Sparkl ML

Homework 1 (Part 1)

(Q2) $A = \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \end{bmatrix}$

$B = \begin{bmatrix} 5 & 6 & 7 \\ a & b & c \\ 4 & 5 & 7 \end{bmatrix}$

* $AB = \begin{bmatrix} a & b & c \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 \\ a & b & c \\ 4 & 5 & 7 \end{bmatrix}$

$= \begin{bmatrix} 5a + ab + 4c & 6a + b^2 + 5c & 7a + bc + 7c \\ 17 + 2a & 21 + 2b & 28 + 2c \end{bmatrix}$

* B.A is not possible since A is 2x3 and B is 3x3 and a matrix multipication of type (3x3) x (2x3) is not possible since the inner dimension does not match.

(Q3) $M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$

* Reducing the matrix M to Echelon form,

Performing $R_2 \leftarrow R_2 - 5R_1$, $R_3 \leftarrow R_3 - 9R_1$, & $R_4 \leftarrow R_4 - 13R_1$,
we get

$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -24 \\ 0 & -12 & -24 & -36 \end{bmatrix}$

Performing $R_3 \leftarrow R_3 - 2R_2$ & $R_4 \leftarrow R_4 - 3R_2$

$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

* Matrix M is now in Echelon form.

* For an Echelon Matrix, the rank of the matrix is determined by the no. of non-zero rows.

* Since there are 2 non-zero rows in M after converting it to Echelon form, the rank of matrix M is 2.

* As mentioned, I converted the matrix to Echelon form to find the rank (algorithm).

(Q4) $M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$

$u_1 = \begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \end{bmatrix}$ $u_2 = \begin{bmatrix} 2 \\ 6 \\ 10 \\ 14 \end{bmatrix}$ $u_3 = \begin{bmatrix} 3 \\ 7 \\ 11 \\ 15 \end{bmatrix}$ $u_4 = \begin{bmatrix} 4 \\ 8 \\ 12 \\ 16 \end{bmatrix}$

* To find out the orthonormal basis, we can use the Gram-Schmidt process.

* According to the Gram-Schmidt process,

$$v_i = u_i - \sum_{j=1}^{i-1} \frac{(u_i \cdot v_j)}{(v_j \cdot v_j)} \cdot v_j.$$

* $v_1 = u_1 = \begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \end{bmatrix}$

* Normalizing v_1, we get.

$$v_1 = \frac{v_1}{|v_1|} = \frac{\begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \end{bmatrix}}{\sqrt{276}} = \frac{\begin{bmatrix} 1 \\ 5 \\ 9 \\ 13 \end{bmatrix}}{16.61} = \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix}$$

So, $v_1 = \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix}$ $\cancel{|v_1| = 0.9936}$

$$v_2 = u_2 - \frac{(u_2 \cdot v_1)}{(v_1 \cdot v_1)} \cdot v_1$$

$$\frac{(u_2 \cdot v_1)}{(v_1 \cdot v_1)} v_1 = \frac{18.24}{0.9936} \times \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix} = 18.36 \times \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix}$$

$$= \begin{bmatrix} 1.1016 \\ 5.508 \\ 9.9144 \\ 14.3208 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 2 \\ 6 \\ 10 \\ 14 \end{bmatrix} - \begin{bmatrix} 1.1016 \\ 5.508 \\ 9.9144 \\ 14.3208 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 0.9 \\ 0.49 \\ 0.08 \\ -0.32 \end{bmatrix} \Rightarrow$$ Normalizing v_2, we get:

$$v_2 = \frac{\begin{bmatrix} 0.9 \\ 0.49 \\ 0.08 \\ -0.32 \end{bmatrix}}{\sqrt{1.1589}}$$

$$v_2 = \frac{\begin{bmatrix} 0.9 \\ 0.49 \\ 0.08 \\ -0.32 \end{bmatrix}}{1.08} = \begin{bmatrix} 0.83 \\ 0.46 \\ 0.08 \\ -0.3 \end{bmatrix}$$

$$v_3 = u_3 - \left[\frac{u_3 \cdot v_1}{(v_1 \cdot v_1)} v_1 + \frac{(u_3 v_2)}{(v_2 \cdot v_2)} \cdot v_2 \right]$$

P.T.O

$$\frac{(u_3 \cdot v_1)}{(v_1 \cdot v_1)} v_1 = \frac{19.92}{0.9936} \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix} = 20 \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix}$$

$$= \begin{bmatrix} 1.2 \\ 6 \\ 10.8 \\ 15.6 \end{bmatrix}$$

$$\frac{(u_3 \cdot v_2)}{(v_2 \cdot v_2)} \cdot v_2 = \frac{2.09}{0.9969} \begin{bmatrix} 0.83 \\ 0.46 \\ 0.08 \\ -0.3 \end{bmatrix} = 2.1 \begin{bmatrix} 0.83 \\ 0.46 \\ 0.08 \\ -0.3 \end{bmatrix}$$

$$= \begin{bmatrix} 1.74 \\ 0.97 \\ 0.17 \\ -0.63 \end{bmatrix}$$

$$v_3 = \begin{bmatrix} 3 \\ 7 \\ 11 \\ 15 \end{bmatrix} - \left(\begin{bmatrix} 1.2 \\ 6 \\ 10.8 \\ 15.6 \end{bmatrix} + \begin{bmatrix} 1.74 \\ 0.97 \\ 0.17 \\ -0.63 \end{bmatrix} \right)$$

$$v_3 = \begin{bmatrix} 3 \\ 7 \\ 11 \\ 15 \end{bmatrix} - \begin{bmatrix} 2.94 \\ 6.97 \\ 10.97 \\ 14.97 \end{bmatrix} = \begin{bmatrix} 0.06 \\ 0.03 \\ 0.03 \\ 0.03 \end{bmatrix}$$

Normalizing v_3, we get

$$v_3 = \frac{\begin{bmatrix} 0.06 \\ 0.03 \\ 0.03 \\ 0.03 \end{bmatrix}}{\sqrt{0.0063}} = \frac{\begin{bmatrix} 0.06 \\ 0.03 \\ 0.03 \\ 0.03 \end{bmatrix}}{0.08} = \begin{bmatrix} 0.75 \\ 0.38 \\ 0.38 \\ 0.38 \end{bmatrix}$$

$$v_4 = u_4 - \left(\frac{(u_4 \cdot v_1)}{(v_1 \cdot v_1)} v_1 + \frac{(u_4 \cdot v_2)}{(v_2 \cdot v_2)} v_2 + \frac{(u_4 \cdot v_3)}{(v_3 v_3)} v_3 \right)$$

$$\frac{(u_4 \cdot v_1)}{(v_1 \cdot v_1)} v_1 = \frac{21.6}{0.9936} \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix} = 21.7 \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix}$$

$$= \begin{bmatrix} 1.3 \\ 6.51 \\ 11.72 \\ 16.93 \end{bmatrix}$$

$$\frac{(u_4 \cdot v_2)}{(v_2 \cdot v_2)} v_2 = \frac{3.16}{0.9969} \begin{bmatrix} 0.83 \\ 0.46 \\ 0.08 \\ -0.3 \end{bmatrix} = 3.17 \begin{bmatrix} 0.83 \\ 0.46 \\ 0.08 \\ -0.3 \end{bmatrix}$$

$$= \begin{bmatrix} 2.63 \\ 1.46 \\ 0.25 \\ -0.95 \end{bmatrix}$$

$$\frac{(u_4 \cdot v_3)}{(v_3 \cdot v_3)} v_3 = \frac{16.68}{0.9957} \begin{bmatrix} 0.75 \\ 0.38 \\ 0.38 \\ 0.38 \end{bmatrix} = 16.75 \begin{bmatrix} 0.75 \\ 0.38 \\ 0.38 \\ 0.38 \end{bmatrix}$$

$$= \begin{bmatrix} 12.56 \\ 6.36 \\ 6.36 \\ 6.36 \end{bmatrix}$$

$$v_4 = \begin{bmatrix} 4 \\ 8 \\ 12 \\ 16 \end{bmatrix} - \left(\begin{bmatrix} 1.3 \\ 6.51 \\ 11.72 \\ 16.93 \end{bmatrix} + \begin{bmatrix} 2.63 \\ 1.46 \\ 0.25 \\ -0.95 \end{bmatrix} + \begin{bmatrix} 12.56 \\ 6.36 \\ 6.36 \\ 6.36 \end{bmatrix} \right)$$

$$V_4 = \begin{bmatrix} 4 \\ 8 \\ 12 \\ 16 \end{bmatrix} - \begin{bmatrix} 16.49 \\ 14.33 \\ 18.33 \\ 22.34 \end{bmatrix} = \begin{bmatrix} -12.49 \\ -6.33 \\ -6.33 \\ -6.34 \end{bmatrix}$$

Normalizing V_4, we get

$$V_4 = \frac{\begin{bmatrix} -12.49 \\ -6.33 \\ -6.33 \\ -6.34 \end{bmatrix}}{\sqrt{276.33}} = \frac{\begin{bmatrix} -12.49 \\ -6.33 \\ -6.33 \\ -6.34 \end{bmatrix}}{16.62} = \begin{bmatrix} -0.75 \\ -0.38 \\ -0.38 \\ -0.38 \end{bmatrix}$$

* $V_1 = \begin{bmatrix} 0.06 \\ 0.3 \\ 0.54 \\ 0.78 \end{bmatrix}$ $\quad V_2 = \begin{bmatrix} 0.83 \\ 0.46 \\ 0.08 \\ -0.3 \end{bmatrix}$

$V_3 = \begin{bmatrix} 0.75 \\ 0.38 \\ 0.38 \\ 0.38 \end{bmatrix}$ $\quad V_4 = \begin{bmatrix} -0.75 \\ -0.38 \\ -0.38 \\ -0.38 \end{bmatrix}$

(Q5) $Xw = y$

* The goal of least squares is to minimize the difference between Xw & y. (assuming this diff. is e)

$$e = Xw - y$$
$$\|e\|^2 = (Xw-y)^T(Xw-y)$$
$$= w^TX^TXw - w^TX^Ty - y^TXw + y^Ty$$

* Differentiating w.r.t w and setting to 0

$$\frac{\partial}{\partial w}\left(w^Tx^TXw - w^TX^Ty - y^TXw + y^Ty\right) = 0$$

$$\frac{\partial}{\partial w}\|e\|^2 = 2X^TXw - 2X^Ty = 0$$

$$X^TXw = X^Ty$$

$$w = (X^TX)^{-1}X^Ty$$ //

(Q6)* The company is essentially using accuracy as their evaluation metric.

P.T.O

* However, this is certainly not a good metric.

* Let us assume that the test is faulty and only gives a COVID-negative result.

* If, out of the 100 people they randomly tested, 99 turned out to be healthy, their test would score a 99% accuracy even though their test is faulty.

* Hence, I suggest the company ~~using~~ use metrics such as precision and recall to evaluate their COVID test.

* I would also suggest the company use a larger set of people which includes both people with and without COVID.