Summer 2020 CX4641/CS7641 Homework 1

Dr. Mahdi Roozbahani

Deadline: May 29, Friday, 11:59 pm

- No extension of the deadline is allowed. Late submission will lead to 0 credit.
- Discussion is encouraged on Piazza as part of the Q/A. However, all assignments should be done individually.

Instructions

- This assignment has no programming, only written questions.
- We will be using Gradescope this semester for submission and grading of assignments.
- Your write up must be submitted in PDF form, you may use either Latex or markdown, whichever you prefer. We will not accept handwritten work.
- Please make sure to start answering each question on a new page. It makes it more organized to map your answers on GradeScope. When submitting your assignment, you must correctly map pages of your PDF to each question/subquestion to reflect where they appear. Improperly mapped questions may not be graded correctly.

1 Linear Algebra [25pts + 8pts]

1.1 Determinant and Inverse of Matrix [11pts]

Given a matrix M:

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

(a) Calculate the determinant of M. [2pts] (Calculation process required.)

(b) Calculate M^{-1}. [5pts] (Calculation process required)
(**Hint:** Please double check your answer and make sure $MM^{-1} = I$)

(c) What is the relationship between the determinant of M and the determinant of M^{-1}? [2pts]

(d) When does a matrix not have an inverse? Provide an example. [2pts]

1.2 Characteristic Equation [8pts] (Bonus)

Consider the eigenvalue problem:

$$Ax = \lambda x, x \neq 0$$

where x is a non-zero eigenvector and λ is eigenvalue of A. Prove that the determinant $|A - \lambda I| = 0$.

1.3 Singular Value Decomposition [14pts]

Given a matrix A:

$$A = \begin{bmatrix} 3 & 3 & 0 \\ -2 & 2 & 0 \end{bmatrix}$$

Compute the Singular Value Decomposition (SVD) by following the steps below. Your full calculation process is required.

(a) Calculate all eigenvalues of AA^T and A^TA. The square roots of the positive eigenvalues make up the singular values, the diagonal entries in Σ. They will be arranged in descending order, all other values in Σ are 0. [4pts]

(b) Calculate all eigenvectors of AA^T normalized to unit length. These will make up the left singular vectors, or the columns of U. [4pts]

(c) Calculate all eigenvectors of A^TA normalized to unit length. These will make up the right singular vectors, or the rows of V^T. [4pts]

(d) Put it all together. Write out the SVD of matrix A in the following form:

$$A = U\Sigma V^T$$

[2pts]

Hint: Reconstruct matrix A from the SVD to check your answer.

2 Expectation, Co-variance and Independence [25pts]

Suppose X, Y and Z are three different random variables. Let X obeys Bernouli Distribution. The probability disbribution function is

$$p(x) = \begin{cases} 0.5 & x = c \\ 0.5 & x = -c. \end{cases}$$

c is a constant here. Let Y obeys the standard Normal (Gaussian) distribution, which can be written as $Y \sim N(0,1)$. X and Y are independent. Meanwhile, let $Z = XY$.

(a) What is the Expectation and Variance of X?(in terms of c) [4pts]

(b) Show that Z also follows a Normal (Gaussian) distribution. Calculate the Expectation and Variance of Z. [9pts]

(c) How should we choose c such that Y and Z are uncorrelated(which means $Cov(Y,Z) = 0$)? [5pts]

(d) Determine whether the following probability is greater than or equal to 0: (1) $P(Y = 0)$; (2) $P(Z = c)$; (3) $P(Y \in (-1,0))$; (4) $P(Z \in (2c,3c))$; (5) $P(Y \in (-1,0), Z \in (2c,3c))$; (6) $P(Y \in (-2,-1), Z \in (c,2c))$.[3pts]

(e) Are Y and Z independent? Make use of the above probabilities to show your conclusion.[4pts]

3 Maximum Likelihood [25 + 10 pts]

3.1 Discrete Example [10 pts]

Suppose we have two types of coins, A and B. The probability of a Type A coin showing heads is θ. The probability of a Type B coin showing heads is 2θ. Here, we have a bunch of coins of either type A or B. Each time we choose one coin and flip it. We do this experiment 10 times and the results are shown in the chart below.

Coin Type	Result
A	Tail
A	Head
A	Tail
B	Head
A	Tail
A	Tail
B	Head
B	Head
B	Head
A	Tail

(a) What is the likelihood of the result given θ? [4pts]

(b) What is the maximum likelihood estimation for θ? [6pts]

3.2 [10 pts]

The C.D.F of independent random variables $X_1, X_2, ..., X_n$ is

$$P(X_i \leq x|\alpha, \beta) = \begin{cases} 0, & x < 0 \\ (\frac{x}{\beta})^{\alpha}, & 0 \leq x \leq \beta \\ 1, & x > \beta \end{cases}$$

where $\alpha \geq 0$, $\beta \geq 0$. Find the MLEs of α and β.

3.3 Poisson distribution [5 pts]

The Poisson distribution is defined as

$$P(x_i = k) = \frac{\lambda^k e^{-\lambda}}{k!}(k = 0, 1, 2, ...).$$

What is the maximum likelihood estimator of λ?

3.4 Bonus [10 pts]

Given n i.i.d. observations $\{(x_i, y_i)\}_{i=1}^n \in \mathbb{R}^d \times \{-1, 1\}$, we assume

$$\mathbb{P}(y_i = 1|x_i) = h(x_i^T\theta) \text{ and } \mathbb{P}(y_i = -1|x_i) = 1 - h(x_i^T\theta)$$

where $h(x) = \frac{1}{1+\exp(-x)}$ and θ is the model parameter and $\theta = (\theta_1, \theta_2, \ldots, \theta_d)^T$.

Write out the likelihood function $L(\theta)$ given (x_i, y_i). Then formulate the log-likelihood function.

4 Information Theory [25pts + 7pts]

4.1 Marginal Distribution [6pts]

Suppose the joint probability distribution of two binary random variables X and Y are given as follows.

$X\|Y$	1	2
0	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{2}$	0

(a) Show the marginal distribution of X and Y, respectively. [3pts]

(b) Find mutual information for the joint probability distribution in the previous question [3pts]

4.2 Mutual Information and Entropy [19pts]

Given a dataset as below.

Player	*Experience*	*NumUtilities*	*BuysBoardwalk?*	*Hunger*	*Outcome*
1	*novice*	2	*no*	*low*	*lose*
2	*intermediate*	0	*no*	*high*	*lose*
3	*novice*	1	*no*	*low*	*win*
4	*expert*	0	*no*	*medium*	*win*
5	*intermediate*	0	*yes*	*high*	*win*
6	*expert*	0	*yes*	*high*	*lose*
7	*intermediate*	2	*yes*	*low*	*win*
8	*intermediate*	1	*no*	*medium*	*win*
9	*expert*	1	*no*	*low*	*lose*
10	*novice*	0	*no*	*medium*	*lose*
11	*novice*	2	*yes*	*low*	*win*
12	*intermediate*	1	*no*	*medium*	*lose*
13	*intermediate*	0	*yes*	*high*	*win*
14	*novice*	0	*yes*	*high*	*lose*

You are analyzing data from your last few Monopoly games in hopes of becoming a world champion. We want to determine what makes a player win or lose. Each input has four features (x_1, x_2, x_3, x_4): Experience, NumUtilities, BuysBoardwalk, Hunger. The outcome (win vs lose) is represented as Y.

(a) Find entropy $H(Y)$. [3pts]

(b) Find conditional entropy $H(Y|x_1)$, $H(Y|x_4)$, respectively. [8pts]

(c) Find mutual information $I(x_1, Y)$ and $I(x_4, Y)$ and determine which one (x_1 or x_4) is more informative. [4pts]

(d) Find joint entropy $H(Y, x_3)$. [4pts]

4.3 Bonus Question [7pts]

(a) Suppose X and Y are independent. Show that $H(X|Y) = H(X)$. [2pts]

(b) Suppose X and Y are independent. Show that $H(X, Y) = H(X) + H(Y)$. [2pts]

(c) Prove that the mutual information is symmetric, i.e., $I(X, Y) = I(Y, X)$ and $x_i \in X, y_i \in Y$ [3pts]

5 Bonus for All [10 pts]

Due to the recent social distancing requirement, Wal-Mart is re-evaluating their delivery policies. In order to properly update their policy, Wal-Mart is analyzing data from previous records. Delivery time can be classified as early, on time or late. Delivery distance can be classified as within 5 miles, between 5 and 10 miles and over 10 miles. From the previous records, 15% of deliveries arrive early, and 55% arrive on time. 70% of orders are within 5 miles and 25% of orders are between 5 and 10 miles. The probability for arriving on time if delivery distance is over 10 miles is 0. The probability of a shipment arriving on time and having a delivery distance between 5 and 10 miles is 10%. The probability for arriving early if delivery distance is within 5 miles is 20%.

(a) What is the probability that the delivery will arrive on time if the distance is between 5 and 10 miles? [2 pts]

(b) What is the probability that the delivery will arrive on time if the distance is within 5 miles? [4 pts]

(c) What is the probability that the delivery will arrive late if the distance is within 5 miles? [4 pts]