Homework #1

Theodor Chakhachiro
UMID: 15801216
ROB501 - Mathematics for Robotics
University of Michigan, Ann Arbor

September 5, 2021

Problem 1

Part a

To prove that $AB = [Ab^1|Ab^2|\ldots|Ab^p]$, we use the direct proof method by solving both the left and hand-side and right hand-side. We then complete the proof by showing an analogy between the two results.
Starting with the left hand-side, we have:

$$A_{n,m} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}, B_{m,p} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,p} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,p} \end{bmatrix}$$

Now introducing a new matrix $C_{n,p}$ as $C_{n,p} = A_{n,m}B_{m,p}$, we get:

$$C_{n,p} = A_{n,m}B_{m,p} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,p} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,p} \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,p} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,p} \end{bmatrix}$$

Where we obtain the following $C_{n,p}$ matrix parameters:

$$C_{n,p} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,p} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,p} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m a_{1,i}b_{i,1} & \sum_{i=1}^m a_{1,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{1,i}b_{i,p} \\ \sum_{i=1}^m a_{2,i}b_{i,1} & \sum_{i=1}^m a_{2,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{2,i}b_{i,p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{n,i}b_{i,1} & \sum_{i=1}^m a_{n,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{n,i}b_{i,p} \end{bmatrix} \tag{1}$$

Now solving for the right hand-side and by deconstructing each term $D_{n,1}^{j=1,\ldots,p} = Ab^j$, where b^j for denote the j-th column vector of matrix $B_{m,p}$, we obtain:

$$D_{n,1}^j = Ab^j = A * \begin{bmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{m,j} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix} \begin{bmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{m,j} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m a_{1,i}b_{i,j} \\ \sum_{i=1}^m a_{2,i}b_{i,j} \\ \vdots \\ \sum_{i=1}^m a_{n,i}b_{i,j} \end{bmatrix}$$

Thus for $j = 1$, we get:

$$D^1_{n,1} = Ab^1 \begin{bmatrix} \sum_{i=1}^m a_{1,i}b_{i,1} \\ \sum_{i=1}^m a_{2,i}b_{i,1} \\ \vdots \\ \sum_{i=1}^m a_{n,i}b_{i,1} \end{bmatrix}$$

Now concatenating all the elements $D^j_{n,1}$ of the matrix $D_{n,p} = [Ab^1|Ab^2|\dots|Ab^p]$ column-wise, we obtain:

$$D_{n,p} = \left[\begin{array}{r|r|c|r} \sum_{i=1}^m a_{1,i}b_{i,1} & \sum_{i=1}^m a_{1,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{1,i}b_{i,p} \\ \sum_{i=1}^m a_{2,i}b_{i,1} & \sum_{i=1}^m a_{2,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{2,i}b_{i,p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{n,i}b_{i,1} & \sum_{i=1}^m a_{n,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{n,i}b_{i,p} \end{array}\right] \tag{2}$$

Comparing equations (1) and (2), we notice that the two matrices $C_{n,p} = A_{n,m}B_{m,p}$ and $D_{n,p} = [Ab^1|Ab^2|\dots|Ab^p]$ are similar. Thus, we can conclude using a direct proof that:

$$\boxed{AB = [Ab^1|Ab^2|\dots|Ab^p] \quad \square}$$

Part b

Again, to prove that $AB = [a^1B|a^2B|\dots|a^nB]^T$, where a^i denotes the i-th row of the matrix $A_{n,m}$, we resort to the direct proof. Since we already have solved the left hand-side in equation 1, we only need to work on the right hand-side and demonstrate that it is equal to 1. Let us denote $D^j_{1,p} = a^jB$, then the right hand-side can be expanded as follows:

$$D^j_{1,p} = a^jB = \begin{bmatrix} a_{j,1} \\ a_{j,2} \\ \vdots \\ a_{j,m} \end{bmatrix}^T \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,p} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,p} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m a_{j,i}b_{i,1} & \sum_{i=1}^m a_{j,i}b_{i,2} & \dots & \sum_{i=1}^m a_{j,i}b_{i,p} \end{bmatrix}$$

Now for $j = 1$, we have:

$$D^1_{1,p} = a^1B = \begin{bmatrix} \sum_{i=1}^m a_{1,i}b_{i,1} & \sum_{i=1}^m a_{1,i}b_{i,2} & \dots & \sum_{i=1}^m a_{1,i}b_{i,p} \end{bmatrix}$$

Now concatenating all the elements $D^j_{1,p}$ of the matrix $D_{n,p} = [a^1B|a^2B|\dots|a^nB]^T$ row-wise, we obtain:

$$D_{n,p} = \left[\begin{array}{rrcr} \sum_{i=1}^m a_{1,i}b_{i,1} & \sum_{i=1}^m a_{1,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{1,i}b_{i,p} \\ \hline \sum_{i=1}^m a_{2,i}b_{i,1} & \sum_{i=1}^m a_{2,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{2,i}b_{i,p} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \sum_{i=1}^m a_{n,i}b_{i,1} & \sum_{i=1}^m a_{n,i}b_{i,2} & \cdots & \sum_{i=1}^m a_{n,i}b_{i,p} \end{array}\right] \tag{3}$$

Comparing equations (1) and (3), we notice that the two matrices $C_{n,p} = A_{n,m}B_{m,p}$ and $D_{n,p} = [Ab^1|Ab^2|\dots|Ab^p]$ are similar. Thus, we can conclude using a direct proof that:

$$\boxed{AB = [a^1B|a^2B|\dots|a^nB]^T \quad \square}$$

Part c

To prove that $[AB]_{ij} = a^i b^j$, yet again we resort to the direct proof method. Equation 1 already gives us the solution of the left hand-side and thus we have:

$$[AB]_{ij} = ab_{ij} = c_{ij} = \sum_{k=1}^{m} a_{i,k} b_{k,j} \tag{4}$$

Therefore, the first elements of the matrix $C_{n,p}$ denoted by $c_{1,1}$ is given by (4):

$$c_{1,1} = ab_{1,1} = \sum_{k=1}^{m} a_{1,k} b_{k,1}$$

For the right hand-side, we start by expanding $a^i b^j$, we obtain:

$$a^i b^j = \begin{bmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,m} \end{bmatrix} \begin{bmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{m,j} \end{bmatrix} = a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \ldots + a_{i,m} b_{1,m} = \sum_{k=1}^{m} a_{i,k} b_{k,j} \tag{5}$$

Comparing equations (4) and (5), we notice that the elements $[AB]_{ij}$ and $a^i b^j$ are similar. Thus, we can conclude using a direct proof that:

$$\boxed{[AB]_{ij} = a^i b^j \quad \square}$$

Problem 2

Part a

To compute the trace of the 3×3 matrix A, we apply the definition of the trace as follows:

$$tr(A) = \sum_{i=1}^{n} a_{i,i} = \sum_{i=1}^{3} a_{i,i} = 1 + 5 + 9 = 15$$

Part b

$x = [x_1, x_2, \ldots, x_n]^T$ is a $n \times 1$ vector and its transpose is a $1 \times n$ vector, thus applying the properties of the trace, we have that:

$$tr(xx^T) = tr(x^T x) = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = x_1^2 + x_2^2 + \ldots + x_n^2 = \sum_{i=1}^{n} x_i^2$$

Part c

To compute $tr(K^TQK)$, we first have to define the Q and K matrices. We have:

$$Q_{n,n} = \begin{bmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,n} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n,1} & q_{n,2} & \cdots & q_{n,n} \end{bmatrix}$$

$$K_{n,m} = \begin{bmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,m} \\ k_{2,1} & k_{2,2} & \cdots & k_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n,1} & k_{n,2} & \cdots & k_{n,m} \end{bmatrix} = \begin{bmatrix} k^1 & k^2 & \dots & k^m \end{bmatrix}$$

where k^i denote the i-th column vector of the matrix K, defined as follows:

$$k^i = \begin{bmatrix} k_{1,i} \\ k_{2,i} \\ \vdots \\ k_{n,i} \end{bmatrix}$$

First, let us separate the matrix $A = K^TQK$ into $A_1 = K^T$ and $A_2 = QK$, we have:

$$A_2 = QK = Q\begin{bmatrix} k^1 & k^2 & \dots & k^m \end{bmatrix} = \begin{bmatrix} Qk^1 & Qk^2 & \dots & Qk^m \end{bmatrix} \tag{6}$$

Moreover:

$$K^T = \begin{bmatrix} k^1 & k^2 & \dots & k^m \end{bmatrix}^T = \begin{bmatrix} k^1 \\ k^2 \\ \vdots \\ k^m \end{bmatrix}$$

Thus:

$$A = A_1A_2 = K^T(QK) = \begin{bmatrix} k^1 \\ k^2 \\ \vdots \\ k^m \end{bmatrix} \begin{bmatrix} Qk^1 & Qk^2 & \dots & Qk^m \end{bmatrix} = \begin{bmatrix} k^1Qk^1 & k^1Qk^2 & \cdots & k^1Qk^m \\ k^2Qk^1 & k^2Qk^2 & \cdots & k^2Qk^m \\ \vdots & \vdots & \ddots & \vdots \\ k^mQk^1 & k^mQk^2 & \dots & k^mQk^m \end{bmatrix}$$

We then can notice that the elements of the diagonal of matrix A have a special pattern. This pattern can be used to compute the trace of K^TQT in terms of k^i and Q only:

$$tr(A) = tr(K^TQT) = tr\left(\begin{bmatrix} k^1Qk^1 & k^1Qk^2 & \cdots & k^1Qk^m \\ k^2Qk^1 & k^2Qk^2 & \cdots & k^2Qk^m \\ \vdots & \vdots & \ddots & \vdots \\ k^mQk^1 & k^mQk^2 & \dots & k^mQk^m \end{bmatrix}\right) = \sum_{i=1}^{m} k^iQk^i \tag{7}$$

Problem 3

Part a

To compute the eigenvalues of the matrix M and their corresponding eigenvector, we solve $det(M - \lambda I) = 0$:

$$det(M - \lambda I) = |M - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda) - 1 = \lambda^2 - 5\lambda + 5$$

Solving for $\lambda_{1,2}$:

$$\lambda^2 - 5\lambda + 5 = 0 \implies \lambda_{1,2} = \frac{5 \pm \sqrt{5}}{2}$$

To compute the eigenvectors, we solve $Mv_{1,2} = \lambda_{1,2} v_{1,2}$. Starting with $\lambda_1 = \frac{5+\sqrt{5}}{2}$, we obtain:

$$Mv_1 = \lambda_1 v_1 \implies \begin{bmatrix} 2-\lambda_1 & 1 \\ 1 & 3-\lambda_1 \end{bmatrix} v_1 = 0$$

$$\begin{bmatrix} \frac{-1-\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} v_{1,1} \\ v_{1,2} \end{bmatrix} = 0 \implies \begin{bmatrix} 1 & \frac{1-\sqrt{5}}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_{1,1} \\ v_{1,2} \end{bmatrix} = 0$$

Setting $v_{1,2} = t$ and then solving for $t = 1$, we get:

$$v_1 = \begin{bmatrix} \frac{\sqrt{5}-1}{2} \\ 1 \end{bmatrix}$$

Now solving for $\lambda_2 = \frac{5-\sqrt{5}}{2}$, we obtain:

$$Mv_2 = \lambda_2 v_2 \implies \begin{bmatrix} 2-\lambda_2 & 1 \\ 1 & 3-\lambda_2 \end{bmatrix} v_2 = 0$$

$$\begin{bmatrix} \frac{-1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} v_{2,1} \\ v_{2,2} \end{bmatrix} = 0 \implies \begin{bmatrix} 1 & \frac{1+\sqrt{5}}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_{2,1} \\ v_{2,2} \end{bmatrix} = 0$$

Setting $v_{2,2} = t$ and then solving for $t = 1$, we get:

$$v_2 = \begin{bmatrix} \frac{-\sqrt{5}-1}{2} \\ 1 \end{bmatrix}$$

Part b

$$(v_1)^T v_2 = \begin{bmatrix} \frac{\sqrt{5}-1}{2} & 1 \end{bmatrix} \begin{bmatrix} \frac{-\sqrt{5}-1}{2} \\ 1 \end{bmatrix} = 1 - \frac{1 - 5 + \sqrt{5} - \sqrt{5}}{4} = 0$$

Part c

To prove that $M = A^T A$ is a symmetric matrix for any real $n \times m$ matrix A, it is enough to show that $M^T = M$. In fact:

$$M^T = (A^T A)^T = (A^T A) = M$$

Thus M is an $m \times m$ symmetric matrix for any real $n \times m$ matrix A.

Part d

The matlab code used to solve this part is displayed below:

```matlab
%% Theodor Chakhachiro ROB 501 HW1 Problem 3
clear all
clc
%% Initialization
n=4;
m=3;
A=zeros(n,m,10);
%% Computing Eigenvalues and Eigenvectors
for i=1:length(A)
    A(:,:,i)=randn(n,m);
    M(:,:,i)=A(:,:,i)'*A(:,:,i);
    [V(:,:,i),L(:,:,i)]=eig(M(:,:,i));
    %% Checking that the sum of the eigenvalues of a matrix is equal to the trace of that matrix
    L_sum(i)=trace(L(:,:,i));
    L_trace(i)=trace(M(:,:,i));
    Trace_check(i)=L_sum(i)-L_trace(i);
    %% Checking that the product of the eigenvalues is equal to the determinant of that matrix
    L_product(i)=prod(diag(L(:,:,i)));
    L_det(i)=det(M(:,:,i));
    Det_check(i)=L_det(i)-L_product(i);
end
%% Checking that the inner product of the eigenvectors of a matrix is equal to zero
for i=1:length(A)
    for j=1:m-1
        VecVal(i,j)=V(1:m,j,i)'*V(1:m,j+1,i);
    end
    VecVal(i,1:m)=[VecVal(i,1:m-1) V(1:m,1,i)'*V(1:m,m,i)];
end
%% End
```

From the previous running code, we can notice that the inner product of any two eigenvectors $(v^i)^T v^j$ of a $n \times n$ square matrix M results in 0, and thus:

$$\forall i, j = 1, \ldots, n \quad \text{s.t.} \quad i \neq j : (v^i)^T v^j = 0$$

We also notice that the sum of all the eigenvalues $\lambda_{1,\ldots,n}$ of a $n \times n$ square matrix M is equal to the trace of that matrix. Therefore, we can write:

$$\sum_{i=1}^{n} \lambda_i = tr(M)$$

Furthermore, we note that the multiplication of all the eigenvalues $\lambda_{1,\ldots,n}$ of a $n \times n$ square matrix M is equal to the determinant of that matrix. Therefore, we can write:

$$\prod_{i=1}^{n} \lambda_i = det(M)$$

Problem 4

Part a

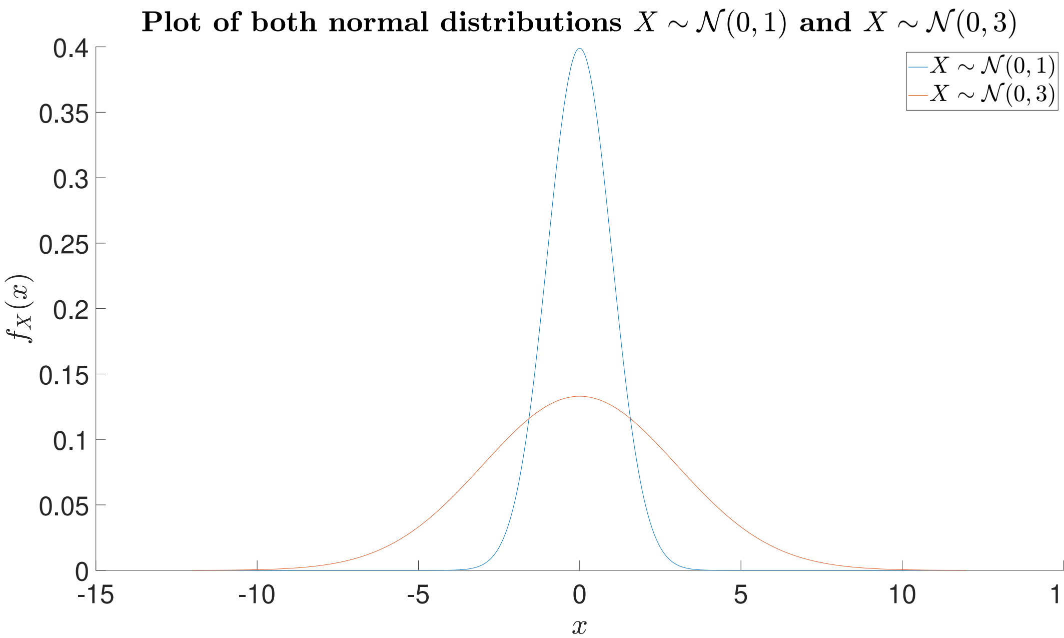

Figure 1: Plot of both normal distributions $X \sim \mathcal{N}(0,1)$ in blue and $X \sim \mathcal{N}(0,3)$ in red

Part b

i)

By either using Matlab quad/integral function or the Z-scores from the Z-tables, we can find the solution to all these probabilities since $X \sim \mathcal{N}(2,5)$:

$$P(X \geqslant 4) = 1 - P(X < 4) = 1 - P(Z < 0.4) = 1 - 0.6554 = 0.3446$$

Or plugging in the following integral in the quad/integral function in Matlab:

$$P(X \geqslant 4) = \int_4^\infty f_X(x)dx = \int_4^\infty \frac{1}{5\sqrt{2\pi}} e^{-\frac{(x-2)^2}{50}} dx = 0.3446$$

ii)

$$P(-2 \leqslant X \leqslant 4) = P(X \leqslant 4) - P(X \leqslant -2) = P(Z \leqslant 0.4) - P(Z \leqslant -0.8) = 0.6555 - 0.2119 = 0.4436$$

Or plugging in the following integral in the quad/integral function in Matlab:

$$P(X \geqslant 4) = \int_{-2}^4 f_X(x)dx = \int_{-2}^4 \frac{1}{5\sqrt{2\pi}} e^{-\frac{(x-2)^2}{50}} dx = 0.4436$$

iii)

$$
\begin{aligned}
P(X \in A) &= P(X \in [-2,4]) + P(X \in [8,100]) \\
&= P(-2 \leqslant X \leqslant 4) + P(8 \leqslant X \leqslant 100) \\
&= 0.4436 + P(X \leqslant 100) - P(X \leqslant 8) \\
&= 0.4436 + P(Z \leqslant 19.6) - P(Z \leqslant 1.2) \\
&= 0.4436 + 0.999 - 0.8849 \\
&= 0.5586
\end{aligned}
$$

Or by plugging in the following integral in the quad/integral function in Matlab:

$$
\begin{aligned}
P(X \in A) &= \int_{-2}^{4} f_X(x)dx + \int_{8}^{100} f_X(x)dx \\
&= \int_{-2}^{4} \frac{1}{5\sqrt{2\pi}} e^{-\frac{(x-2)^2}{50}} dx + \int_{8}^{100} \frac{1}{5\sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{50}} dx \\
&= 0.5586
\end{aligned}
$$

```matlab
%% Theodor Chakhachiro ROB 501 HW1 Problem 4
clear all
clc
%% Initializing
bigstd_x = 3;
n = 4*bigstd_x;
x = [-n:.01:n];
y1 = normpdf(x,0,1);
y2 = normpdf(x,0,3);
figure(1);
hold on
plot(x,y1)
plot(x,y2)
legend('$X \sim \mathcal{N}(0,1)$','$X \sim \mathcal{N}(0,3)$','Interpreter','latex')
title('\textbf{Plot of both normal distributions $X \sim \mathcal{N}(0,1)$ and $X \sim \
    mathcal{N}(0,3)$}','Interpreter','latex')
xlabel('$x$','Interpreter','latex')
ylabel('$f_X(x)$','Interpreter','latex')
set(gca,'fontsize',40)
%% Calculating Probabilities
% i) P(x>=4)
Q1 = integral(@(x)mynormaldist(2,5,x),4,inf);
% ii) P(-2<=x<=4)
Q2 = integral(@(x)mynormaldist(2,5,x),-2,4);
% iii
Q3 = integral(@(x)mynormaldist(2,5,x),-2,4) + integral(@(x)mynormaldist(2,5,x),8,100);
%% End
```

```matlab
function y = mynormaldist(mu,sigma,x)
y=exp(-((x-mu).^2)/50)/(sigma*sqrt(2*pi));
end
```

Part c

We know that $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ with $\mu = 2$ and $\sigma_X^2 = 5$, so:

$$
P(X < a) = \int_{-\infty}^{a} f_X(x)dx = \int_{-\infty}^{a} \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma_X^2}} dx
$$

Thus for $Y = 2X + 4$, we have:

$$P(Y < a) = P(X < \frac{a-4}{2}) = \int_{-\infty}^{\frac{a-4}{2}} f_X(x)dx = \int_{-\infty}^{\frac{a-4}{2}} \frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}dx$$

It is important that we modify our integral according to the change of variable adopted, therefore we notice that $dx = \frac{dy}{2}$. As $x \to \frac{a-4}{2}$, we have $y \to a$. Therefore we obtain:

$$\begin{aligned} P(Y < a) &= \int_{-\infty}^{a} \frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{(\frac{y-4}{2}-\mu_X)^2}{2\sigma_X^2}}\frac{dy}{2} \\ &= \int_{-\infty}^{a} \frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{(\frac{y-4}{2}-\mu_X)^2}{2\sigma_X^2}}\frac{dy}{2} \\ &= \int_{-\infty}^{a} \frac{1}{2\sigma_X\sqrt{2\pi}}e^{-\frac{(y-(4+2\mu_X))^2}{2*2^2\sigma_X^2}}dy \end{aligned}$$

We can now define from the above the normal distribution parameters:

$$\mu_Y = 4 + 2\mu_X$$
$$\sigma_Y^2 = 2^2\sigma_X^2$$

Finally, we can define the normal distribution of Y as $Y \sim \mathcal{N}(2\mu_X + 4, 2^2\sigma_X^2)$

Problem 5

Part a

To find the value of the constant K, we integrate the joint density $f_{X,Y}(x, y)$ over the whole domain of x, y and then equate it to 1. Thus we have:

$$\begin{aligned} 1 &= \int_0^2 \int_0^1 f_{X,Y}(x,y)dxdy \\ &= \int_0^2 \int_0^1 K(x+y)^2dxdy \\ &= K\int_0^2 \int_0^1 (x^2 + 2xy + y^2)dxdy \\ &= K\int_0^2 [\frac{1}{3}x^3 + yx^2 + y^2x]_0^1 dy \\ &= K\int_0^2 (\frac{1}{3} + y + y^2)dy \\ &= K[\frac{1}{3}y + \frac{1}{2}y^2 + \frac{1}{3}y^3]_0^2 \\ &= K(\frac{2}{3} + 2 + \frac{8}{3}) = \frac{16}{3}K \\ &\boxed{\implies K = \frac{3}{16}} \end{aligned}$$

Part b

To get the marginal distributions of X and Y, we integrate the joint density $f_{X,Y}(x,y)$ over the domain of the respective random variable. That is, we have:

$$\begin{aligned}
f_X(x) &= \int_0^2 f_{X,Y}(x,y)dy \\
&= \int_0^2 K(x+y)^2 dy \\
&= \int_0^2 K(x^2+2xy+y^2)dy \\
&= K[x^2y + xy^2 + \frac{1}{3}y^3]_0^2 \\
&= K(2x^2 + 4x + \frac{8}{3})
\end{aligned}$$

$$\boxed{f_X(x) = \frac{3}{8}x^2 + \frac{3}{4}x + \frac{1}{2}}$$

We do the same for $f_Y(y)$ and we obtain:

$$\begin{aligned}
f_Y(y) &= \int_0^1 f_{X,Y}(x,y)dx \\
&= \int_0^1 K(x+y)^2 dx \\
&= \int_0^1 K(x^2+2xy+y^2)dx \\
&= K[\frac{1}{3}x^3 + x^2y + y^2x]_0^1 \\
&= K(\frac{1}{3} + y + y^2)
\end{aligned}$$

$$\boxed{f_Y(y) = \frac{3}{16}y^2 + \frac{3}{16}y + \frac{1}{16}}$$

Part c

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \frac{K(x+y)^2}{K(y^2+y+\frac{1}{3})}
\end{aligned}$$

$$\boxed{f_{X|Y}(x|y) = \frac{(x+y)^2}{(y^2+y+\frac{1}{3})}}$$

Problem 6

We want to solve the following optimization problem using the Lagrange multipliers method for $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$:

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & x_1 + 3x_2 = 4 \end{aligned}$$

The Lagrangian function $\mathcal{L}(x_1, x_2, \lambda)$ is defined as follows:

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) - g(x_1, x_2)$$

where $f(x_1, x_2)$ is our optimization function and $g(x_1, x_2)$ is the constraint function, both defined as follows:

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2^2 \\ g(x_1, x_2) &= x_1 + 3x_2 - 4 \end{aligned}$$

Now let us solve this minimization problem by first getting the partial derivatives of $\mathcal{L}(x_1, x_2, \lambda)$, we have:

$$\begin{aligned} \mathcal{L}_{x_1}(x_1, x_2, \lambda) &= 2x_1 - \lambda \\ \mathcal{L}_{x_2}(x_1, x_2, \lambda) &= 2x_2 - 3\lambda \\ \mathcal{L}_{\lambda}(x_1, x_2, \lambda) &= 4 - 3x_2 - x_1 \end{aligned}$$

Now equating each of these partial derivatives to 0, we get:

$$\begin{aligned} \mathcal{L}_{x_1}(x_1, x_2, \lambda) = 0 &\implies x_1 = \lambda/2 \\ \mathcal{L}_{x_2}(x_1, x_2, \lambda) = 0 &\implies x_2 = 3\lambda/2 \\ \mathcal{L}_{\lambda}(x_1, x_2, \lambda) = 0 &\implies g(x_1, x_2) = 0 \end{aligned}$$

Now expressing $\mathcal{L}_{\lambda}(x_1, x_2, \lambda) = 0$ in terms of λ only, we get:

$$4 - \frac{9}{2}\lambda - \frac{1}{2}\lambda = 0 \implies \lambda = 4/5$$

Thus we obtain the values of x_1 and x_2 to minimize the function $f(x_1, x_2)$:

$$\boxed{\left.\begin{aligned} x_1 &= 2/5 \\ x_2 &= 6/5 \end{aligned}\right\} \implies f(x_1 = 2/5, x_2 = 6/5) = 8/5 \quad | \quad g(x_1 = 2/5, x_2 = 6/5) = 0}$$

Problem 7

Part a

Let $\chi = [x \quad y]^T$, the joint distribution is given by:

$$f_\chi(x, y) = \frac{1}{\sqrt{(2\pi)^2 det(\Sigma_\chi)}} e^{-\frac{1}{2}(\chi - \mu_\chi)^T \Sigma_\chi^{-1} (\chi - \mu_\chi)}$$

Where Σ_χ , the covariance matrix, and μ_χ , the mean vector, are defined as:

$$\Sigma_\chi = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma y^2 \end{bmatrix} = \begin{bmatrix} 3 & \sqrt{5} \\ \sqrt{5} & 2 \end{bmatrix}$$
$$\mu_\chi = \begin{bmatrix} \mu_x & \mu_y \end{bmatrix}^T = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$$

Computing the Mahalanobis distance and replacing it in the joint distribution, we get:

$$\begin{aligned}
(\chi - \mu_\chi)^T \Sigma_\chi^{-1} (\chi - \mu_\chi) &= \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix}^T \begin{bmatrix} 2 & -\sqrt{5} \\ -\sqrt{5} & 3 \end{bmatrix} \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix} \\
&= \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix}^T \begin{bmatrix} 2(x - \mu_x) - \sqrt{5}(y - \mu_y) \\ 3(y - \mu_y) - \sqrt{5}(x - \mu_x) \end{bmatrix} \\
&= 2(x - \mu_x)^2 - 2\sqrt{5}(x - \mu_x)(y - \mu_y) + 3(y - \mu_y)^2 \\
&= 2(x - 1)^2 - 2\sqrt{5}(x - 1)(y - 2) + 3(y - 2)^2 \\
\implies f_\chi(x, y) &= \frac{1}{\sqrt{(2\pi)^2 \begin{vmatrix} 3 & \sqrt{5} \\ \sqrt{5} & 2 \end{vmatrix}}} e^{-\frac{1}{2}(2(x-1)^2 - 2\sqrt{5}(x-1)(y-2) + 3(y-2)^2)} \\
&= \frac{1}{2\pi} e^{-\frac{1}{2}(2(x-1)^2 - 2\sqrt{5}(x-1)(y-2) + 3(y-2)^2)}
\end{aligned}$$

The marginal distributions of X and Y can be computed using the following formula:

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_\chi(x, y) dy \\
&= \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(2(x-1)^2 - 2\sqrt{5}(x-1)(y-2) + 3(y-2)^2)} dy \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(2(x-1)^2 - 2\sqrt{5}(x-1)(y-2) + 3(y-2)^2)} dy \\
&= \frac{1}{2\pi} e^{-(x-1)(x+2\sqrt{5}-1)} \int_{-\infty}^{\infty} e^{-\frac{3(y-2)^2 - 2\sqrt{5}(x-1)y}{2}} dy \\
&= \frac{1}{2\pi} e^{-(x-1)(x+2\sqrt{5}-1)} \int_{-\infty}^{\infty} e^{\sqrt{5}xy - \sqrt{5}y - \frac{3(y-2)^2}{2}} dy \\
&= \frac{1}{2\pi} e^{-(x-1)(x+2\sqrt{5}-1)} \int_{-\infty}^{\infty} e^{-(\frac{\sqrt{3}y}{\sqrt{2}} - \frac{\sqrt{5}x - \sqrt{5} + 6}{\sqrt{6}}) + \frac{(\sqrt{5}x - \sqrt{5} + 6)^2}{6} - 6} dy
\end{aligned}$$

Now we define $u = \frac{3y - \sqrt{5}x + \sqrt{5} - 6}{\sqrt{6}}$ and $\frac{du}{dy} = \frac{3}{\sqrt{6}}$ and we get:

$$
\begin{aligned}
f_X(x) &= \frac{1}{2\pi}e^{-(x-1)(x+2\sqrt{5}-1)} * \frac{\pi e^{\frac{\sqrt{5}x-\sqrt{5}+6}{6}-6}}{\sqrt{6}}\int_{-\infty}^{\infty}\frac{2e^{-u^2}}{\sqrt{\pi}}du \\
&= \frac{1}{2\pi}e^{-(x-1)(x+2\sqrt{5}-1)} * \frac{\pi e^{\frac{\sqrt{5}x-\sqrt{5}+6}{6}-6}}{\sqrt{6}}[erf(u)]_{-\infty}^{\infty} \\
&= \frac{e^{-x^2-2\sqrt{5}x+2x+2\sqrt{5}-7+\frac{(\sqrt{5}(1-x)-6)^2}{6}} * [erf(\frac{\sqrt{3}y}{2}+\frac{\sqrt{5}(1-x)-6}{\sqrt{6}})]_{-\infty}^{\infty}}{2\sqrt{6\pi}}
\end{aligned}
$$

$$
\boxed{f_X(x) = \frac{e^{-\frac{(x-1)^2}{6}}}{\sqrt{6\pi}}}
$$

We do the same for $f_Y(y)$ and obtain:

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_X(x,y)dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(2(x-1)^2-2\sqrt{5}(x-1)(y-2)+3(y-2)^2)}dx
\end{aligned}
$$

$$
\boxed{f_Y(y) = \frac{e^{-\frac{(y-2)^2}{4}}}{2\sqrt{\pi}}}
$$

Part b

$$
\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \frac{\frac{1}{2\pi}e^{-\frac{1}{2}(2(x-1)^2-2\sqrt{5}(x-1)(y-2)+3(y-2)^2)}}{\frac{e^{-\frac{(y-2)^2}{4}}}{2\sqrt{\pi}}}
\end{aligned}
$$

$$
\boxed{f_{X|Y}(x|y) = \frac{e^{-((x-1)^2-\sqrt{5}(x-1)(y-2)+\frac{5}{4}(y-2)^2)}}{\sqrt{\pi}}}
$$

Part c

The variance of X given $Y = y$ denoted as $\sigma^2_{X|Y=y}$ is given by:

$$
\sigma^2_{X|Y=y} = (1-\rho^2)\sigma^2_X
$$

where $\rho \in (-1, 1)$ is computed as follows:

$$
\rho = \frac{\sigma_{xy}}{\sigma_x\sigma_y} = \sqrt{\frac{5}{6}}
$$

Even if we do not know the value of ρ, we can still compare $\sigma^2_{X|Y=y}$ and σ^2_X using the properties of ρ as follows:

$$
-1 < \rho < 1
$$

$$0 < \rho^2 < 1$$
$$0 < 1 - \rho^2 < 1$$
$$0 < \frac{\sigma^2_{X|Y=y}}{\sigma^2_X} < 1$$

Thus, we can see that for any value $Y = y$, the variance of X given $Y = y$ denoted as $\sigma^2_{X|Y=y}$ is always lesser than the variance of X denoted by σ^2_X.

Part d

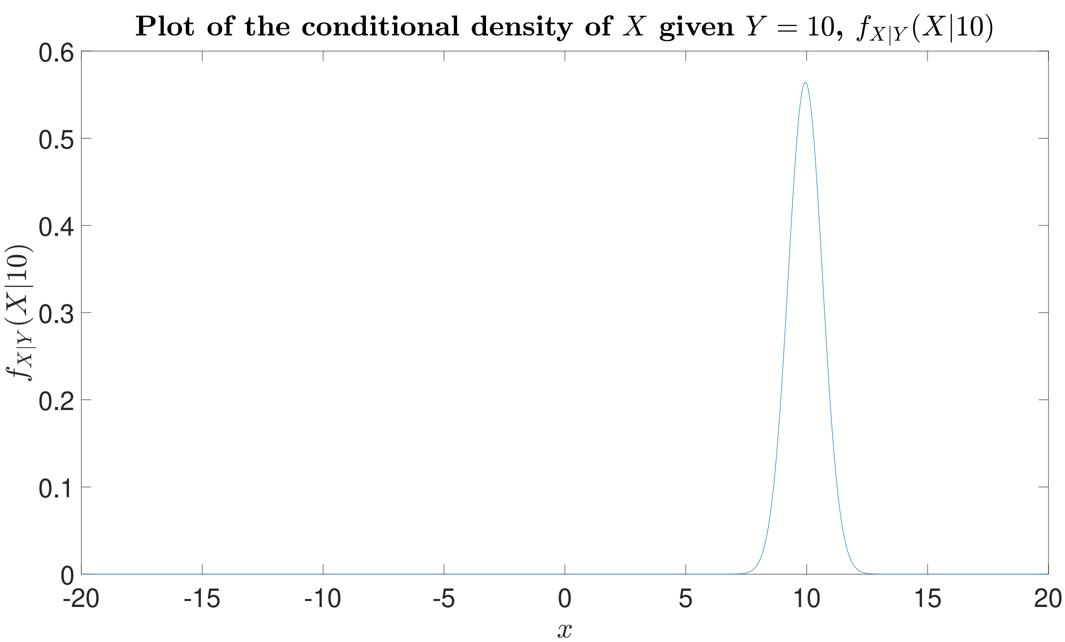

Figure 2: Plot of the conditional density of X given $Y = 10$, $f_{X|Y}(X|10)$

```
%% Theodor Chakhachiro ROB 501 HW1 Problem 7
clear all
clc
%% Initializing
n = 20;
x = [-n:.01:n];
f = p7func(x,10);
figure(1);
plot(x,f)
title('\textbf{Plot of the conditional density of $X$ given $Y=10$, $f_{X|Y}(X|10)$}','
    Interpreter','latex')
xlabel('$x$','Interpreter','latex')
ylabel('$f_{X|Y}(X|10)$','Interpreter','latex')
set(gca,'fontsize',40)
%% End
```

```
function f = p7func(x,y)
f=1/sqrt(pi)*exp(-((x-1).^2-sqrt(5)*(x-1).*(y-2)+5/4*(y-2).^2));
end
```