Name: Asadullah Hill Galib

CSE 847

Homework 1

(1) Introduction

(1) Given, $P(r) = 0.2$, $P(b) = 0.2$, $P(g) = 0.6$

1.3 $p(apple) = ?$

$p(g \mid orange) = ?$

$$P(apple) = P(apple \mid b) P(b) + P(apple \mid r) P(r) + P(apple \mid g) P(g)$$

$$\cancel{= \frac{5}{10} \cdot 0.2 +}$$

$$= \frac{5}{10} \cdot 0.2 + \frac{3}{10} \cdot 0.2 + \frac{3}{10} \cdot 0.6$$

$$= 0.34$$

$\cancel{P(g \mid orange) = \frac{P(o \mid g) P}{}}$

$$P(g \mid orange) = \frac{P(orange \mid g) P(g)}{P(orange)}$$

$$= \frac{P(orange \mid g) P(g)}{\sum_{box} P(orange \mid box) P(box)}$$ → using sum and product rule

$$= \frac{\frac{3}{10} \times 0.6}{\underbrace{\frac{4}{10} \times 0.2}_{\text{red box}} + \underbrace{\frac{5}{10} \times 0.2}_{\text{blue box}} + \underbrace{\frac{3}{10} \times 0.6}_{\text{green box}}}$$

$$= 0.5 \quad \text{(Ans).}$$

(2) Given, two variables x and y are independent.

1.6

$$cov[x,y] = E[xy] - E[x]\,E[y]$$

$$= \sum_{xy} xy\, P(xy) - \sum_{x} x\, p(x) \cdot \sum_{y} y\, P(y)$$

$$\cancel{= \sum_{xy} xy\, P(xy) - \sum_{x} x\, P(x}$$

$$= \sum_{xy} xy\, P(x).P(y) - \sum_{xy} xy\, P(x)\, P(y)$$

$[P(xy) = P(x).P(y)$ as x & y are independent$]$

$$= 0 \quad \text{[Showed]}$$

~~But~~ for continuous variables,

$$cov[x,y] = E[xy] - E[x]\,E[y]$$

$$= \iint xy\, p(xy)\, dx\, dy - E[x]\,E[y]$$

$$= \int x\, p(x)\, dx \int y\, p(y)\, dy - E[x]\,E[y]$$

$$= E[x]\,E[y] - E[x]\,E[y]$$

$$= 0$$

[Showed]

(3) 1.11

$$\ln p(x \mid \mu, \sigma^2) = -\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_n-\mu)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi) \quad \dots (1.54)$$

$$\mu_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_n \quad \dots (1.55)$$

$$\sigma^2_{ML} = \frac{1}{N}\sum_{i=1}^{N}(x_n-\mu_{ML})^2 \quad \dots (1.56)$$

By setting the Derivative of (1.54) with respect to μ equal to zero,

$$\frac{d}{d\mu}\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_n-\mu)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi)\right) = 0$$

$$\Rightarrow \frac{d}{d\mu}\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_n-\mu)^2\right) - 0 - 0 = 0$$

$$\cancel{\Rightarrow \sum_{i=1}^{N}\frac{d}{d\mu}\left(-\frac{1}{2\sigma^2}(x\right.}$$

$$\Rightarrow \sum_{i=1}^{N}\frac{d}{d\mu}\left(-\frac{1}{2\sigma^2}(x_n-\mu)^2\right) = 0$$

$$\Rightarrow \cancel{\sum} - \frac{1}{2\sigma^2}\sum_{i=1}^{N} 2(x_n-\mu).(-1) = 0$$

$$\Rightarrow \sum_{i=1}^{N}(x_n-\mu) = 0$$

$$\Rightarrow \sum_{i=1}^{N} x_n - N\mu = 0$$

$$\Rightarrow \mu_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_n \quad \text{[Showed]}$$

By setting the derivative of (1·54) with respect to σ^2 equal to zero,

$$\frac{d}{d\sigma^2}\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_n-\mu)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi)\right) = 0$$

$$\Rightarrow \cancel{\frac{d}{d\sigma^2}} \sum_{i=1}^{N}\left((x_n-\mu)^2 \cdot \frac{d}{d\sigma^2}\left(-\frac{1}{2\sigma^2}\right)\right) - \frac{N}{2}\cdot\frac{1}{\sigma^2} - 0 = 0$$

$$\Rightarrow \sum_{i=1}^{N}\left((x_n-\mu)^2 \cdot \left(-\frac{1}{2}\right)(-1)\left(\sigma^{-2}\right)^{-2}\right) - \frac{N}{2\sigma^2} = 0$$

$$\Rightarrow \sum_{i=1}^{N}\left((x_n-\mu)^2 \cdot \frac{1}{2\sigma^4}\right) - \frac{N}{2\sigma^2} = 0$$

$$\Rightarrow \frac{1}{2\sigma^2}\left(\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_n-\mu)^2 - N\right) = 0$$

$$\Rightarrow \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_n-\mu)^2 = N$$

$$\Rightarrow \sigma^2_{ML} = \frac{1}{N}\sum_{i=1}^{N}(x_n-\mu)^2$$

[Showed]

(2) Linear Algebra

(1) $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$, $C = \begin{bmatrix} 3 & -1 & 3 \\ 4 & 1 & 5 \\ 2 & 1 & 3 \end{bmatrix}$, $D = \begin{bmatrix} 2 & -4 & 5 \\ 0 & 1 & 4 \\ 3 & 2 & 1 \end{bmatrix}$,

$E = \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix}$

(a) $(2A)^T$

$$= \begin{bmatrix} 2 & 4 & 6 \\ 4 & 2 & 8 \end{bmatrix}^T = \begin{bmatrix} 2 & 4 \\ 4 & 2 \\ 6 & 8 \end{bmatrix} \text{ (Ans).}$$

(b) $(A - B)^T$

$A - 2 \times 3$

$B - 3 \times 2$

So, $A - B$ is undefined [To subtract, Have to be same dimensions]

(Not Possible) (Ans).

(c) $(3B^T - 2A)^T$

$$= \left(3 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} - 2\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}\right)^T$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ -4 & 1 & -2 \end{pmatrix}^T = \begin{bmatrix} 1 & -4 \\ 2 & 1 \\ 3 & -2 \end{bmatrix} \text{ (Ans).}$$

(1) (d) $(-A)^T E$

$$= \begin{bmatrix} -1 & -2 & -3 \\ -2 & -1 & -4 \end{bmatrix}^T \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -2 \\ -2 & -1 \\ -3 & -4 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -7 & -6 \\ -8 & 0 \\ -17 & -10 \end{bmatrix} \text{(Ans).}$$

(e) $(C + D^T + E)^T$

$$= \left(\underset{\substack{C \\ 3\times3}}{\begin{bmatrix} 3 & -1 & 3 \\ 4 & 1 & 5 \\ 2 & 1 & 3 \end{bmatrix}} + \underset{\substack{D^T \\ 3\times3}}{\begin{bmatrix} 2 & 0 & 3 \\ -4 & 1 & 2 \\ 5 & 4 & 1 \end{bmatrix}} + \underset{\substack{E \\ 2\times2}}{\begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix}} \right)^T$$

Dimensions mismatched

Addition with E is not possible.

Not Possible (Ans).

(3) $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ $\quad B = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}$

$$AB = \begin{bmatrix} 2-6 & -1+8 \\ 6-6 & -3+8 \end{bmatrix} = \begin{bmatrix} -4 & 7 \\ 0 & 5 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 2-3 & 4-2 \\ -3+12 & -6+8 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 9 & 2 \end{bmatrix}$$

So, $AB \neq BA$

② ⓐ $S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2=0\}$

(i) $\begin{bmatrix}0\\0\end{bmatrix} \in S$ as $0^2+0^2=0$

(ii) Let, $v = \begin{bmatrix}a\\b\end{bmatrix} \in S \Rightarrow a^2+b^2=0$

then $cv = c\begin{bmatrix}a\\b\end{bmatrix} = c\,(a^2+b^2) = 0$

So, $cv \in S \quad \forall c \in \mathbb{R}$ & S is closed under multiplication

(iii) Let, $v_1 = \begin{bmatrix}a\\b\end{bmatrix} \quad v_2 = \begin{bmatrix}c\\d\end{bmatrix}$

$v_1, v_2 \in S \Rightarrow a^2+b^2=0$ & $c^2+d^2=0$

then, $v_1+v_2 = \begin{bmatrix}\cancel{a+b}\\\cancel{c+d}\end{bmatrix} \; a^2+b^2+c^2+d^2=0$

So, $v_1+v_2 \in S$ & S is closed under addition

$\therefore$ So, S is a subspace of $\mathbb{R}^2$

Alternative Justification:

$\begin{bmatrix}0\\0\end{bmatrix} \in S$ [and it is the only solution for $x^2+y^2=0$ in $\mathbb{R}^2$]

For any scalars α & β,

$\alpha\begin{bmatrix}0\\0\end{bmatrix} + \beta\begin{bmatrix}0\\0\end{bmatrix} \in \cancel{\mathbb{R}^2}\, S$

So, S is a subspace.

② ⓑ $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 0\}$

let, $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in S$ & $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in S$

then,

$$\alpha v_1 + \beta v_2$$

$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \left[\text{let } \alpha = \beta = 1\right]$$

$$= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \notin S \quad \left[\text{as, } 1^2 - 0^2 \neq 0\right]$$

So, S is not a subspace.

② ⓒ $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y \cancel{b} = 0\}$

let, $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in S$ & $v_2 = \begin{bmatrix} -1 \\ \cancel{0}\,1 \end{bmatrix} \in S$

then,

$$\alpha v_1 + \cancel{\alpha} \beta v_2$$

$$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \left[\text{let, } \alpha = \beta = 1\right]$$

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix} \notin S \quad \left[\text{as } 0^2 - 2 \neq 0\right]$$

So, S is not a subspace

2 d $S = \{(x,y) \in \mathbb{R}^2 \mid x - y = 0\}$

~~So~~ $x - y = 0 \Rightarrow x = y$

So, ~~for~~ any vector in form $\begin{bmatrix} c \\ c \end{bmatrix}$, $c \in \mathbb{R}$ belongs

to S.

let, $v_1 = \begin{bmatrix} a \\ a \end{bmatrix}$ & $v_2 = \begin{bmatrix} b \\ b \end{bmatrix}$; $v_1, v_2 \in S$

then,

$$\alpha V_1 + \cancel{\alpha} \beta V_2 = \begin{bmatrix} \alpha a + \beta b \\ \alpha a + \beta b \end{bmatrix} \in S \quad [\text{as } x = y]$$

So, S is a subspace.

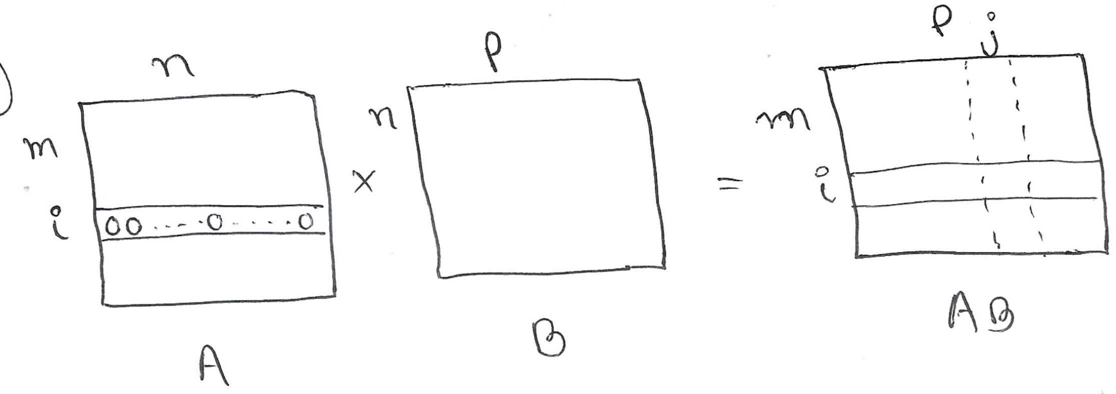
2 e $S = \{(x,y) \in \mathbb{R}^2 \mid x - y = 1\}$

$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \notin S$ as $0 - 0 \neq 1$

~~So~~, zero vector not in S

So, S is not a subspace.

4 a



Let, the ith row of A contains only zeros.

According to matrix multiplication, any element (i,j) in the ith row of AB,

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

But, for all k $(k = 1, 2, \dots n)$, $A_{ik} = 0$

So, for ~~any~~ all j, $(AB)_{ij} = 0$

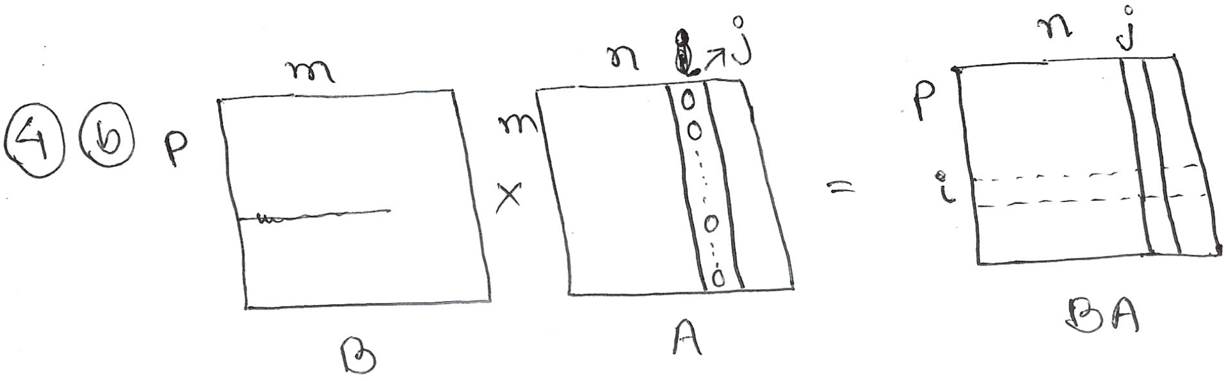
So, AB has a row of zeros [Showed]

Alternative proof:

$$AB = \underbrace{\begin{bmatrix} - a_1^T - \\ - a_2^T - \\ \vdots \\ - a_i^T - \\ \vdots \\ - a_m^T - \end{bmatrix}}_{A} B = \begin{bmatrix} - a_1^T B - \\ - a_2^T B - \\ \vdots \\ - a_i^T B - \\ \vdots \\ - a_m^T B - \end{bmatrix}$$

If a row of A, let $a_i^T = 0$'s then the ith row of AB, $a_i^T B$ also be all 0's [Showed]

4) b)



Let, j-th column of A contains all zeros.

Then according to matrix multiplication, for any element (i,j) in the j-th column of BA,

$$(BA)_{ij} = \sum_{k=1}^{m} B_{ik} A_{kj}$$

But, for all k $(k=1,2,\dots m)$, $A_{kj} = 0$

So, for all i, $(BA)_{ij} = 0$

So, BA has a column of all zeros. [Showed]

<u>alternative proof :</u>

$$BA = B \cdot \underbrace{\begin{bmatrix} | & | & & | & & | \\ a_1 & a_2 & \cdots & a_i & \cdots & a_n \\ | & | & & | & & | \end{bmatrix}}_{A} = \begin{bmatrix} | & | & & | & & | \\ Ba_1 & Ba_2 & \cdots & Ba_i & \cdots & Ba_n \\ | & | & & | & & | \end{bmatrix}$$

If a column of A, let a_i contains all zeros.

Then the i-th column of BA, Ba_i also contains all zeros [Showed]

(5) $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{p \times n}$

$$AB = A \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ Ab_1 & Ab_2 & \cdots & Ab_n \\ | & | & & | \end{bmatrix}$$

~~We~~ So, we can represent each column of AB as a combination of the columns of A.

Therefore, $\operatorname{ran}(AB) \subseteq \operatorname{ran}(A)$

$$\Rightarrow \dim(\operatorname{ran}(AB)) \leq \dim(\operatorname{ran}(A))$$

$$\Rightarrow \operatorname{rank}(AB) \leq \operatorname{rank}(A) \quad \cdots (i)$$

Again,

$$\operatorname{rank}(AB) = \operatorname{rank}(AB)^T$$

$$= \operatorname{rank}(B^T A^T)$$

$$\leq \operatorname{rank}(B^T) \quad \text{[using equation (i)]}$$

$$= \operatorname{rank}(B)$$

So, $\operatorname{rank}(AB) \leq \operatorname{rank}(B) \quad \cdots (ii)$

Equation (i) and (ii) implies that,

$$\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$$

[Showed]