This document contains a list of problems that are useful for making sure that you have the right background for CSC411/2515. The problems are designed to be straightforward to people with an in-depth understanding of the required math.

Each problem is built around one topic that is important in CSC411/2515. For example, when talking about neural networks and logistic regression, we will be discussion the Sigmoid function $\sigma(x) = \frac{1}{1+exp(-x)}$, and it will be important for students to have an intuitive grasp of its shape. Understanding the techniques that are used for solving the problems in Section 1 will make it easy for students to understand our discussion of the Sigmoid function in class.

If you are having trouble with a problem in this problem set, it is likely not enough to find out what the solution is. Instead, you should read up about the entire topic. For example, it is easy to look up the solution to Problem 3.4 in any linear algebra textbook, but what would really be useful when we are discussion Principal Component Analysis is (perhaps recent) experience with thinking about rotations of clouds of points in 2D and 3D space.

We are planning to use Eigenvectors and rotation matrices in the second half of the course. All the other problems will be relevant to material in the first several weeks.

1 Properties of Functions

1.1 Elementary functions

Let

$$f(x) = \exp(2x - 5),$$
$$g(x) = \log(1/x),$$
$$h(x) = x^2.$$

Sketch the functions f, h, and g. Label the axes, and compute the coordinates of at least three points on the sketch without using a calculator.

1.2 Argmin/Argmax

Consider the function

$$f(x) = \log\left(1 + \frac{1}{1 + |x - 10|}\right).$$

Where is this function's maximum? Where is this function's minimum? Do not use calculus.

2 Calculus

2.1 Definition of the derivative

Let

$$f(x) = x^2 + 1.$$

Using the definition of the derivative and limit notation, write down the formula for the derivative of f at point a. Simplify the expression as much as possible.

2.2 The chain rule and optimization

Let

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right).$$

At what value of x is the function $f(x)$ maximized? Show how to obtain the answer using calculus and how to check whether you obtained the minimum or the maximum. There is no restriction on the value of σ. Why is that important for correctly solving this problem?

3 Linear Algebra

3.1 Matrix multiplication

Re-express the following formula more concisely by using matrix multiplication notation

$$bm + dn + am + zn.$$

In the expression that you obtain, each entry in each matrix should be a single variable. An example of the kind of expression that is required is
$\begin{bmatrix} a & b \\ d & c \end{bmatrix} \begin{bmatrix} w & y \\ y & z \end{bmatrix}$.

3.2 Eigenvectors

Let $M = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$.

Which of the following are eigenvectors of the matrix M? Briefly explain how you got your answer. How can you solve the problem without explicitly computing the eigenvectors?

(a) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (b) $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ (b) $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

3.3 Rotation around the origin

Let a point on a 2D plane be $p = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

What is the matrix M such that Mp is the point p rotated clockwise around the origin by angle θ? What about counterclockwise?

3.4 Rotation around an arbitrary point

Let a point on a 2D plane be $p = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Write down an expression using matrices for rotation p around the point (a, b) clockwise, by angle θ.

3.5 Orthogonal matrices

Show that the rotation matrices you obtained in the previous questions (note that the final result there was not always just a rotation matrix) are orthogonal (i.e., $M^T M = MM^T = I$ and $M^T = M^-1$).

3.6 Basis and dimension

Give an example of a 2-dimensional linear subspace of a 3-dimensional space. Characterize the spaces both geometrically and by specifying their basis vectors.

3.7 Basis and dimension + rotation

Suppose all vectors in a 2D subspace of a 3D linear space are rotated using a rotation matrix M. Characterize the resultant subspace. Is it still 2D? What are its basis vectors?

Suppose all vectors in a 2D subspace of a 3D linear space are shifted by $(5, 2, 15)$. Under what conditions is this set of vectors still a linear space after the shift? What are the basis vectors for the 2D subspace?

3.8 Change of basis: Rotation

~~Suppose all vectors in a 2D subspace of a 3D linear space are rotated using a rotation matrix M. Under what conditions (on M) does the 2D subspace not change? Your answer should refer to the entries in the 3×3 matrix M.~~

The problem is recalled by the manufacturer for being too difficult/ambiguous. You are still encouraged to work on it, but it's not necessary.

4 Probability

4.1 Standard distributions: Poisson

Suppose you sample 10 independent variables from the Poisson distribution with rate λ. What is the probability that they are all larger than 0?

4.2 Standard distributions: Bernoulli

A fair coin is tossed three times. What is the probability that the outcome of all three tosses is the same?

4.3 Standard distributions: Bernoulli and Binomial

Suppose you sample 20 independent variables from Bernoulli distribution with parameter θ. What is the expected sum? What is the probability that the sum is greater than 15?

4.4 Standard distributions: Bernoulli and Binomial, independence

In the previous problem, we assumed that the variables are independent. Give an example where the variables are not independent (both mathematically, and by describing a plausible real-life scenario where that is the case). In the scenario you described, what is the expected sum of the 20 variables? What is the probability that the sum is greater than 15?

4.5 Independence

Give an example of a probability mass function $P(X, Y)$ such that the random variables X and Y are *not* statistically independent.

4.6 Conditional probability: basic idea

Two fair dice are tossed (independently). Given that the sum of of the outcomes is 6, what is the probability that the outcome of the first die is 3? Given that the sum of the outcomes is s, what is the probability that the outcome for the first die is a? What is the probability that the sum is s and the outcome for the first die is a? What is the probability that the sum is s and the outcome for the first die is a and the outcome for the second die is b?

4.7 Conditional probability: the Monty Hall Problem

At the end of the Let's Make a Deal with Monty Hall game show, a contestant needs to choose between three closed doors. There is a car behind one of the doors, and there are goats behind two of the other doors. The contestant chooses a door, and Monty Hall, who knows where the car is, opens one of the doors that the contestant didn't choose, and reveals that there is a goat behind it. The contestant is then offered to switch to another, still-closed, door. Should the contestant switch? Assume that Monty Hall will always open one of the doors to reveal a goat, regardless of which door the contestant chooses.

Suppose the contestant picked door #1, and Monty Hall revealed that there is a goat behind door# 2. Compute the probability $P(d1 = \text{car}, \text{Monty opens Door \# 2})$, and use that to compute $P(d1 = \text{car}|\text{Monty opens Door \# 2})$. Compute the probability $P(d3 = \text{car}, \text{Monty opens Door\# 2})$, and use that to compute $P(d3 = \text{car}|\text{Monty opens Door\# 2})$.

Why is it necessary to note that Monty Hall will *always* open one of the doors to reveal a goat?

4.8 The Law of Total Probability and Bayes' Theorem

The probability that that a truck hits a house at any given second is 1×10^{-10}. The probability of an earthquake happening at any given second is 1×10^{-11}. The probability of a house jumping if there is an earthquake is 0.7, and the probability of a house jumping if a truck hits it is 0.9.

What is the probability of a house jumping in any given second, assuming (implausibly) that trucks hitting the house and earthquakes or independent events?

Given that the house jumped, what is the probability that an earthquake happened?

Given that the house jumped, what is the probability that there was both an eathquake, and a truck hit the house?