2 \cdots + x_n^2)} \cdot \sqrt{C(y_1^2+y_2^2+\cdots y_n^2)}$$

$$= \sqrt{RC} \sqrt{x_1^2+x_2^2+\cdots x_n^2} \cdot \sqrt{y_1^2+y_2^2+\cdots y_n^2}$$

$$= \sqrt{RC} \, \|x\|_2 \|y\|_2 = R.H.S.$$

$\Rightarrow |x^T A y| \leq \sqrt{RC} \, \|x\|_2 \|y\|_2$ Hence Proved

$\boxed{4(b)}$ (i) given: A is pos-def

To prove: A is invertible

Proof: by def, $x^T A x > 0 \quad \forall x$

Let $A = U \Lambda U^T$ s.t. $UU^T = I$ & $\|u_i\| = 1$

$\Rightarrow x^T (U \Lambda U^T) x > 0 \quad \forall x$

Let x = first eigenvector $u_1 \Rightarrow U^T x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$

$\Rightarrow [1\ 0\ 0\ 0 \cdots] \Lambda \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix} > 0 \Rightarrow \lambda_1 > 0$

$\|^{ly}$ $\lambda_i > 0 \quad \forall i = 1-n$

Now $\det(A) = \prod_{i=1}^{n} \lambda_i > 0$ as $\lambda_i > 0 \ \forall i = 1 \to n$

$\Rightarrow \det(A) > 0$

$\therefore$ A is invertible

(9ii) given: Hessian is invertible i.e. $\det(H) \neq 0$

& Hessian is not positive definite anywhere

$\Rightarrow \lambda_i < 0$

To find: $f \to \mathbb{R}^2 \to \mathbb{R}$ which has the above Hessian

One such hessian is $H = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ $\det(H) = -3$

& $\lambda(H) = -1, -3$

$$f(\underline{x}) = x^T\left(\frac{1}{2}H\right)x = \frac{1}{2}[x_1, x_2]\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{x_1^2 + x_2^2 + 4x_1x_2}{2}$$

$$\Rightarrow \boxed{f(\underline{x}) = \frac{x_1^2 + x_2^2 + 4x_1x_2}{2}}$$

<u>Check</u>

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

(iii) given: A s.t. $x^T A x \geq 0 \;\; \forall\, x \in \mathbb{R}^n$

To find: condition for $x^T A x > 0 \;\; \forall\, x \in \mathbb{R}^n$

Proof: Let us do the eigen decomposition of A

$$A = U \Lambda U^T \quad \text{s.t.} \quad UU^T = I$$

$$\Rightarrow \quad x^T (U \Lambda U^T) x \geq 0 \quad \forall\, x \in \mathbb{R}^n \quad — ①$$

Let $x = u_i$ the ith eigen vector

$$\Rightarrow \quad U^T x = U^T \underline{u}_i = [1\ 0 \cdots 0]^T \quad — ②$$

From ① & ② we get

$$\lambda_i \geq 0 \quad \forall\, i = 1 \dots n$$

Now,

for $x^T A x > 0$ we would get $\lambda_i > 0$

$$\left.\begin{array}{l} \Rightarrow \quad \lambda_i \neq 0 \\ \text{Equivalently} \quad \det(A) = \prod_{i=1}^{n} \lambda_i \neq 0 \end{array}\right\} \text{Additional condition on } A$$

4(c) To prove: $A A^{+} A = A$

Using Hint,

$$A = U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^T \quad \text{where} \quad \Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \quad \lambda_i \neq 0$$

with $U U^T = I$

$$A^{+} = U \Lambda^{-} U^T \quad \text{where} \quad [\Lambda^{-}]_{j,j} = \begin{cases} [\Lambda]^{-1}_{j,j} & \forall\ 1 \le j \le k \\ 0 & \text{elsewhere} \end{cases}$$

$$\Rightarrow A A^{+} A = \left(U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^T\right)\left(U \Lambda^{-} U^T\right)\left(U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^T\right)$$

$$= U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Lambda^{-} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^T$$

$$= U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Lambda^{-}\Lambda & 0 \\ 0 & 0 \end{bmatrix} U^T = U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} U^T$$

$$= U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^T = A = \text{R.H.S.}$$

hence proved