Assignment 0: Foundations

CSE 151B: Deep Learning

Summer 2026

Instructions

All submissions should be made to PrairieLearn before the deadline specified on the syllabus and course webpage.

As stated in the syllabus, you may use any tools or resources you'd like, and you must complete a **Learning Process and Resource Reflection**. Further, we will ask you about portions of the assignment in an oral assessment as described in the syllabus.

For this first assignment, we will work through some portions in class and/or discussion.

For questions that we do not work on together in class and/or discussion, you are also welcome to collaborate. Working together on homework should follow the spirit of the **Gilligan's Island Rule** (Dymond, 1986): No notes can be made (or recording of any kind) during a conversation about the problems in this assignment, and you must watch one hour of Gilligan's Island or something equally insipid before writing anything down.

If you're working on problems and learning together in a way that doesn't follow the spirit of the above rule, meaning that you *do* write things down together (which can definitely be useful for helping each other learn), do it with problems other than those directly from this assignment. There are lots of deep learning problems online in the resources linked on the syllabus and course webpage, and you are always welcome to ask teaching staff for more practice problems if you need help finding some.

Part I

Algorithms & Architectures

For the mathematical computation questions, you are expected and encouraged to *work through the problem yourself on paper (e.g. in your roughbook)* as if they were open-ended/free response questions. Multiple choice and Parsons problems simply make it possible for you to receive real-time assessment and feedback for you to learn from.

Some resources for reviewing matrix calculus:

- Notes from course 18.063 in MIT's OpenCourseWare
- The Matrix Calculus You Need For Deep Learning
- Matrix Calculus for 10-301/601

You are also welcome and encouraged to share any other materials that you find helpful for this assignment.

Question 1

A logistic regression model predicts $\hat{y} = \sigma(\mathbf{w} \cdot \mathbf{x} + b)$ and assigns class 1 when $\hat{y} \geq 0.5$. A student argues that because σ is a nonlinear function, the resulting decision boundary between the two classes must be nonlinear. What is your assessment of this argument?

Question 2

For the next problem, you'll be computing the derivative of the sigmoid function and writing it in terms that *only use the function itself* (and constants). In many of the techniques you'll learn, a large part of the computations needed are for derivatives. Being able to write derivatives in terms of quantities that you've already computed is helpful for avoiding redundant computations.

In the example of the sigmoid function, imagine that you've already computed the sigmoid of certain values in order to find out what the activations are of different units. Let's say you then want to do an update to your model that involves taking the derivative of the sigmoid. Being able to represent the derivatives in terms of the activations you've already calculated means you don't have to redo those computations.

Derive $\frac{d}{dx}\sigma(x)$ (where $\sigma(x) = \frac{1}{1+e^{-x}}$), and factor the solution into terms that only use $\sigma(x)$ and constants.

Question 3

Let's do a warm up to remember multivariable derivatives before we get to vectors.

Find the first, second, and cross partial derivatives of $f(x, y) = x^2 + 5xy + 2y^2$.

Question 4

Find $\frac{\partial f}{\partial x_i}$ where $f(\mathbf{x}) = \|\mathbf{x}\|_2^2 \quad (1 \leq i \leq n)$. (The subscripted 2 indicates the L2 norm).

Question 5

Find $\frac{\partial f}{\partial x_i}$ where $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x}$ and $\mathbf{w}$ is a constant vector $(1 \leq i \leq n)$.

Part II

Programming

Getting started with Pytorch

Make a copy of the following Stanford CS224N PyTorch Tutorial in your Google Drive, and complete the sections from the beginning ("Part 1: Tensors") up to and not including "Autograd". You're welcome to work ahead if you'd like–the later portions of this tutorial will be useful for the second part of next week's assignment.

This week's oral assessment will be only from portions before the "Autograd" section.

Acknowledgments

Parts of this assignment's instructions were adapted from Professor Gary Cottrell's materials.

Parts of the matrix calculus portions of this assignment were adapted from CMU's Matrix Calculus for 10-301/601.

The programming portion of this assignment was adapted from the Stanford CS224N PyTorch Tutorial.