Self-Assessment Problem Set

NAME: Harsh Bhate **GTID: 903424029**

Problem	**Points**
Problem 1	
Problem 2	
Problem 3	
Problem 4	
Problem 5	
Problem 6	
Problem 7	
Problem 8	
Problem 9	
Problem 10	
Total	

Problem 1

(a). Since the variables X and Y are uncorrelated,

$$\begin{aligned} E((X - E(X))(Y - E(Y))) = 0 \\ \implies \mathbb{E}(XY - X \cdot \mathbb{E}(X) - \mathbb{E}(Y) \cdot Y + \mathbb{E}(X)\mathbb{E}(Y)) = 0 \\ \implies \mathbb{E}(XY) - 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y) = 0 \\ \implies \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

(b). The Variance $var(X)$ is given by:

$$var(X) = \mathbb{E}[(X - E(X))^2]$$

Let us now consider $var(\sum_{i=1}^{n} X_i)$,

$$\begin{aligned} var(\sum_{i=1}^{n} X_i) &= \mathbb{E}\left[\sum_{i=1}^{n}(X_i - E(X_i))^2\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{n} X_i \sum_{j=1}^{n} X_j\right] - 2 \cdot \mathbb{E}[X_i]\mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{n} X_i\right]\right] + \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{n} X_i\right]\mathbb{E}\left[\sum_{i=1}^{n} X_i\right]\right] \end{aligned}$$

$\because cov(A, B) = E(AB) = 0,$

$$\begin{aligned} &= \sum_{i=1}^{n} \mathbb{E}\left[X_i X_i\right] - 2 \cdot \mathbb{E}[X_i] \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}\left[X_i\right]\right] + \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}\left[X_i\right]\mathbb{E}\left[X_i\right]\right] && \text{[Linearity]} \\ &= \sum_{i=1}^{n} \mathbb{E}\left[X_i^2 - 2X_i\mathbb{E}[X_i] + \mathbb{E}[X_i]\mathbb{E}[X_i]\right] \\ &= \sum_{i=1}^{n} \mathbb{E}\left[(X_i - \mathbb{E}[X_i)^2\right] \\ &= \sum_{i=1}^{n} var(X_i) \end{aligned}$$

(c). Let us consider two independent random variables X and Y. Since X and Y are independent, $P(X, Y) = P(X) \cdot P(Y)$. The expectation function $\mathbb{E}$ is given by:

$$\mathbb{E}(X) = \sum_{x} x \cdot P(X = x)$$

Let us now evaluate $\mathbb{E}(XY)$,

$$\begin{aligned}\mathbb{E}(XY) &= \sum_x \sum_y x \cdot y \cdot P(X = x, Y = y) \\ &= \sum_x \sum_y x \cdot y \cdot P(X = x) \cdot P(Y = y) \qquad [\because \text{X,Y are independent}] \\ &= \sum_x x \cdot P(X = x) \sum_y y \cdot P(Y = y) \\ &= \mathbb{E}(X) \cdot \mathbb{E}(Y)\end{aligned}$$

Thus, X and Y are uncorrelated. Extending the same line of argument, it can be proved that n independent random variables are all uncorrelated.

(d). The mean of X and Y is given by:

$$\begin{aligned}\mu_X &= \mathbb{E}(X) \\ &= \int_{-\pi}^{\pi} cos(\theta) d\theta \\ &= 0\end{aligned}$$

$$\begin{aligned}\mu_Y &= \mathbb{E}(Y) \\ &= \int_{-\pi}^{\pi} sin(\theta) d\theta \\ &= 0\end{aligned}$$

The variables X and Y are unorrelated if:

$$cov(X, Y) = \mathbb{E}\Big((X - \mu_X)(Y - \mu_Y)\Big) = 0$$

Thus,

$$\begin{aligned}cov(X, Y) &= \mathbb{E}\Big((X - \mu_X)(Y - \mu_Y)\Big) \\ &= \mathbb{E}\Big(XY\Big) \\ &= \int_{-\pi}^{\pi} cos(\theta) sin(\theta) d\theta \\ &= 0\end{aligned}$$

Thus, X and Y are uncorrelated. Recall the identity,

$$\begin{aligned} sin^2(\theta) + cos^2(\theta) = 1 \\ \implies cos(\theta) = \sqrt{1 - sin^2(\theta)}\end{aligned}$$

Thus, X and Y are dependent.

Problem 2

(a). The Expectation function, $\mathbb{E}$, is given by,

$$\mathbb{E}(XY) = \int_x \int_y (x \cdot y) \cdot P_{XY} dydxs$$

Thus,

$$\begin{aligned} E(XY) &= \int_0^1 \int_0^u (u \cdot v) \cdot (4u^2) \cdot dv \cdot du \\ &= \int_0^1 \int_0^u v \cdot (4u^3) \cdot dv \cdot du \\ &= \int_0^1 (4u^3) \cdot \left(\frac{v^2}{2}\right)_0^u \cdot du \\ &= \int_0^1 2u^4 \cdot du \\ &= \left(2 \cdot \frac{u^6}{6}\right)_0^1 \\ &= \frac{1}{3} \end{aligned}$$

(b).

$$\begin{aligned} p_Y &= \int_{-\infty}^{\infty} p_{uv} du \\ &= \int_v^1 4u^2 du \\ &= \left.\frac{4u^3}{3}\right|_v^1 \\ &= \frac{4}{3} - \frac{4v^3}{3} \\ &= \frac{4}{3}(1 - v^3) \end{aligned}$$

(c).

$$\begin{aligned} p_{X|Y} &= \frac{p_{XY}}{p_Y} \\ &= \frac{p_{XY}}{\frac{4}{3}(1 - v^3)} \\ &= \frac{3}{4(1 - v^3)} p_{xy} \end{aligned}$$

Thus,

$$p_{X|Y} = \begin{cases} \frac{3u^2}{(1-v^3)} & \text{if } 0 < v < u < 1 \\ 0 & \text{otherwise} \end{cases}$$

(d).

$$\begin{aligned}
\mathbb{E}(X^2|Y=y) &= \int_0^y x^2 p_{X|Y}(x|y)dx && [\because 0 < y < 1] \\
&= \int_0^y x^2 \frac{3x^2}{1-y^3} dx \\
&= \frac{3}{1-y^3} \int_0^y x^4 dx \\
&= \frac{3}{1-y^3} \left[\frac{x^5}{5}\right]_0^y \\
&= \frac{3}{1-y^3} \left[\frac{y^5}{5}\right] \\
&= \frac{3y^5}{5(1-y^3)}
\end{aligned}$$

Problem 3

The distribution of U and V is:

$$U(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$V(y) = \begin{cases} 1 & \text{if } 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Since both U and V are uniform distribution, the PDF of $X = UV$ is merely the multiplication of the PDFs of U and V respectively because of the independent nature of distribution. Thus,

$$X(x, y) = \begin{cases} 1 & \text{if } 0 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

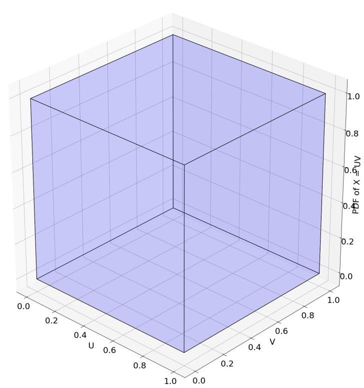
The CDF of $X(F_X)$ is thus, given by:

$$\begin{aligned} F_X &= \int_{-\infty}^{y} \int_{-\infty}^{x} X(x, y) dx dy \\ &= \int_{0}^{y} \int_{0}^{x} dx dy \\ &= xy \end{aligned}$$

Thus,

$$F_X(x, y) = \begin{cases} 0 & \text{if } 0 < x, y \\ xy & \text{if } 0 < x, y < 1 \\ 1 & \text{otherwise} \end{cases}$$

The plot of the PDF of U, V is:



Problem 4

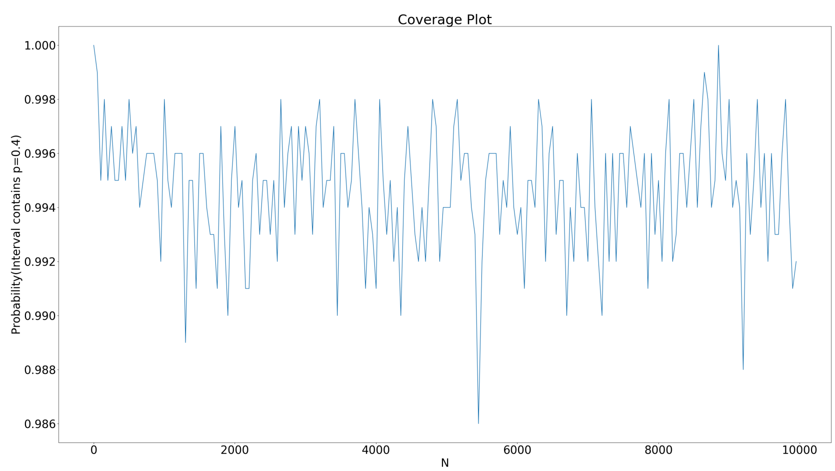
(a). In the special case of independent and identically distributed Bernouilli distribution, the Hoeffding Equality translates to:

$$\mathbf{P}\big((p-\epsilon)n \leq H(n) \leq (p+\epsilon)n\big) \geq 1 - 2\exp\left(-2\epsilon^2 n\right)$$

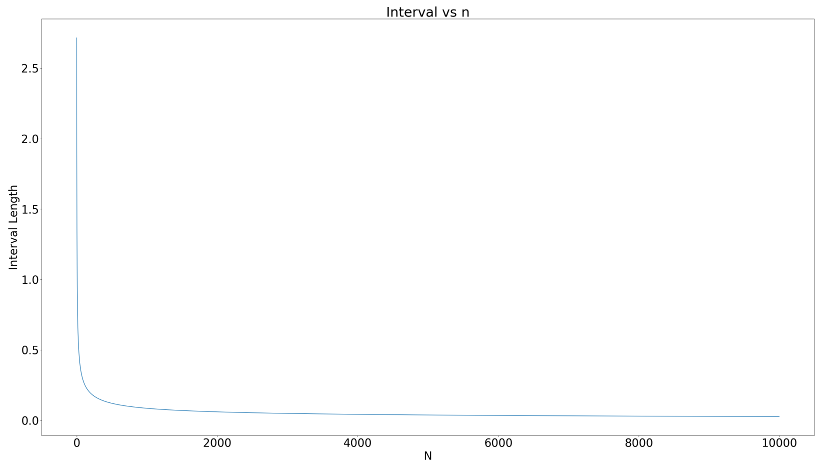
Substituting the values provided in the problem, we get,

$$\begin{aligned}\mathbf{P}\big((\hat{p}_n-\epsilon)n \leq p \leq (\hat{p}_n+\epsilon)n\big) &\geq 1 - 2\exp\left(-2\epsilon^2 n\right)\\ &\geq 1 - 2\exp\Big(-2\Big(\sqrt{\frac{1}{2n}\ log\frac{2}{\alpha}}\Big)^2 n\Big)\\ &\geq 1 - 2\exp\Big(-log\Big(\frac{2}{\alpha}\Big)\Big)\\ &\geq 1 - 2\cdot\frac{\alpha}{2}\\ &\geq 1-\alpha\end{aligned}$$

(b).



(c). For interval to be no more than 0.05, the value of n should be $n \geq 2952$. The plot of interval length versus n is shown below:


Interval vs n
Interval Length
2.5
2.0
1.5
1.0
0.5
0.0
0
2000
4000
6000
8000
10000
N

Problem 5

The parameters of the Normal distribution are:

$$\mathcal{N}(0,1) \implies \mu = 0, \sigma^2 = 1$$

Thus, the pdf ($f(x)$) of Z is:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)$$

Let us now consider the expression $\mathbb{P}(|Z|> t)$.

$$\mathbb{P}(|Z|> t) = \mathbb{P}(Z^- < t) + \mathbb{P}(Z^+ > t)$$

Since Z is a normal distribution with zero mean, it is symmetric about 0. Thus,

$$\mathbb{P}(|Z|> t) = 2\mathbb{P}(Z > t)$$

Using Markov's inequality, we can expand the above statement as:

$$\begin{aligned}
\mathbb{P}(|Z|> t) &= 2\mathbb{P}(Z > t) \\
&\leq 2\frac{\mathbb{E}(X)}{t} \\
&\leq \frac{2}{t}\int_t^\infty x f(x) dx \\
&\leq \frac{2}{t}\int_t^\infty x \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big) dx \\
&\leq \frac{2}{t}\Bigg[\frac{1}{\sqrt{2\pi}}\Big(-\exp\Big(-\frac{x^2}{2}\Big)\Big)\Bigg]_t^\infty \\
&\leq \frac{2}{t}\frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{t^2}{2}\Big) \\
&\leq \sqrt{\frac{2}{\pi}}\frac{e^{\frac{-t^2}{2}}}{t}
\end{aligned}$$

Problem 6

(a). The conditional PDF is given by:

$$f_{X|Y} = \frac{f_{XY}}{f_Y}$$

In case of a bivariate normal distribution, the conditional distribution is given by:

$$X|Y = \mathcal{N}(\mu_X + \frac{\sigma_X}{\sigma_Y} k(y - \mu_Y), (1 - k^2)\sigma_X^2)$$

where

$$\begin{aligned} k &= corr(X, Y) \\ &= \frac{cov(X, Y)}{\sigma_X \sigma_Y} \\ &= \frac{\rho}{1 \cdot 1} && \text{[From Covariance Matrix } K\text{]} \\ &= \rho \end{aligned}$$

Thus,

$$\begin{aligned} \mu_X|Y &= \mu_X + \frac{\sigma_X}{\sigma_Y} k(y - \mu_Y) \\ &= 0 + \frac{1}{1}\rho(y - 0) && [X, Y \text{ are zero mean}] \\ &= \rho y \\ \sigma_X|Y &= (1 - k^2)\sigma_X^2 \\ &= (1 - \rho^2)(1)^2 \\ &= 1 - \rho^2 \end{aligned}$$

Also the PDF is given by,

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

Now,

$$P(X \leq x|Y = y) = \int_{-\infty}^{x} f_{X|Y} dx \tag{1}$$

The CDF (ϕ) for 1-D normal distribution is given by:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} \tag{2}$$

From (1) and (2), we can conclude that the solution is:

$$P(X \leq 1|Y = y) = \phi\left(\frac{1 - \rho y}{\sqrt{1 - \rho^2}}\right)$$

(b).

$$\mathbb{E}((X-Y)^2|Y=y) = \mathbb{E}(X^2|Y=y) - 2\mathbb{E}(XY|Y=y) + \mathbb{E}(Y^2|Y=y)$$

where

$$\mathbb{E}(g(x,y)|Y=y) = \int g(x,y) f_{X|Y}\, dx$$

Upon solving and simplifying, we get,

$$\mathbb{E}((X-Y)^2|Y=y) = (\rho-1)^2 y^2 + (1-\rho^2)$$

Problem 7

(a).

$$||\mathbf{x}||_2^2 = \sum_i x_i^2 = \mathbf{x} \cdot \mathbf{x} = x^2$$

Thus,

$$\begin{aligned} ||x+y||_2^2 &= ||x||_2^2 + ||y||_2^2 \\ \sqrt{(x+y)^2}^2 &= \sqrt{x^2}^2 + \sqrt{y^2}^2 \\ x^2 + y^2 + 2xy &= x^2 + y^2 \\ \implies 2xy &= 0 \\ \implies x \cdot y &= 0 \end{aligned}$$

Thus, $||x+y||_2^2 = ||x||_2^2 + ||y||_2^2$ holds true only when x and y are normal to each other.

(b).

$$||\mathbf{x}||_2 = \sqrt{\sum_i x_i^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x^2}$$

$$\begin{aligned} ||x+y||_2 &= ||x||_2 + ||y||_2 \\ \sqrt{(x+y)^2} &= \sqrt{x^2} + \sqrt{y^2} \\ \sqrt{(x+y)^2}^2 &= \left(\sqrt{x^2} + \sqrt{y^2}\right)^2 \\ x^2 + y^2 + 2xy &= x^2 + y^2 + 2\sqrt{x^2 y^2} \\ xy &= \sqrt{(xy)^2} \\ \implies |x||y|cos(\theta) &= |x||y| \\ \implies x \cdot cos(\theta) &= |1| && [\because \text{Square root}] \\ \implies \theta &= 0 \text{ or } \pi \end{aligned}$$

Thus, for $||x+y||_2 = ||x||_2 + ||y||_2$ to be true, x and y need to be parallel (in the same direction or otherwise) to each other.

Problem 8

(a). Since $\mathbf{A}$ and $\mathbf{B}$ are square symmetric matrices, they are diagonalizable. Thus,

$$\mathbf{A} = \mathbf{Q_1 D_1 Q_1}^{-1}$$

$$\mathbf{B} = \mathbf{Q_2 D_2 Q_2}^{-1}$$

Also, the matrices $\mathbf{A}$ and $\mathbf{B}$ have the same Eigenvectors. Thus, the matrices $\mathbf{Q_1} = \mathbf{Q_2} = \mathbf{Q}$.

$$\begin{aligned}
\mathbf{AB} &= \mathbf{Q D_1 Q}^{-1}\mathbf{Q D_2 Q}^{-1} \\
&= \mathbf{Q D_1 D_2 Q}^{-1} && [\because \mathbf{Q}^{-1}\mathbf{Q} = \mathbb{I}] \\
&= \mathbf{Q D_2 D_1 Q}^{-1} && [\because \mathbf{D_1}, \mathbf{D_2} \text{ are diagonal } \implies \mathbf{D_1}, \mathbf{D_2} \text{ Commute}] \\
&= \mathbf{Q D_2} \mathbb{I} \mathbf{D_1 Q}^{-1} \\
&= \mathbf{Q D_2 Q}^{-1}\mathbf{Q D_1 Q}^{-1} && [\because \mathbf{Q}^{-1}\mathbf{Q} = \mathbb{I}] \\
&= \mathbf{BA}
\end{aligned}$$

$$\therefore \mathbf{AB} = \mathbf{BA}$$

(b). Let $\mathbf{v}$ be an eigenvector of A corresponding the the eigenvale λ. We know that,

$$\mathbf{AB} = \mathbf{BA}$$

Multiplying both sides by $\mathbf{v}$, we get,

$$\begin{aligned}
&\mathbf{ABv} = \mathbf{BAv} \\
\implies &\mathbf{ABv} = \mathbf{B}\lambda\mathbf{v} \\
\implies &\mathbf{ABv} = \mathbf{B}\lambda\mathbf{v} \\
\implies &\mathbf{A}(\mathbf{Bv}) = \lambda(\mathbf{Bv}) \qquad (1)
\end{aligned}$$

From (1), we can conclude that both $\mathbf{v}$ and $\mathbf{Bv}$ are eigenvectors of A corresponding to the same eigenvalue (λ). Since, A has no repeated eigenvalues and thus no repeated eigenvectors exist, we can safely conclude that:

$$k\mathbf{v} = \mathbf{Bv} \text{ for some constant } k$$

Observing the structure of the statement above, we can conclude that k is an eigenvalue of $\mathbf{B}$ with the eigenvector $\mathbf{v}$.

Therefore, $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvectors.

Problem 9

(a). Performing Column Reduction operation on A, we obtain:

$$\hat{A} = \begin{bmatrix} -2 & 2 & 2 & -2 & 0 \\ -2 & 2 & 2 & -2 & 0 \\ 2 & -2 & -2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & -2 & 0 \\ -2 & 0 & 0 & -2 & 0 \\ 2 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \left[\frac{C_2 \rightarrow C_2 + C_1}{C_3 \rightarrow C_3 + C_1}\right]$$

$$= \begin{bmatrix} -2 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{bmatrix} \left[\frac{C_2 \leftrightarrow C_4; C_3 \leftrightarrow C_5}{C_2 \rightarrow C_2 - C_1}\right] = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \text{[Normalizing]}$$

Since, rank of matrix is the dimension of it's column space, the rank of A is 3.

(b). The SVD of A is:

$$\mathbf{U} = \begin{bmatrix} -0.628 & -0.325 & 0 & 0.707 & 0 \\ -0.628 & -0.325 & 0 & -0.707 & 0 \\ 0.460 & -0.888 & 0 & 0 & 0 \\ 0 & 0 & 0.707 & 0 & -0.707 \\ 0 & 0 & 0.707 & 0 & 0.707 \end{bmatrix}$$

$$\mathbf{\Sigma} = diag\left(6.153, 3.185, 2.828, 0, 0\right)$$

$$\mathbf{V^T} = \begin{bmatrix} 0.558 & -0.558 & -0.558 & 0.259 & 0 \\ -0.149 & 0.149 & 0.149 & 0.966 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0.791 & 0.571 & 0.221 & 0 & 0 \\ 0.202 & -0.584 & 0.786 & 0 & 0 \end{bmatrix}$$

(c). The column space of a matrix is the set of independent column vectors the span of which is the set of all the column vectors of the matrix. From the Column Reduction Operations in part (a):

$$\text{Column Space}\{A\} \in span\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

(d). The row space of a matrix is the set of independent row vectors the span of which is the set of all the row vectors of the matrix. Performing Column Operations similar to the one in part (a) on the matrix A^T,

$$\text{Row Space}\{A\} \in span\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}^T ; \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}^T ; \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \right\}$$

Problem 10

To find the Orthogonal Projection, Gram-Schmidt orthogonalization process was used. The process involves finding a set of orthonormal basis $\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}$ from the vectors $\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}$. The set of orthnormal vectors are calculated as follows:

$$\begin{aligned} \mathbf{q_i} &= \frac{\mathbf{u_i}}{\sqrt{\mathbf{u_i}\mathbf{u_i}}} \\ \mathbf{u_1} &= \mathbf{p_1} \\ \mathbf{u_2} &= \mathbf{p_2} - \frac{\mathbf{u_1}\mathbf{p_2}}{\mathbf{u_1}\mathbf{u_1}}\mathbf{u_1} \\ \mathbf{u_2} &= \mathbf{p_3} - \frac{\mathbf{u_1}\mathbf{p_3}}{\mathbf{u_1}\mathbf{u_1}}\mathbf{u_1} - \frac{\mathbf{u_2}\mathbf{p_3}}{\mathbf{u_3}\mathbf{u_3}}\mathbf{u_3} \end{aligned}$$

The Orthonormal Basis $Q = [\mathbf{q_1}, \mathbf{q_2}.\mathbf{q_3}]$ is:

$$Q = \begin{bmatrix} -0.1826 & 0.9425 & -0.2747 \\ -0.3651 & 0.1817 & 0.7708 \\ -0.5477 & -0.2385 & -0.5647 \\ -0.7303 & -0.1476 & 0.1068 \end{bmatrix}$$

Next, the vector $\mathbf{x}^*$ is a linear combination of $\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}$ and can be found using:

$$\mathbf{x}^* = (\mathbf{x} \cdot \mathbf{q_1})\mathbf{q_1} + (\mathbf{x} \cdot \mathbf{q_2})\mathbf{q_2} + (\mathbf{x} \cdot \mathbf{q_3})\mathbf{q_3}$$

And $\mathbf{x_e}$ can be found using $\mathbf{x_e} = \mathbf{x} - \mathbf{x}^*$. Thus,

$$\begin{aligned} \mathbf{x}^* &= [0.8946, 2.9665, 4.1247, 5.6996]^T \\ \mathbf{x_e} &= [0.1054, -0.9665, -1.1247, 1.3004]^T \end{aligned}$$