CS534 — Written Homework Assignment 0 —

Overview and Objectives. In this homework, you are going to practice some of the skills we'll be using in class. If you can solve most of these problems (even if you have to Google some identities or brush up on your knowledge) then you should be very well-equipped for the course. Otherwise, you'll need to spend extra time revising these topics. **This assignment is to help you (and us) gauge your familiarity with these concepts and might be a bit challenging depending on your background.**

How to Do This Assignment. We prefer solutions typeset in LaTeX but will accept scanned written work if it is legible. If a TA can't read your work, they can't give you credit. Submit your solutions to Canvas as a PDF.

Advice. Start early. Start early. Start early. You may be rusty on some of this material. Some of it might be new to you depending on your background – seek out resources to refresh yourself if so. Some helpful references:

- **Probability Refresher:** CS229 Probability Review from Stanford
- **Linear Algebra Refresher:** Zico Koltur's Linear Algebra Review
- **Differential Calculus Refresher:** Jackie Nicholas' booklet from University of Sydney
- **Integration Refresher:** Mary Barnes' booklet from University of Sydney

Linear algebra

(a) **Transpose and Associative Property, Positive Semi-definite matrices [2pt]** Define a matrix $B = bb^T$, where $b \in R^{d\times 1}$ is a column vector that is not all-zero. Show that B is a positive semi-definite matrix.

[*Hint*: To show that B is positive semi-definite, we need to show that B is symmetric, and for any vector $x \in R^{d\times 1}$, $x^T Bx \geq 0$. For the latter, try to get $x^T Bx$ to look like the product of two identical scalars. Note that $b^T x = (x^T b)^T$, that $a^T = a$ for scalar value a, and that matrix multiplication is associative.]

(b) **Solving systems of linear equations with matrix inverse. [2pt]** Consider the following set of linear equations:

$$\begin{cases} x_1 + x_2 - x_3 - x_4 = 1 \\ 2x_1 + 5x_2 - 7x_3 - 5x_4 = -2 \\ 2x_1 - x_2 + x_3 + 3x_4 = 4 \\ 5x_1 + 2x_2 - 4x_3 - 2x_4 = 6 \end{cases}$$

(a) (1 pt) Please express the system of equations as $A\mathbf{x} = \mathbf{b}$ by specifying the matrix A and vector $\mathbf{b}$

(b) (1 pt) Solve for $A\mathbf{x} = \mathbf{b}$ by using the matrix inverse of A (you can use software to compute the inverse).

Vector Calculus

(a) **Derivatives.[2pt]**. Compute the derivative $f'(x)$ for

(a) (1 pts) the logistic (aka sigmoid) function $f(x) = \frac{1}{1+\exp(-x)}$

(b) (1 pts) $f(x) = \exp(-\frac{1}{2\sigma^2}(x-\mu)^2)$

(b) **Gradients.** **[3pt]**Compute the gradient $\nabla_{\mathbf{x}} f$ of the following functions. Please clearly specify the dimension of the gradient.

(a) (1pt)

$$f(z) = \log(1+z),\ z = \mathbf{x}^T\mathbf{x},\ \mathbf{x} \in R^D$$

(b) (2pt)

$$f(z) = \exp\left(-\frac{1}{2}z\right)$$

$$z = g(\mathbf{y}) = \mathbf{y}^T S^{-1}\mathbf{y}$$

$$\mathbf{y} = h(\mathbf{x}) = \mathbf{x} - \mu$$

where $\mathbf{x}, \mu \in R^D, S \in R^{D\times D}$ is a symmetric matrix.

Probability

(a) **Joint, Marginal, and Conditional Probabilities [2pt]** Consider two discrete random variables X and Y with the following joint distribution:

Y \ X	x_1	x_2	x_3	x_4	x_5
y_1	0.01	0.02	0.03	0.1	0.1
y_2	0.05	0.1	0.05	0.07	0.2
y_3	0.1	0.05	0.03	0.05	0.04

Please compute:

(a) (1 pt) The marginal distributions $p(x)$ and $p(y)$

(b) (1 pt) The Conditional distribution $p(x|Y=y_1)$ and $p(y|X=x_3)$

(b) **Conditional probabilities, Marginalization and Bayes Rule [5pt]** Consider two coins, one is fair and the other one has a 1/10 probability for head. Now you randomly pick one of the coins, and toss it twice. Answer the following questions.

(a) (1pt) What is the probability that you picked the fair coin? What is the probability of the first toss being head?

(b) (2pts) If both tosses are heads, what is the probability that you have chosen the fair coin (Hint: you should apply Bayes Rule for this)?

(c) (2pts) If both tosses are heads, what is the probability that the third coin toss will be head? (you should build on results of c)

(c) **Linearity of Expectation [2 pt]** A random variable x distributed according to a standard normal distribution (mean zero and unit variance) has the following probability density function (pdf):

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Using the properties of expectations, evaluate the following integral

$$\int_{-\infty}^{\infty} p(x)(ax^2 + bx + c)dx$$

[*Hint:* This is *not* a calculus question. The simple solution relies on linearity of expectation and the provided mean/variance of p(x).]

(d) **Cumulative Density Functions / Calculus [2 pt]** X is a continuous random variable over the interval [0,1], show that the following function p is a valid probability density function (PDF) and derive the corresponding cumulative density function (CDF).

$$p(x) = \begin{cases} 4x & 0 \leq x \leq 1/2 \\ -4x + 4 & 1/2 \leq x \leq 1 \end{cases}$$

[*Hint:* Recall that a function is a valid PDF function if it integrates to 1: $\int_{-\infty}^{\infty} p(x)\ dx = 1$. And the cumulative density function (CDF) is defined as $C(x) = P(X \leq x)$ or the probability that a sample from p is less than x – which can be computed as $C(x) = \int_{-\infty}^{x} p(x)\ dx$. This *is* a calculus question. But the PDf is a piece-wise linear function, hence it is straightforward.]