Homework 1 — Numerical Analysis Fall 2024

Instructions:

- Due 09/10 at 5:00pm on Gradescope.
- You must follow the submission policy in the syllabus

Problem 1. Floating point operations (such as adding two floating point numbers) are some of the most basic units of modern computation (for instnace, using a neural network require adding and multiplying lots of numbers).

The Frontier supercomputer at Oak Ridge National Lab is the largest supercomputer in the world. Frontier consists of roughly 10,000 nodes (individual computers) connected together. Each node can do 10^{14} floating point operations per second.

Image of Frontier from Wikipedia.

(a) Suppose you want to train the AlexNet image classifier (state of the art in 2012). This will require roughly 10^{17} floating point operations.[1]

You are allocated a single node on Frontier. Roughly how long will it take (in minutes)?

[1] https://epochai.org/data/notable-ai-models#explore-the-data

(b) Suppose that the furthest nodes are separated by 25 meters. Roughly how many floating point operations could the whole supercomputer do in the time it takes to send a signal from one of these node to the other? (You can assume the signal travels at the speed of light)

(c) Suppose you want to train GPT-4. This requires roughly 10^{25} floating point operations, 10^8 times more operations than in AlexNet. However, this time you can use the entire supercomputer.

Because the data set is so huge, we need to store different parts of the data on different nodes.

We are doing a computation 10^8 times larger, using 10^4 times as many resources. Do you think the runtime will be more, less, or about the same than 10^4 times the runtime in (a)? Why?

Relate you answer to your response in (b).

(d) A Macbook Pro M3 can do roughly 10^{11} floating point operations per second. How long would it (in years) to do 10^{25} flops?

Problem 2. For (a), (b), and (c) you must *prove* the results for the general case.

(a) Suppose $\mathbf{A}$ is both an upper triangular matrix and a lower triangular matrix. Show that $\mathbf{A}$ is diagonal.

(b) Suppose $\mathbf{A}$ is a matrix. Show that $\mathbf{A}^\mathsf{T}\mathbf{A}$ is symmetric.

(c) Suppose $\mathbf{A}$ is a diagonal matrix and $\mathbf{B}$ is upper triangular. Show that $\mathbf{AB}$ is upper triangular.

(d) Find an example of a matrix $\mathbf{A}$ with orthonormal columns (i.e. $\mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{I}$) for which $\mathbf{A}\mathbf{A}^\mathsf{T} \neq \mathbf{I}$.

Problem 3. Define

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & -1 & -4 \\ -7 & 0 & -1 \\ 3 & 4 & 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ -2 \end{bmatrix}.$$

(a) Compute $\mathbf{AB}$.

(b) Compute $\mathbf{Bx}$.

(c) Compute $\mathbf{A}(\mathbf{Bx})$.

(d) Compute $(\mathbf{AB})\mathbf{x}$.

Show your work at each step.

Problem 4. Make sure to read the course guidelines on submitting code!

(a) Make the matrices $\mathbf{A}$, $\mathbf{B}$, and vector $\mathbf{x}$ in numpy.

(b) Verify each of the answers you generated in Problem 3 using numpy.

Problem 5. Consider linear equations $f(x) = ax + b$ and $\tilde{f}(x) = ax + (b + \epsilon)$, where a and b are constants and $\epsilon = 10^{-5}$.

(a) Find the solutions to $f(x) = 0$ and $\tilde{f}(x) = 0$ in terms of a and b.

(b) Find values of a and b so that the solutions to these two equations are very different. Here "very different" means the solutions differ by much more than ϵ.

(c) Using matplotlib, plot the functions $f(x)$ and $\tilde{f}(x)$ corresponding to your values of a and b. Make sure the plots are labeled, and that the axes bounds are chosen such that the two lines can be clearly differentiated from one another and the zeros are clearly visible.