CSE 847 Homework One

Reuben Lewis

05 February, 2021

1 Introduction

1.1 Pg 58, 1.3

Box	Apples	Oranges	Limes	Selection Prob p
tab$r$	3	4	3	0.2
b	1	1	0	0.2
g	3	3	4	0.6

(1) What is the probability of selecting is an apple? $p(A) = p(A, r) + p(A, b) + p(A, g)$
$p(A) = p(A|r) * p(r) + p(A|b) * p(b) + p(A|g) * p(g)$
$p(A) = 0.3 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6$
$p(A) = 0.34$

(2) If the selected fruit is an orange, what is the probability it came from the green box?
$p(O) = p(O, r) + p(O, b) + p(O, g)$
$p(O) = p(O|r) * p(r) + p(O|b) * p(b) + p(O|g) * p(g)$
$p(O) = 0.4 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6$
$p(O) = 0.36$
$p(g|O) = \frac{p(O|g)p(g)}{p(O)}$
$p(g|O) = \frac{0.3*0.6}{0.34}$
$p(g|O) = 0.5$

1.2 Pg 59, 1.6

$cov[x, y] = 0$
$\mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y] = 0$
$\mathbb{E}[xy] = \mathbb{E}[x]\mathbb{E}[y]$
$\sum_{x,y} xyP_{x,y}(x, y) = \sum_x xP(x) * \sum_y yP(y)$
$P(x, y) = P(x) * P(y)$
This is the definition of independence, so, if x and y are independent, their covariance is going to be 0.

1.3 Pg 59, 1.11

Solve first for μ_{ML}, then for σ^2 to verify the results of (1.55) and (1.56). This involves first taking the partial derivative with respect to the value we are solving for, then setting that equation to 0 and solving for the value.

(1) μ_{ML}

$$\begin{gathered}
-\frac{1}{2\sigma^2}\sum_{n-1}^{N}(x_n-\mu)^2-\frac{N}{2}\ln(\sigma^2)-\frac{N}{2}\ln(2\pi)\\
\frac{\partial}{\partial\mu}\Rightarrow-\frac{1}{2\sigma^2}\sum_{n-1}^{N}\frac{\partial}{\partial\mu}(x_n-\mu)^2\\
-\frac{1}{2\sigma^2}\sum_{n-1}^{N}2*-1*(x_n-\mu)\\
\frac{1}{\sigma^2}\sum_{n-1}^{N}(x_n-\mu)\\
0=\frac{1}{\sigma^2}\sum_{n-1}^{N}(x_n-\mu)\\
0=\sum_{n-1}^{N}(x_n-\mu)\\
0=\sum_{n-1}^{N}x_n-N*\mu\\
N*\mu=\sum_{n-1}^{N}x_n\\
\mu_{ML}=\frac{1}{N}\sum_{n-1}^{N}x_n
\end{gathered}$$

(2) σ^2

$$\begin{gathered}
-\frac{1}{2\sigma^2}\sum_{n-1}^{N}(x_n-\mu)^2-\frac{N}{2}\ln(\sigma^2)-\frac{N}{2}\ln(2\pi)\\
\frac{\partial}{\partial\sigma^2}\Rightarrow-\frac{\partial}{\partial\sigma^2}\frac{1}{2\sigma^2}\sum_{n-1}^{N}(x_n-\mu)^2-\frac{\partial}{\partial\sigma^2}\frac{N}{2}\ln(\sigma^2)\\
\frac{1}{\sigma^4}\sum_{n-1}^{N}(x_n-\mu)^2-\frac{N}{2\sigma^2}\\
0=\frac{1}{2\sigma^4}\sum_{n-1}^{N}(x_n-\mu)^2-\frac{N}{2\sigma^2}\\
\frac{N}{2\sigma^2}=\frac{1}{2\sigma^4}\sum_{n-1}^{N}(x_n-\mu)^2\\
N*\sigma^2=\sum_{n-1}^{N}(x_n-\mu)^2\\
\sigma^2=\frac{1}{N}\sum_{n-1}^{N}(x_n-\mu)^2
\end{gathered}$$

2 Linear Algebra

1. Matrix Calculations

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

   (a) $(2A)^T$ First, scale A by 2, then transpose it.

$$\begin{gathered}
2*A=\begin{bmatrix}2&4&6\\4&2&8\end{bmatrix}\\
(2*A)^T=\begin{bmatrix}2&4\\4&2\\6&8\end{bmatrix}
\end{gathered}$$

(b) $(A - B)^T$ Because A and B don't have the same dimension, this equation doesn't work. $A - B^T$ would, however, since B would then have the same dimensions as A.

(c) $(3B^T - 2A)^T$ First, scale the matrices, then apply the transpose onto B. Next, subtract the matrices and transpose that output.

$$2 * A = \begin{bmatrix} 2 & 4 & 6 \\ 4 & 2 & 8 \end{bmatrix}$$
$$3 * B = \begin{bmatrix} 3 & 0 \\ 6 & 3 \\ 9 & 6 \end{bmatrix}$$
$$3 * B^T = \begin{bmatrix} 3 & 6 & 9 \\ 0 & 3 & 6 \end{bmatrix}$$
$$3 * B^T - 2 * A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & 1 & -2 \end{bmatrix}$$
$$(3 * B^T - 2 * A)^T = \begin{bmatrix} 1 & -4 \\ 2 & 1 \\ 3 & -2 \end{bmatrix}$$

(d) $(-A)^T E$ First, distribute the $-1 * A$, and then transpose it. Next, multiply it by E, since there inner dimensions match (3x2 * 2x2).

$$-1 * A = \begin{bmatrix} -1 & -2 & -3 \\ -2 & -1 & -4 \end{bmatrix} \; (-1 * A)^T = \begin{bmatrix} -1 & -2 \\ -2 & -1 \\ -3 & -4 \end{bmatrix}$$
$$(-1 * A)^T E = \begin{bmatrix} -1 * 3 + -2 * 2 & -2 * -1 + -2 * 4 \\ -2 * 3 + -1 * 2 & -2 * 2 + -1 * 4 \\ -3 * 3 + -4 * 2 & -3 * -2 + -4 * 4 \end{bmatrix}$$
$$(-1 * A)^T E = \begin{bmatrix} -7 & -6 \\ -8 & 0 \\ -17 & -10 \end{bmatrix}$$

(e) $(C + D^T + E)^T$
Because E is not the same dimensions as C or D, it can not be added to the others, so this operation cannot be done.

2. Subspace of $\mathbb{R}^2$? Justify your answer.

(a) $\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 0\}$
This is not a subspace, since two vectors $(1, 0) + (1, 1) = (2, 1)$ and show's it's not closed under addition.

(b) $\{(x, y) \in \mathbb{R}^2 | x^2 - y^2 = 0\}$
This is not a subspace, since two vectors $(1, 1) + (1, -1) = (2, 0)$ and show's it's not closed under addition.

(c) $\{(x, y) \in \mathbb{R}^2 | x^2 - y = 0\}$
This is not a subspace, since two vectors $(1, 1) + (1, -1) = (2, 0)$ and show's it's not closed under addition.

(d) $\{(x,y) \in \mathbb{R}^2 | x - y = 0\}$
This is a subspace, since it satisfies the three main properties. The 0 vector is a part of the space, it's closed under vector addition, and it's closed under scalar multiplication. This is a special case of a line with a slope of 1.

Closed under Vector addition:
$$u = (x_1, y_1), v = (x_2, y_2)$$
$$u + v = (x_1 + x_2, y_1 + y_2)$$
$$u + v = (x_1 + x_2, x_1 + x_2)$$

This process is similar for scaling as well.

(e) $\{(x,y) \in \mathbb{R}^2 | x - y = 1\}$
This is not a subspace because it doesn't contain the origin.

3.
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}.$$

Is $AB = BA$? No.

$$AB = \begin{bmatrix} 1*2+2*-3 & 1*-1+2*4 \\ 3*2+2*-2 & 3*-1+2*4 \end{bmatrix}$$
$$AB = \begin{bmatrix} -4 & 7 \\ 2 & 5 \end{bmatrix}$$
$$BA = \begin{bmatrix} 2*1+-1*3 & 2*2+-1*2 \\ -3*1+4*3 & -3*2+4*2 \end{bmatrix}$$
$$BA = \begin{bmatrix} -1 & 2 \\ 9 & 2 \end{bmatrix}$$

4. (a) Because of how matrix multiplication works, you multiply the rows of the first matrix by the columns of the second. This means that if there is a row of 0s in the first matrix, that row would be multiplied against every column of the second matrix. This would cause a row of 0s in the resulting matrix, since multiplying any value by 0 results in 0. As an example, imagine $A = 2X3$ matrix and $B = 3X4$, with a row of 0s at $i = 2$. Row 2 of the resulting matrix would be calculated like so:

$$A_{2k} = \sum_j 0_{2,j} * b_{j,k}$$
$$\begin{bmatrix} 0*b_{0,0} + 0*b_{1,0} + 0*b_{2,0} & 0*b_{0,1} + 0*b_{1,1} + 0*b_{2,1} \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 \end{bmatrix}$$

(b) The best way to show this is through an example. Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

If we multiply BA, then we get the following matrix. This is because when we multiply rows of B through the columns of A, the column of 0s in A nullifies the values in the row of B, resulting in a column of 0s in the resulting matrix.

$$BA = \begin{bmatrix} 1*1+2*1 & 1*0+2*0 \\ 3*1+4*1 & 3*0+4*0 \end{bmatrix}$$
$$BA = \begin{bmatrix} 3 & 0 \\ 7 & 0 \end{bmatrix}$$

5. Show that $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$.

Remember that $rank(AB) = dim(ran(AB)), rank(A) = dim(ran(A))$ By definition, $ran(AB) = y \in \mathbb{R}^n : y = ABxforsomex \in \mathbb{R}^n$ Substitute $z = Bx \Rightarrow y = A(Bx) \Rightarrow y \in \mathbb{R}^n : y = Azforsomex \in \mathbb{R}^n = ran(a)$

This shows that the vector space $ran(AB)$ is a subspace of $ran(A)$, which allows us to say $rank(AB) = dim(ran(AB)) \leq dim(ran(A)) = rank(A)$. Now what about $rank(B)$?

Remember that $rank(A) = rank(A^T)$ Let $C = A^T$ and $D = B^T$ $rank(DC) \leq rank(D) \Rightarrow rank(C^T D^T) \leq rank(D^T)$ Translating that back over, we get $rank(AB) \leq rank(B)$

So, by that logic, $rank(AB)$ is less than or equal to whichever of A or B has the lower rank.