CSE4088
Introduction to Machine Learning
Homework 1
Prerequisite Concepts

Prepared by:

İsmail ÖKSÜZ
150119516

BASIC QUESTIONS

1-Vectors and Matrices

1- I calculated the inner product of y and z by using the formula.

$$\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a} = a_1b_1 + a_2b_2 + \cdots + a_mb_m = \sum_{i=1}^{m} a_ib_i.$$

9*7+8*6=111

2-I calculated the product Xy by using the formula.

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{pj}$$

X is 2x2 matrix and y is a vector so the result is not a scalar number.
9*9+8*8=145
7*9+6*8=111

The result is:
145
111

3- I added augmented matrix on the right side. Then made the first column pivot. After that I eliminated the first row. Then I made the second column pivot. And I eliminated the second row. The left part gives me the identity matrix and the right part gives me the inverse matrix.

The result becomes:
-3 4
7/2 -9/2

4-X is an invertible matrix, so the rank of X = number of rows.
Consequently, the answer is 2.

2-Calculus

1- When take the derivative with respect to x, it becomes
3*4*x^(3-1)-2*x^(2- = 12x^2-2x

2- Firstly, I divided y into 2 partial derivative.

For the left part, I take out the values which are not respected to x from the derivative. Then I did my operation.
The result of the left part becomes: tanz*6*z*x^(6*z-1)

After that, I did the right part. For partial derivative of ln, the inside value of the ln takes out with flipping. After that, the inside value was decimal, I used the Quotient Rule. Then I multiplied the value by the first value that I taked out. The result becomes: (-21x-4z)/(7x^2+xz)

When I substract the right part from the left part, I got the result as:
[tanz*6z*x^(6z-1)]-[(-21x-4z)/(7x^2+xz)]

3-Probability and Statistics

1- Number of numbers in the data is 7.
The sum of those numbers are 4.
So, the sample mean= 4/7

2-To calculate sample variance, I used the formula.

$$s^2 = \frac{\sum (X - \bar{X})^2}{N - 1}$$

After doing calculations, the result becomes 2/7

3- If the number in data is 0, I multiplied by 0.3
Else if the number is 1, I multiplied by 0.7
Number of 0=3 and Number of 1=4
So (0.3)^3+(0.7)^4 give me the result as 0.0065

4-The mean value maximizes the probability of the sample.
So the result is same as the mean value which is: 4/7

5-
a- P(A=0,B=0) was given in the table. 0.1

b- To find P(A=1), we sum every column in the probability table where A=1
P(A=1,B=0)+P(A=1,B=1)=0.2+0.3=0.5

c- P(A=0|B=1) = P(A=0,B=1)/P(B=1) = [P(A=0,B=1)]/[P(B=1,A=0)+P(B=1,A=1)]
=(0.4)/(0.4+0.3)=4/7

d-P(A=0vB=0) = P(A=0) + P(B=0) – P(A=0,B=0)
=P(A=0,B=0) + P(A=0,B=1) + P(B=0,A=0) + P(B=0,A=1) – P(A=0,B=0)
=P(A=0,B=1) + P(B=0,A=0) + P(B=0,A=1)
= 0.4+0.1+0.2 = 0.7

4- Big-O Notation

1- O(f(n))=O(n/2)=O(n)
O(g(n))=O(log2n)=O(logn)
They are not equal to each other.
g(n)=O(f(n)) is true.

2- O(f(n))=O(lnn)=O(logn)
O(g(n))=O(log2n)=O(logn)
They are equal to each other.
g (n)=O(f(n)) and f(n)=O(g(n)) are both true.

3- O(f(n))=O(n^100)
O(g(n))=O(100^n)
They are not equal to each other.
f(n)=O(g(n)) is true.

MEDIUM-LEVEL QUESTIONS

5-Algorithm

-I created a function findZero
-Lower is a variable, initially “0”
-Upper is a variable, initially “len(arr)-1”
-Arr is array to search
-Mid is the average of lower and upper

-The function continues only if lower is less than upper.
-If arr[mid]==0, function returns mid.
-If arr[mid]>0, function calls findZero but upper value becomes mid-1.
findZero(arr,lower,mid-1)
-If arr[mid]<0, function calls findZero but lower value becomes mid+1.
findZero(arr,mid+1,upper)
-If there is no zero in range, returns 0.

```
findZero(arr,lower,upper)
        while(lower<upper)
                mid=(lower+upper)/2
                if(arr[mid]==0)
                        return mid
                else if(arr[mid]<0)
                        return findZero(arr,mid+1,upper)
                else
                        return findZero(arr,lower,mid-1)
        return 0
```

6-Probability and Random Variables

6.1- Probability

1- By using Bayer Rule, P(A|B) becomes [P(B|A)*P(A)]/P(B)
So P(A|B)*P(B) = P(B|A)*P(A)
The Answer is True.

2- P(AuB) = P(A)+P(B)-P(A,B)
P(A|B) = P(A,B)/P(B)
If P(B) = 1, the equation becomes true. But B cannot be 1 .
Therefore, the answer is False.

3- Let BuC=X
After doing some operations, I got the equation. Not all values satisfies the condition but some of them satisfies.
Therefore the answer is True.

4- P(A^C|B) + P(A|B) = 1
But P(A|B) not equal to P(B|A).
So, the answer is False.

5- The answer is True.

6.2- Discrete and Continuous Distributions

Laplace is h

Multinomial is i

Poisson is l

Dirichlet is k

Gamma is j

6.3- Mean and Variance

1-

a- I find the mean of binomial distribution using this formula.

Mean

$$\mu_x = np$$

The result is n*p

b- I find the variance of binomial distribution by taking the square of standard deviation. Standard deviation can be found by using this formula:

$$\sigma_x = \sqrt{np(1-p)}$$

The square of standard deviation is: n*p*(1-p)

2- I know that E[X]=1 and Var(X)=1

a- E[3X] = 3*E[X] = 3*1 = 3
b- Var(3X) = (3^2)*Var(X) = 3*3*1 = 9
c- Var(X+3) = Var(X) + 0 = 1

6.4- Mutual and Conditional Independence

1- The integral can be divided into two parts because those two are independent. After the separation, I get two integral.
x*f(x)dx = E[X] and y*f(y)dy = E[Y]
The result is multiplication of those two integrals.
So E[X]*E[Y] = E[XY] is provided.

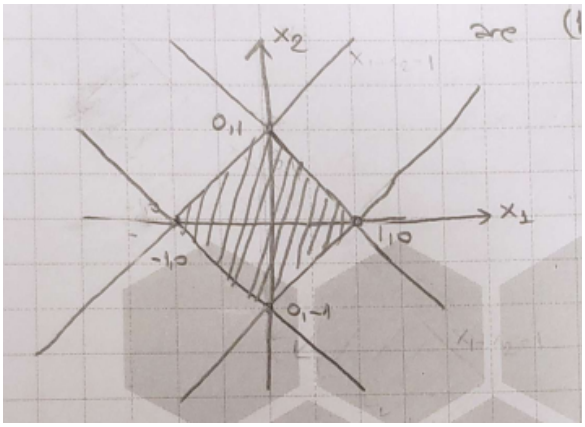
2-Covariance of X and Y becomes 0. Because for independent variables, the covariance is 0. We know that X and Y are independent so Cov(X,Y) = 0.
Var(X+Y) = Var(X) + 2*Cov(X,Y) + Var(Y) was given.
We know that Cov(X,Y) = 0
So Var (X+Y) = Var(X) + Var(Y) is provided.

3-As those two dices behave independently of each other, the result of the second die will be totally independent from the first die. Because they behave independently, they are not related to each other

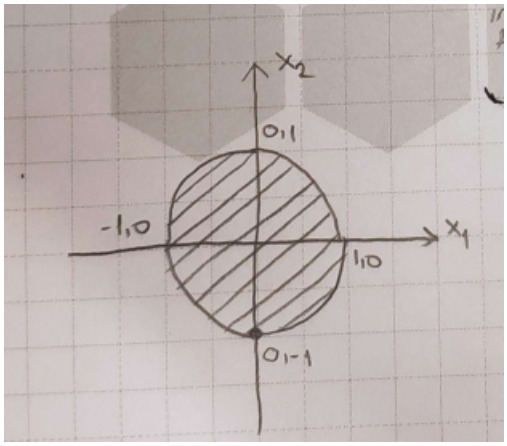
7-Linear Algebra

7.1- Norms

1- The equation is |x1|+|x2|<=1
Intersection points on top of the region are (1,0), (-1,0), (0,1) and (0,-1)



2- The equation is x1^2+x2^2<=1
Intersection points on top of the region are (1,0), (-1,0), (0,1) and (0,-1)



3-The equation is, highest of x1 and x2 is less than 1.
If x1>x2, we take x1 as max.
If x1<x2, we take x2 as max.
If x1=x2, we take the value of those two as max.