Homework 0: Math Refresher

Shichang Zhang

June 2024

This homework briefly goes over some of the mathematics that we will use in the course. If you have questions regarding them, you may reach out to the TAs during the lab session in the afternoon, or you can ask questions on Piazza. In all cases, you are only required to know basic concepts and manipulation rules. This homework is a self-practice. There is no due date, and it **won't** be counted towards your grades.

1 Elementary Notations

1.1 Number Systems

We will use the basic building blocks of mathematics:

Example 1 *The **Natural Numbers** are represented by the symbol* $\mathbb{N}$*, which are numbers* $0, 1, 2, 3, \cdots$.

Example 2 *The **Integers** are represented by the symbol* $\mathbb{Z}$*, which are numbers* $\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots$.

Example 3 *The **Reals** are represented by the symbol* $\mathbb{R}$*, which are numbers that have decimal expansions.*

We will not concern ourselves with the details, only the notations. By default we assume all numbers are Reals in this course.

1.2 Function

In our context, a function is a map (that is, some arrow such that the source is sent to the target) from a number system to another number system. For example, the notation:

$$f : \mathbb{R} \to \mathbb{R}$$

is a function from the reals to the reals, an example is the function $f(x) = 2x$, which returns $2x$ for every real number x.

$$f : \mathbb{R} \to \{0, 1\}$$

is a function that takes any real number to either 0 or 1. Here the bracket notation $\{0, 1\}$ is a **set** of two numbers, 0 and 1.

1.3 Algebraic Symbols

The **summation symbol** $\sum$ is an abbreviate for summation:

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$$

where i are indices. The **product symbol** $\prod$ is an abbreviation for products:

$$\prod_{i=1}^{n} x_i = x_1 \times x_2 \times x_3 \times \cdots \times x_n$$

2 Calculus

you are only required to know basic calculus, this includes knowing:

- How to find the derivative of a one-variable differentiable function.
- How to find the partial derivatives of a two-variable differentiable function.
- How to apply Chain Rule to the two cases above.

2.1 One Variable Differential Calculus

You are expected to know how to do basic one-variable differential calculus. Here are some examples:

Example 4 *Let* $f(x) = x^5 + 4x^3 + 2x^2 + 1$, *find* $f'(x)$

Example 5 *Let* $f(x) = e^x$, *find* $f'(x)$.

Example 6 *Let* $f(x) = \ln(2x)$, *find* $f'(x)$.

Example 7 *Let* $f(x) = \frac{1}{x}$, *find* $f'(x)$. *Is the derivative defined everywhere?*

Example 8 *What are other notations for the expression* $f'(x)$*?*

2.2 Two Variable Differential Calculus

You are expected to know how to find **partial derivatives** of a two-variable differentiable function:

Example 9 *Let* $f(x, y) = 2x + 3y$, *find* $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$.

Example 10 *Let* $f(x, y) = e^{x+y}$, *find* $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$.

Example 11 *Let* $f(x, y) = \ln(2x + 3y)$, *find* $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$.

2.3 Chain Rule

The Chain Rule of Calculus is crucial for machine learning and data science.

Example 12 *Let f, g be two differentiable functions, and*

$$f : \mathbb{R} \to \mathbb{R}$$

$$g : \mathbb{R} \to \mathbb{R}$$

Then:

$$(f(g(x)))' = f'(g(x))g'(x)$$

In the Leibniz Notation, let $h(x) = f(g(x))$, then:

$$\frac{dh}{dx} = \frac{dh}{dg}\frac{dg}{dx}$$

Example 13 *Let $f(x, y) = e^{x^2+y^2}$, find $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$.*

Answer:

$$\frac{\partial f}{\partial x} = 2xe^{x^2+y^2}$$
$$\frac{\partial f}{\partial y} = 2ye^{x^2+y^2}$$

3 Basic Linear Algebra

Example 14 *A d-dimensional **vector** is an element of $\mathbb{R}^d$, the d-dimensional space. Notice that the dimension is a positive integer. We often denote $\mathbf{x} \in \mathbb{R}^d$, where $\in$ is the relationship of "belonging", and $\mathbf{x}$ can be written as $\mathbf{x} = [x_1, \cdots, x_d]^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$. Note that by default we treat vectors are **column vectors** with the "vertical shape".*

Example 15 *A $d \times d$ **matrix** is an linear transformation that can be written as: $X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \cdots & \cdots & \cdots \\ x_{d1} & x_{d2} & \cdots & x_{dd} \end{bmatrix}$, and its **transpose** is just a rearrangement of entries: $X^T = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{d1} \\ x_{12} & x_{22} & \cdots & x_{d2} \\ \vdots & \cdots & \cdots & \cdots \\ x_{1d} & x_{2d} & \cdots & x_{dd} \end{bmatrix}$*

Example 16 *We often denote a $m \times n$ matrix with m as the number of rows and n the number of columns. If X is a $m \times n$ matrix, we say that $X \in \mathbb{R}^{m \times n}$.*

Example 17 *The vector-vector **dot-product** is the algebraic operation between two $d-$dimensional vectors:*

$$\mathbf{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_d y_d = \sum_{i=1}^{d} x_i y_i$$

Example 18 ***shape-matching for matrix and vectors**: If a matrix X has shape $m \times n$ (i.e., $X \in \mathbb{R}^{m \times n}$), then a matrix-vector product Xv is defined only if the vector is dimension-n (i.e., $v \in \mathbb{R}^n$), in which case $Xv \in \mathbb{R}^m$ is a m-dimensional vector:*

$$X \in \mathbb{R}^{m \times n}, v \in \mathbb{R}^n, \textit{Then } Xv \in \mathbb{R}^m$$

Example 19 *Similarly, two matrices can be multiplied together if the dimensions match:*

$$X \in \mathbb{R}^{m \times n}, Y \in \mathbb{R}^{n \times d},\ \textit{Then } XY \in \mathbb{R}^{m \times d}$$