1

10-601

Machine Learning

HW #1

MAHER KHAN (mhkhan)

1.a) Let the event that I want to go up be $\uparrow Me$

Let the event that I want to go down be $\downarrow Me$

Let the event that the lift is going down be $\downarrow Ele$

" " " " " " " " up be $\uparrow Ele$

Let the number of elevators be m

Now, the required probability is equivalent to :

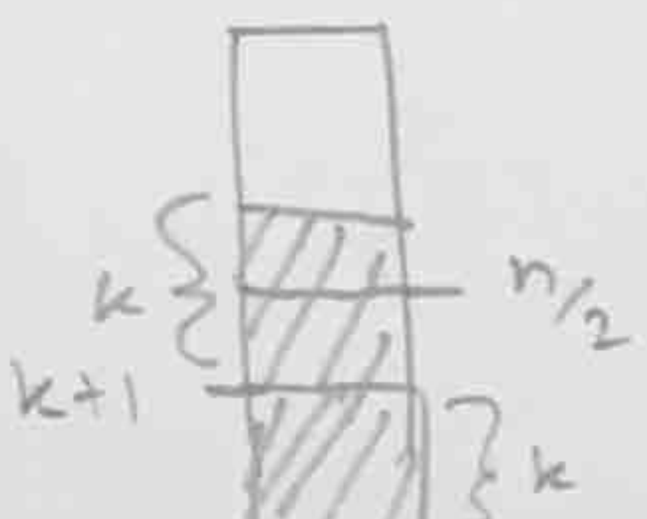
$$P\left((\uparrow Me \wedge \uparrow Ele) \vee (\downarrow Me \wedge \downarrow Ele)\right)$$

Let $k+1$ be the floor I am on right now.

So, $0 \leq k \leq n$

Now, two cases:

1) $k \leq \frac{n}{2}$



$$\boxed{1 - P\left(\text{none in shaded area}\right)}$$

$$=$$

$$P(\uparrow Ele) = P\left(\uparrow Ele \mid \begin{matrix}\text{at least}\\ 1\ Ele\\ \text{in shaded}\\ \text{area}\end{matrix}\right) \times P\left(\begin{matrix}\text{at least}\\ 1\ Ele\ \text{in}\\ \text{shaded}\\ \text{area}\end{matrix}\right)$$

$$+ P\left(\uparrow Ele \mid \begin{matrix}\text{none in}\\ \text{shaded area}\end{matrix}\right) \times P\left(\begin{matrix}\text{none in}\\ \text{shaded area}\end{matrix}\right)$$

$$= \frac{1}{2} \times \left(1 - \left(\frac{n-2k}{n}\right)^m\right) + 0$$

$$P(\downarrow Ele) = 1 - P(\uparrow Ele)$$

$$= 1 - \left(\frac{1}{2} \times \left(1 - \left(\frac{n-2k}{n}\right)^m\right)\right)$$

1) $k > \frac{n}{2}$

$$P(\downarrow Ele) = P\left(\downarrow Ele \,\middle|\, \begin{array}{c}\text{at least}\\ \text{1 Ele in}\\ \text{shaded}\\ \text{area}\end{array}\right) \times P\left(\begin{array}{c}\text{at least}\\ \text{1 Ele in}\\ \text{shaded}\\ \text{area}\end{array}\right)$$

$$+ P\left(\downarrow Ele \,\middle|\, \begin{array}{c}\text{none in}\\ \text{shaded area}\end{array}\right) \times P\left(\begin{array}{c}\text{none in}\\ \text{shaded}\\ \text{area}\end{array}\right)$$

$$= \frac{1}{2} \times \left(1 - \left(\frac{2n-2k}{n}\right)^m\right) + 0$$

$$P(\uparrow Ele) = 1 - P(\downarrow Ele)$$

$$= 1 - \left(\frac{1}{2} \times \left(1 - \left(\frac{2n-2k}{n}\right)^m\right)\right)$$

$$P\left((\uparrow Me \wedge \uparrow Ele) \vee (\downarrow Me \wedge \downarrow Ele)\right)$$

$$= \sum_{k=0}^{n/2} \left[\left(\frac{n-k}{n}\right) \times \left(\frac{1}{2} \times \left(1 - \left(\frac{n-2k}{n}\right)^m\right)\right) \times \frac{1}{n+1}\right]$$

$$+ \sum_{k=n/2+1}^{n} \left[\left(\frac{n-k}{n}\right) \times \left(1 - \left(\frac{1}{2} \times \left(1 - \left(\frac{2n-2k}{n}\right)^m\right)\right)\right) \times \frac{1}{n+1}\right]$$

$$+ \sum_{k=0}^{n/2} \left[\frac{k}{n} \times \left(1 - \left(\frac{1}{2} \times \left(1 - \left(\frac{n-2k}{n}\right)^m\right)\right)\right) \times \frac{1}{n+1}\right]$$

$$+ \sum_{k=\frac{n}{2}+1}^{n} \left[\frac{k}{n} \times \left(\frac{1}{2} \times \left(1 - \left(\frac{2n-2k}{n}\right)^m\right)\right) \times \frac{1}{n+1}\right]$$

By Symmetry:

$$P\big((\uparrow Me \wedge \uparrow Ele) \vee (\downarrow Me \wedge \downarrow Ele)\big)$$

$$= 2\sum_{k=0}^{n/2}\left[\left(\frac{n-k}{n}\right)\times\left(\frac{1}{2}\times\left(1-\left(\frac{n-2k}{n}\right)^m\right)\right)\times\frac{1}{n+1}\right]$$

$$+\,2\sum_{k=0}^{n/2}\left[\frac{k}{n}\times\left(1-\left(\frac{1}{2}\times\left(1-\left(\frac{n-2k}{n}\right)^m\right)\right)\right)\times\frac{1}{n+1}\right]$$

$$= 2\sum_{k=0}^{n/2}\left[\left(\frac{n-k}{n}\right)\times\frac{1}{2}\times\left(1-\left(\frac{n-2k}{n}\right)^m\right)\times\frac{1}{n+1} + \frac{k}{n}\times\left(1-\left(\frac{1}{2}\times\left(1-\left(\frac{n-2k}{n}\right)^m\right)\right)\right)\times\frac{1}{n+1}\right]$$

$$= 2\sum_{k=0}^{n/2}\left[\frac{(n-k)\left(1-\left(\frac{n-2k}{n}\right)^m\right)}{2n(n+1)} + \frac{k\left(1-\left(\frac{1}{2}\left(1-\left(\frac{n-2k}{n}\right)^m\right)\right)\right)}{n(n+1)}\right]$$

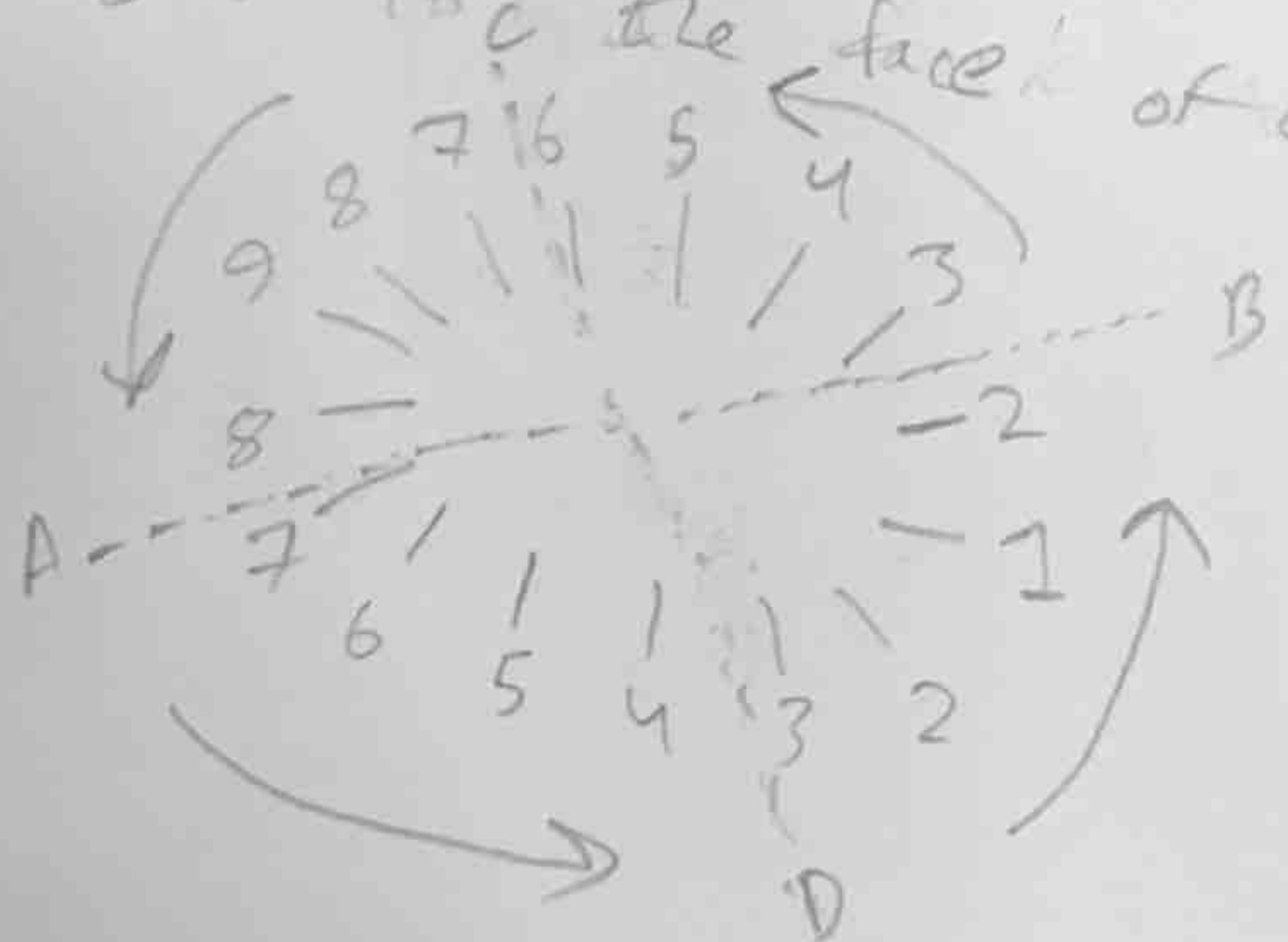
$$= 2\sum_{k=0}^{n/2}\left[\frac{(n-k)\left(1-\left(\frac{n-2k}{n}\right)^m\right) + k\left(2-1+\left(\frac{n-2k}{n}\right)^m\right)}{2n(n+1)}\right]$$

$$= 2\sum_{k=0}^{n/2}\left[\frac{(n-k)\left(1-\left(\frac{n-2k}{n}\right)^m\right) + k\left(1+\left(\frac{n-2k}{n}\right)^m\right)}{2n(n+1)}\right]$$

1.b)

Now, the two elevators can be on any of the floor and traveling strictly up or down. However, for the next elevator to come to me must be traveling downwards in order for me to use it. For example, if an elevator is currently on the 4th floor and traveling downwards, it needs to reach 1st floor, then go all the way up to 9th floor and then come down to me on the 7th floor.

We can simplify this problem into something analogous to the face of a clock as follows:



Assuming we travel only andiclockwise on the clock face, we can easily see the above example fitting nicely on this clock face.

Now, without loss of generality if we had only one elevator, and this elevator is equally likely to be on any floor then by law of total expectation, it is expected to take 8 minutes to reach me since there are 16 floors and each floor takes 1 minute. On the clock face, this means the elevator is expected to be on 3rd floor and travelling up. Now, the second elevator can be anywhere between me and the first elevator, and the expeted time is thus 4 minutes. In terms of expetation of joint probability

Let X be first elevator
" Y " second elevator

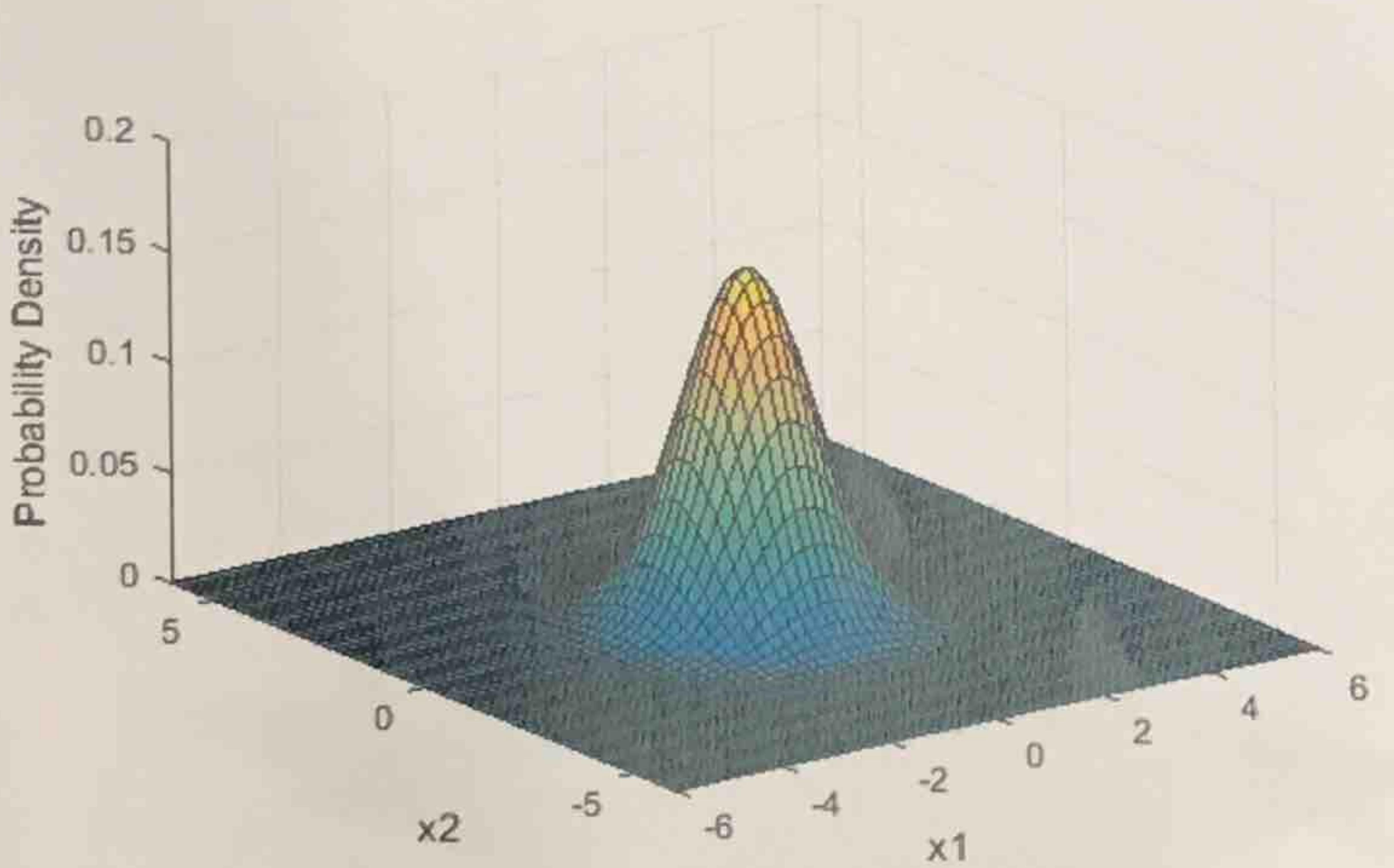
Now, $E[XY] = E[X]\,E[Y]$

$= \frac{1}{2} \times \frac{16}{2}$

$= 4$ min [ans]

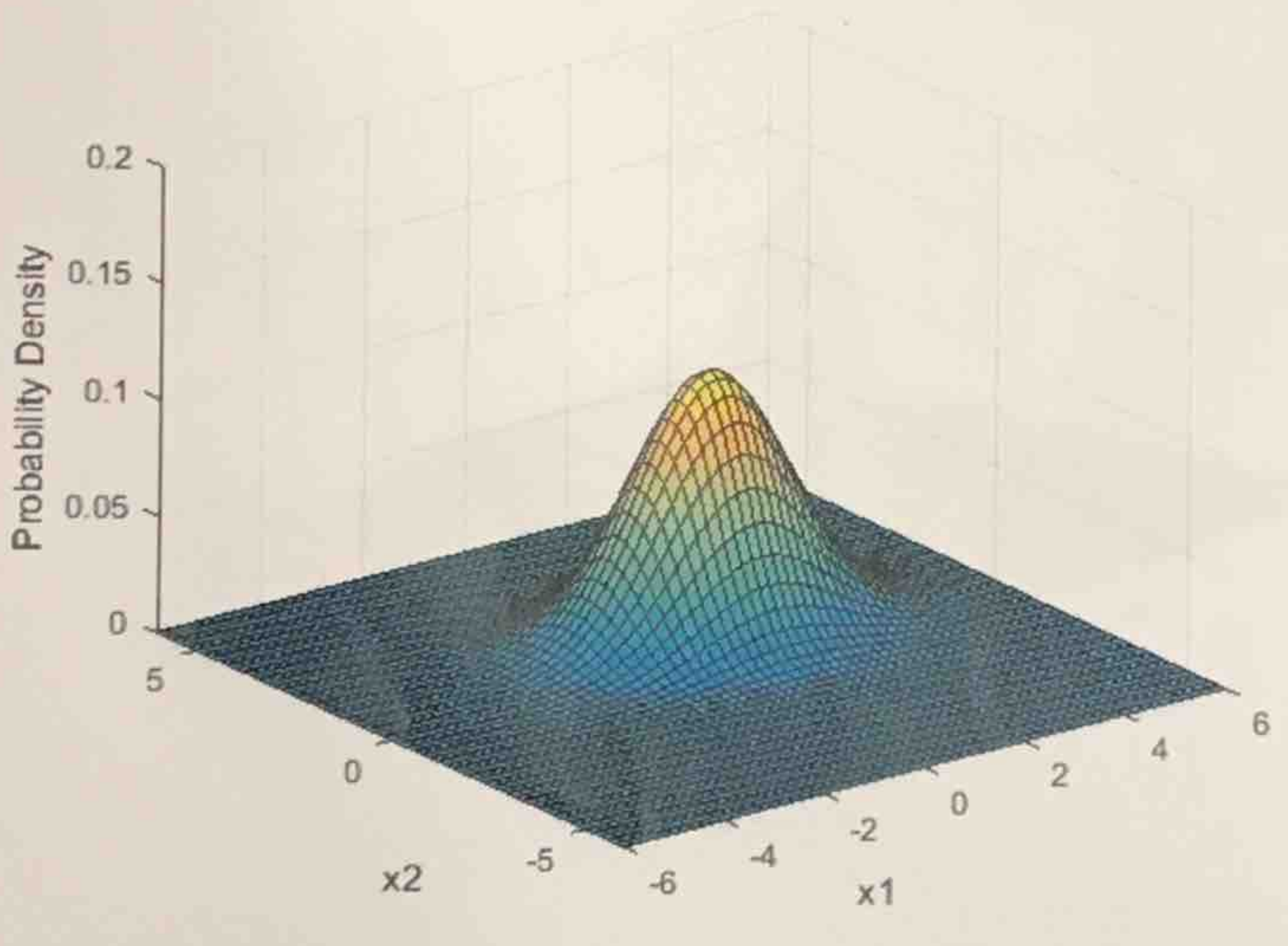
2.a.I.i

```
mu = [0 0];
Sigma = [1 0; 0 1];
x1 = -6:.2:6; x2 = -6:.2:6;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));
surf(x1,x2,F);
caxis([min(F(:))-.5*range(F(:)),max(F(:))]);
axis([-6 6 -6 6 0 .2])
xlabel('x1'); ylabel('x2'); zlabel('Probability Density');
```



2.a.I.ii

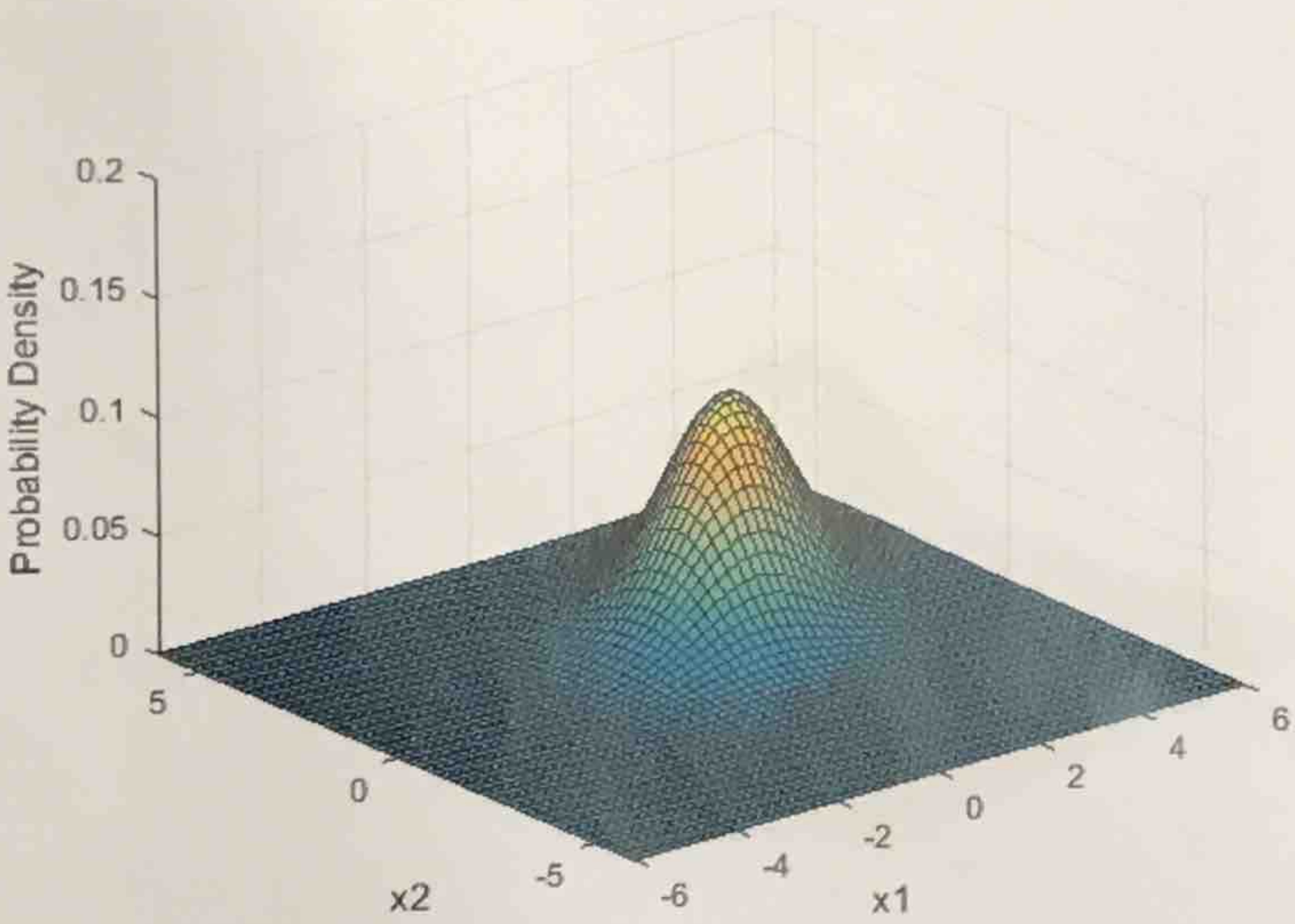
```
mu = [1 1];
Sigma = [2 0; 0 1];
x1 = -6:.2:6; x2 = -6:.2:6;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));
surf(x1,x2,F);
caxis([min(F(:))-.5*range(F(:)),max(F(:))]);
axis([-6 6 -6 6 0 .2])
xlabel('x1'); ylabel('x2'); zlabel('Probability Density');
```



2.a.I.ii

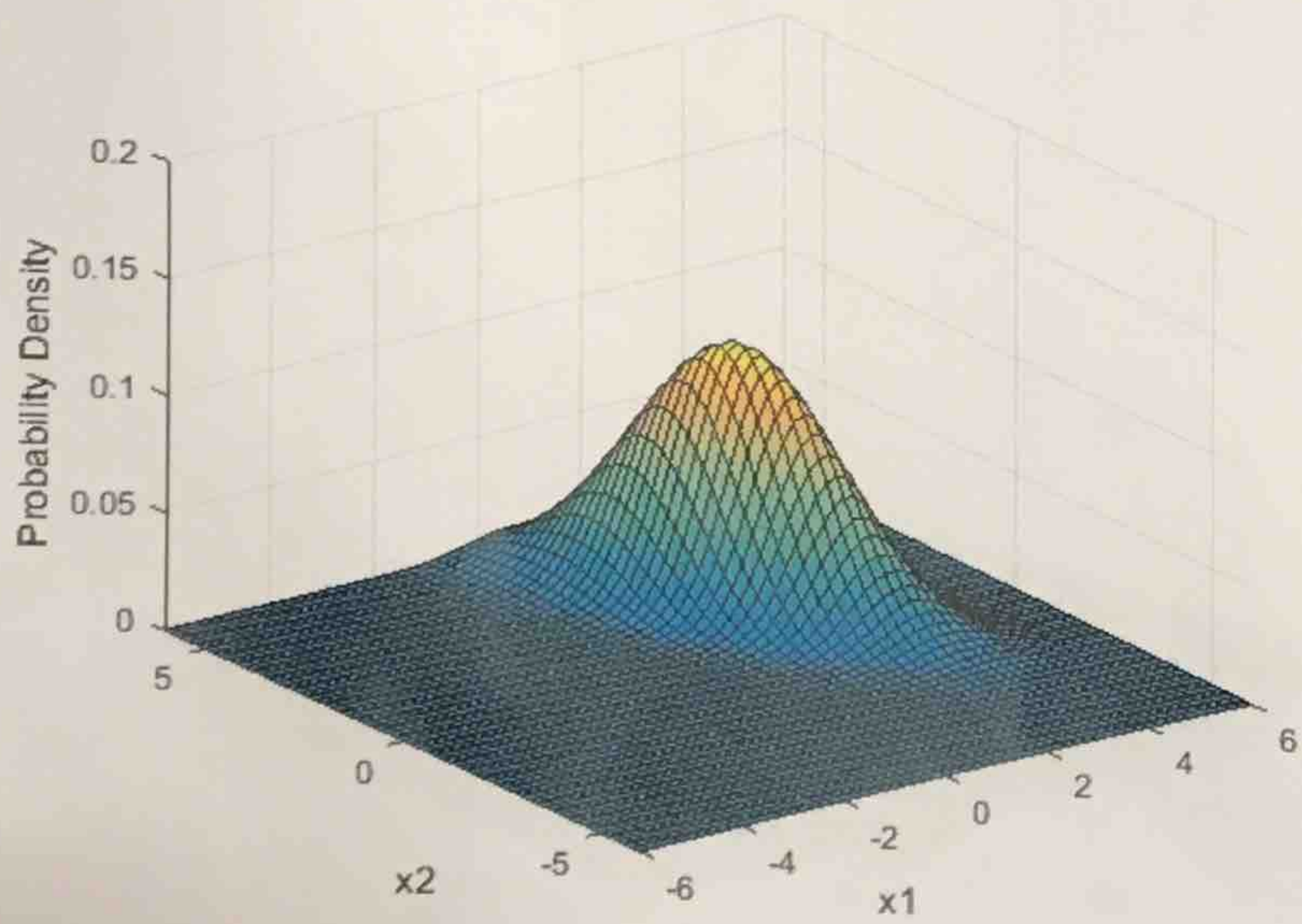
2.a.I.iii

```
mu = [2 2];
Sigma = [2 0.95; 0.95 2];
x1 = -6:.2:6; x2 = -6:.2:6;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));
surf(x1,x2,F);
caxis([min(F(:))-.5*range(F(:)),max(F(:))]);
axis([-6 6 -6 6 0 .2])
xlabel('x1'); ylabel('x2'); zlabel('Probability Density');
```



2.a.I.iv

```
mu = [2 2];
Sigma = [1 -1; -1 3];
x1 = -6:.2:6; x2 = -6:.2:6;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));
surf(x1,x2,F);
caxis([min(F(:))-.5*range(F(:)),max(F(:))]);
axis([-6 6 -6 6 0 .2])
xlabel('x1'); ylabel('x2'); zlabel('Probability Density');
```



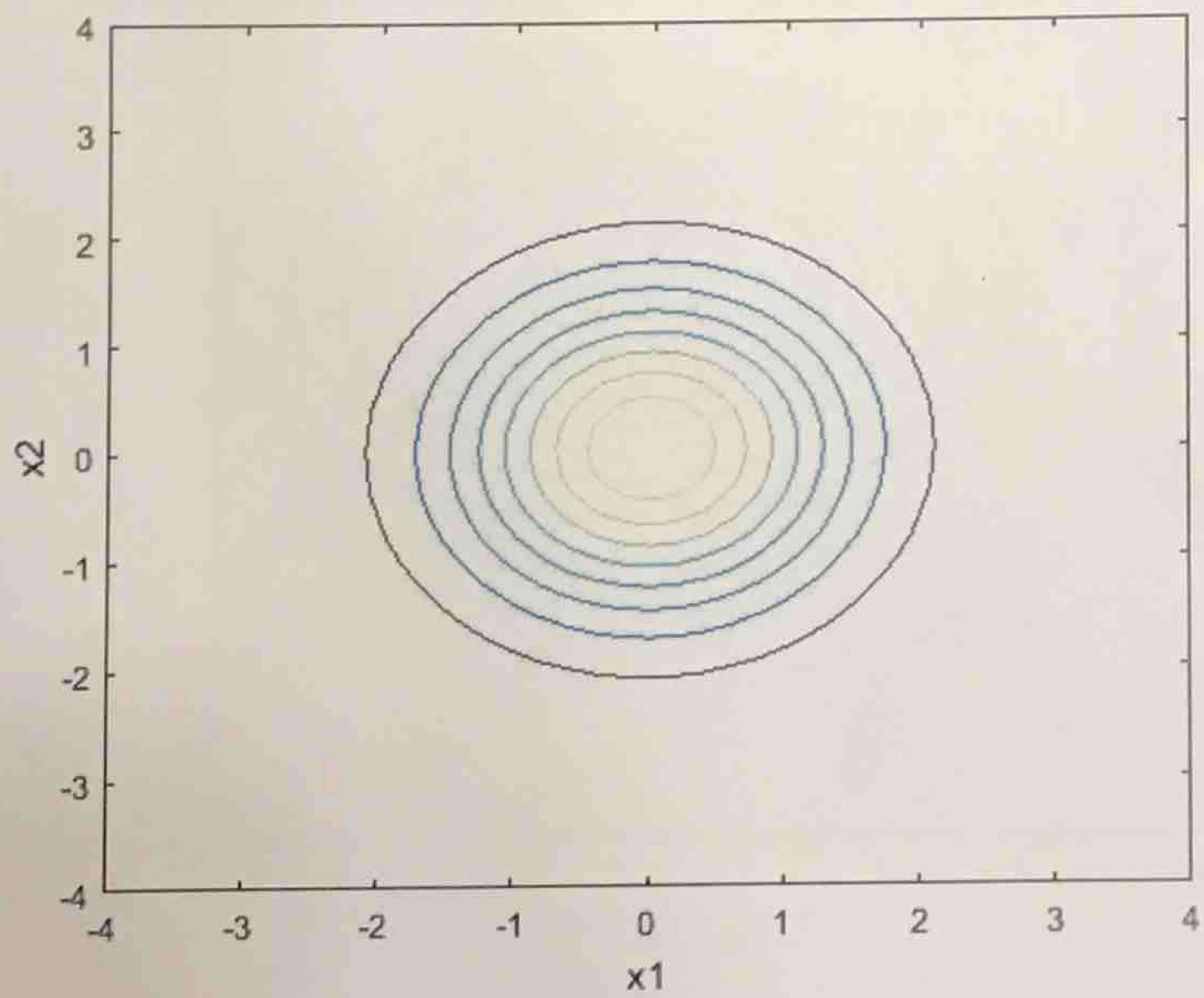
2.a.II.i

```
mu = [0 0];
Sigma = [1 0; 0 1];

max_c = 1/(sqrt(((2*pi)^2) * det(Sigma)))
c_range = linspace(0, max_c, 10)

x1 = -4:.2:4; x2 = -4:.2:4;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));

mvncdf([0 0],[1 1],mu,Sigma);
contour(x1,x2,F,c_range);
xlabel('x1'); ylabel('x2');
```



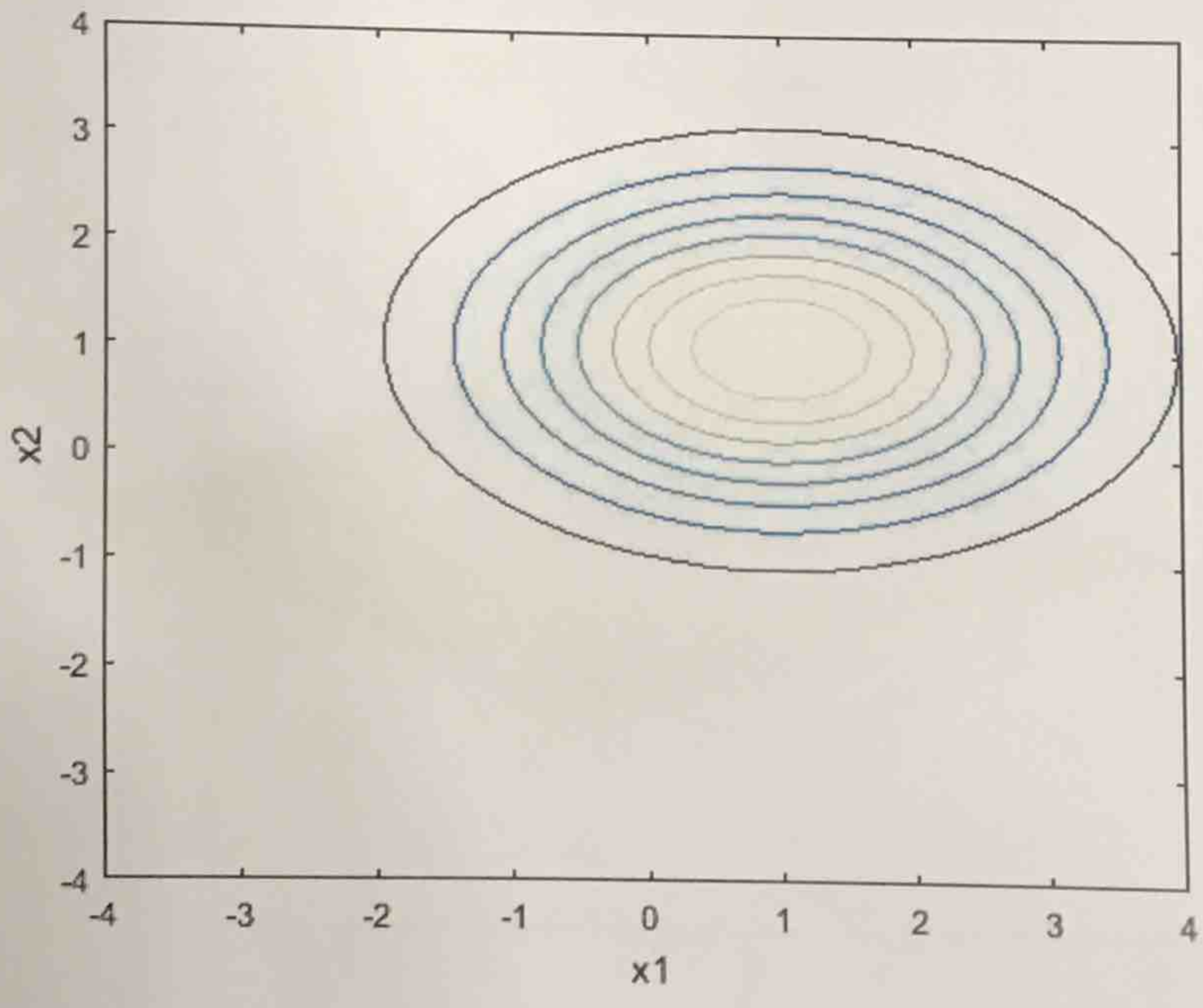
2.a.II.ii

```
mu = [1 1];
Sigma = [2 0; 0 1];

max_c = 1/(sqrt(((2*pi)^2) * det(Sigma)))
c_range = linspace(0, max_c, 10)

x1 = -4:.2:4; x2 = -4:.2:4;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));

mvncdf([0 0],[1 1],mu,Sigma);
contour(x1,x2,F, c_range);
xlabel('x1'); ylabel('x2');
```



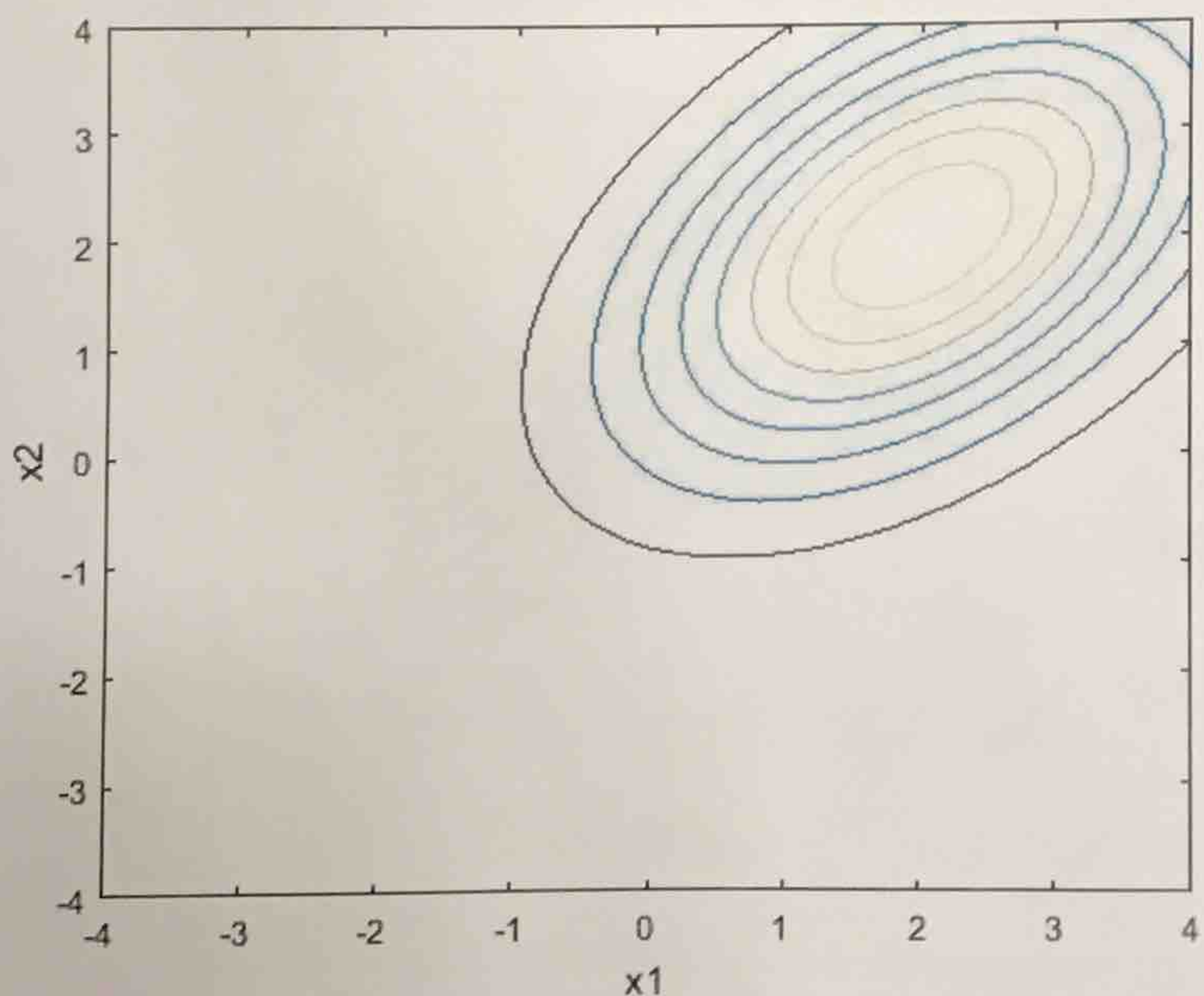
2.a.II.iii

```
mu = [2 2];
Sigma = [2 0.95; 0.95 2];

max_c = 1/(sqrt(((2*pi)^2) * det(Sigma)))
c_range = linspace(0, max_c, 10)

x1 = -4:.2:4; x2 = -4:.2:4;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));

mvncdf([0 0],[1 1],mu,Sigma);
contour(x1,x2,F,c_range);
xlabel('x1'); ylabel('x2');
```



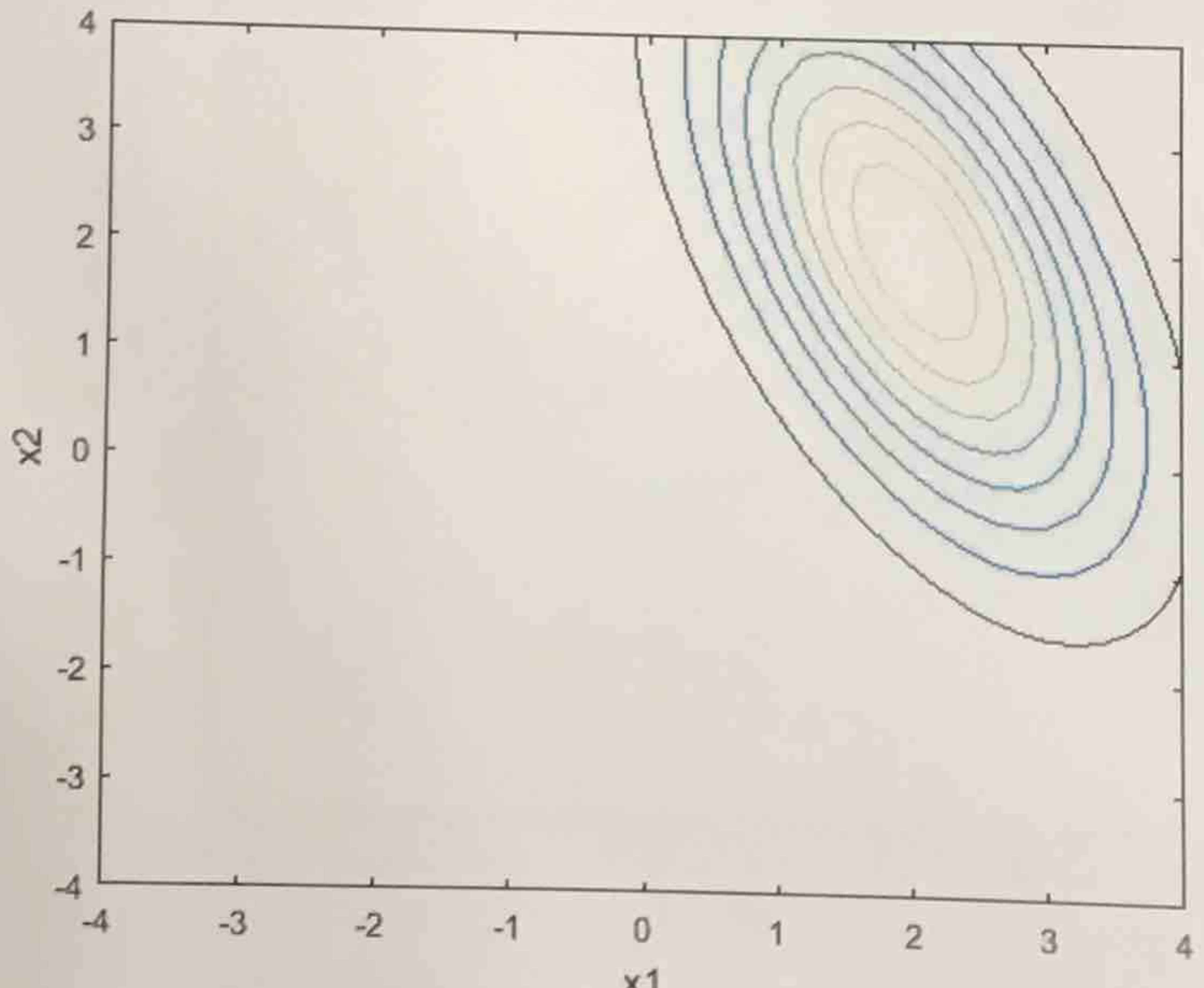
2.a.II.iv

```
mu = [2 2];
Sigma = [1 -1; -1 3];

max_c = 1/(sqrt(((2*pi)^2) * det(Sigma)))
c_range = linspace(0, max_c, 10)

x1 = -4:.2:4; x2 = -4:.2:4;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));

mvncdf([0 0],[1 1],mu,Sigma);
contour(x1,x2,F,c_range);
xlabel('x1'); ylabel('x2');
```



b) No, this is not a valid Gaussian.
This is because the covariance matrix Φ is wrong.
This is because by definition, each matrix cell of Φ, $\sigma_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)]$.
This means that σ_{ij} must be equivalent to σ_{ji} as product of $x \times y$ is same as $y \times x$. However in the given Φ, $\sigma_{ij} \neq \sigma_{ji}$ when $i \neq j$. Also, when $i = j$, $(x_i - \mu_i)(x_j - \mu_j) = (x_i - \mu_i)^2$, and thus σ_{ij} when $i = j$ must be a positive number. Howeve σ_{22} in the given Φ is -1, which is wrong. Lastly, the determinant of the given Φ is negative and we cannot square root a negative number while calculating the multivariate Gaussian probability density function.

5. We can represent the multivariate Gaussian PDF as $P_x(X; \mu, \Phi)$

We will do these in parts:

1) simply outputs X

So, $0.5 \times P_x\left(X; [0\ 0], \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$

2) multiplies x_1 by 2 and outputs $Y = [2x_1\ x_2]^T$

$$Y = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Now, we need to find the mean and covariance of this

$$E[Y] = E\left(\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \cdot E[X] = 0$$

$$\Phi(Y) = E[(Y - \mu_y)(Y - \mu_y)^T]$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} [x_1\ x_2]}_{I} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore,

$$0.2 \times P_X\left(X; 0, \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

iii)

$$Y = \begin{bmatrix} x_1 \\ x_2 + Z \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ Z \end{bmatrix}$$

$$E[Y] = E\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + E\begin{bmatrix} 0 \\ Z \end{bmatrix}$$

$$= E[X] + \begin{bmatrix} 0 \\ E[Z] \end{bmatrix}$$

$$= 0 + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Phi(Y) = E\left[(Y - \mu_y)(Y - \mu_y)^T\right]$$

$$= E\left[\left(X + Z - (\mu_x + \mu_z)\right)\left(X + Z - (\mu_x + \mu_z)\right)^T\right]$$

$$= E\left[\left((x - \mu_x) + (Z - \mu_z)\right)\left((x - \mu_x) + (Z - \mu_z)\right)^T\right]$$

$$= E\begin{bmatrix} (x - \mu_x)(x - \mu_x)^T + (Z - \mu_z)(Z - \mu_z)^T \\ + (x - \mu_x)(x - \mu_z)^T + (Z - \mu_z)(x - \mu_x)^T \end{bmatrix}$$

$$= E\left((x-\mu_x)(x-\mu_x)^T\right) + E\left((z-\mu_z)(z-\mu_z)^T\right)$$

$$+ E\left((x-\mu_x)(z-\mu_z)^T\right) + E\left((z-\mu_z)(x-\mu_x)^T\right)$$

$$= I + \Phi(z)$$

$$= I + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

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

So, we have:

$$0.3 \times P_x\left(X; \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right)$$

Therefore,

$$P_Y = 0.5 \times P_x\left(X; 0, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

$$+ 0.2 \times P_x\left(X; 0, \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

$$+ 0.3 \times P_x\left(X; \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right)$$

[Ans]