Preliminaries - Assignment 0

Instructions : Student - Shreyas Shukla

- **This assignment is graded on submission, not correctness.**
- This assignment will help you build and evaluate 1) the necessary mathematical background and, 2) the necessary programming skills to succeed in this course. Your performance will also indicate whether you will need to do significant additional work outside of assignments and class time to meet course expectations.
- **Please be honest in completing this assignment without referring to GenAI or external resources except what is linked here.** Seemingly good performance on this assignment is not going to impress me or boost your grade in any way.[1]
- Do not try to 'find' the right answers in an effort to be perceived as a strong student - it simply does not matter to me how you do on this assignment. I only want to know your weaknesses so that I can help you overcome them.
- For each section, if you have relevant background, feel free to jump straight into solving the questions. Otherwise, I have linked some resources that should help you build the necessary background - which you should go through first.
- **Please do not hesitate to reach out to me directly** if you find anything difficult to understand. My favorite part of my job is meeting and talking to students outside of class, getting to know each other, and pacing explanations on an individual basis.
- For programming questions, please use Python 3 as efficiently as possible. This course expects you to not only write code that 'works', but also works within tight time-requirements. (More on this in section 2.)
- **Section 3 is aimed at students with prior AI experience.** If you do not have prior AI experience, please feel free to skip this section entirely. By the end of the semester, you will be able to answer these questions confidently.
- I do not expect anyone in the class to be able to answer all questions in Section 3. If you are, in fact, able to do so from your own understanding without any external help, then unless CS5100 is a required course for your degree, it may not be the right fit for you - please reach out to me, and I will be happy to help you pick.

[1] You may be asked to explain your answers in person. Pretending to know something or cheating in assignments, and not being able to explain your answers later, is a sure-fire way to lose my trust, and ruin your chances of being hired as a TA/RA at Khoury.

Remember that I'm always here to help you, and you can reach out to me for any reason throughout the semester. I am committed to fostering a nonjudgmental learning environment for all my students.

1 Mathematical Background

1.1 Vector/Matrix Operations

External Resources:

[Vector Operations - LibreTexts], [PDF Version]
[Dot Product, Vector Projection], [PDF Version]
[Matrix Operations - LibreTexts], [PDF Version]

Q1. Given the vectors

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix},$$

compute the quantities $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} \cdot \mathbf{y}$.
Solution:

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} 1+4 \\ 2+5 \\ 3+6 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}$$

$$\mathbf{x} \cdot \mathbf{y} = (1)(4) + (2)(5) + (3)(6) = 4 + 10 + 18 = 32$$

Q2. If the Euclidean (ℓ^2) norm of a vector, $\mathbf{w}$ is given by:

$$||\mathbf{w}||_2 = \sqrt{\sum_{w_i \in \mathbf{w}} w_i^2},$$

compute the Euclidean norm of $\mathbf{x} + \mathbf{y}$.
Solution:

$$||\mathbf{x} + \mathbf{y}||_2 = || \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} ||_2 = \sqrt{5^2 + 7^2 + 9^2} = \sqrt{25 + 49 + 81} = \sqrt{155}$$

Q3. Find the cosine of the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$.
Solution:

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||_2 ||\mathbf{y}||_2}$$

$$||\mathbf{x}||_2 = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

$$||\mathbf{y}||_2 = \sqrt{4^2 + 5^2 + 6^2} = \sqrt{16 + 25 + 36} = \sqrt{77}$$

$$\cos\theta = \frac{32}{\sqrt{14}\sqrt{77}} = \frac{32}{\sqrt{1078}} = \frac{32}{7\sqrt{22}}$$

Q4. Compute the projection of the vector $\mathbf{x}$ onto the vector $\mathbf{y}$.
Solution:

$$\text{proj}_{\mathbf{y}}\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{y}}{||\mathbf{y}||_2^2}\mathbf{y} = \frac{32}{77}\begin{bmatrix}4\\5\\6\end{bmatrix} = \begin{bmatrix}128/77\\160/77\\192/77\end{bmatrix}$$

Q5. Given the matrices,

$$\mathbf{A} = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix},\ \mathbf{B} = \begin{bmatrix}5 & 6 & 7\\7 & 8 & 9\end{bmatrix},$$

compute $\mathbf{A}+\mathbf{B}$, $\mathbf{A}^{\mathrm{T}}$, and $\mathbf{A}^{\mathrm{T}}\mathbf{B}$.
Solution:

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix}1+5 & 2+6 & 3+7\\4+7 & 5+8 & 6+9\end{bmatrix} = \begin{bmatrix}6 & 8 & 10\\11 & 13 & 15\end{bmatrix}$$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{B} = \begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}\begin{bmatrix}5 & 6 & 7\\7 & 8 & 9\end{bmatrix} = \begin{bmatrix}(1)(5)+(4)(7) & (1)(6)+(4)(8) & (1)(7)+(4)(9)\\(2)(5)+(5)(7) & (2)(6)+(5)(8) & (2)(7)+(5)(9)\\(3)(5)+(6)(7) & (3)(6)+(6)(8) & (3)(7)+(6)(9)\end{bmatrix} = \begin{bmatrix}33 & 38 & 43\\45 & 52 & 59\\57 & 66 & 75\end{bmatrix}$$

Q6. Compute $5\mathbf{x}$ and $5\mathbf{A}$.
Solution:

$$5\mathbf{x} = 5\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}5\\10\\15\end{bmatrix}$$

$$5\mathbf{A} = 5\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} = \begin{bmatrix}5 & 10 & 15\\20 & 25 & 30\end{bmatrix}$$

Q7. Compute $\mathbf{A}\mathbf{x}$.
Solution:

$$\mathbf{A}\mathbf{x} = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}(1)(1)+(2)(2)+(3)(3)\\(4)(1)+(5)(2)+(6)(3)\end{bmatrix} = \begin{bmatrix}1+4+9\\4+10+18\end{bmatrix} = \begin{bmatrix}14\\32\end{bmatrix}$$

Q8. If for a matrix $\mathbf{M}$,

$$\mathbf{M}^n = \prod_{i=1}^{n}\mathbf{M}$$

then compute $(\mathbf{A}^{\mathrm{T}}\mathbf{B})^2$.

Solution:

Let $\mathbf{C} = \mathbf{A}^{\mathrm{T}}\mathbf{B} = \begin{bmatrix}33 & 38 & 43\\45 & 52 & 59\\57 & 66 & 75\end{bmatrix}$. Then:

$$(\mathbf{A}^{\mathrm{T}}\mathbf{B})^2 = \mathbf{C}^2 = \mathbf{C}\cdot\mathbf{C}$$

$$\mathbf{C}^2 = \begin{bmatrix} (33)(33)+(38)(45)+(43)(57) & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & (57)(43)+(66)(59)+(75)(75) \end{bmatrix} = \begin{bmatrix} 5250 & 6068 & 6886 \\ 6738 & 7796 & 8854 \\ 8226 & 9524 & 10822 \end{bmatrix}$$

1.2 Probability Theory

External Resources:

[Seeing Theory - A Visual Intro to Prob/Stats] (Recommended)
[Seeing Theory - Full PDF Textbook]

Q9. Two fair dice are rolled together, what is the probability that the sum of the rolls is 6? **Solution:** There are 5 outcomes that sum to 6: (1,5), (2,4), (3,3), (4,2), (5,1). The total number of possible outcomes is $6 \times 6 = 36$. The probability is $5/36$.

Q10. What is the probability of drawing an ace from a shuffled deck of cards? **Solution:** There are 4 aces in a standard deck of 52 cards. The probability is $4/52 = 1/13$.

Q11. From a shuffled deck, a card drawn at random is the 7 of spades. What is the probability of drawing the ace of spades from the remaining deck at random? **Solution:** After drawing one card, 51 cards remain. The ace of spades is still in the deck. The probability of drawing it is $1/51$.

Q12. A medical test has a 95% chance of detecting a disease if the patient has it (true positive rate) and a 5% chance of a false positive. If 1% of the population has the disease, what is the probability that a person testing positive actually has the disease? **Solution:** Let D be having the disease and + be a positive test. P(D)=0.01, P(+|D)=0.95, P(+|not D)=0.05. P(not D)=0.99. Using Bayes' Theorem, $P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D)+P(+|\text{not D})P(\text{not D})} = \frac{(0.95)(0.01)}{(0.95)(0.01)+(0.05)(0.99)} = \frac{0.0095}{0.0095+0.0495} = \frac{0.0095}{0.059} = \frac{95}{590} = \frac{19}{118}$.

1.3 Calculus

External Resources:

[Differentiation Rules - LibreTexts], [PDF version]

[Partial Derivatives - LibreTexts], [PDF version]

Q13. Find the derivative of $x^2 + 4x + 1$ with respect to x, at $x = 3$.

Solution:

$$\frac{d}{dx}(x^2 + 4x + 1) = \frac{d}{dx}(x^2) + \frac{d}{dx}(4x) + \frac{d}{dx}(1) = 2x + 4$$

$$\text{At } x = 3: \quad 2(3) + 4 = 6 + 4 = \boxed{10}$$

Q14. Find the partial derivatives of the function $f(x, y) = 5x^2 + 4xy + 2y^2 + 5$ with respect to x and y, at the point $(x = 3, y = 5)$.

Solution:

Partial derivative with respect to x:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(5x^2 + 4xy + 2y^2 + 5) = 10x + 4y$$

$$\text{At } (x = 3, y = 5): \quad 10(3) + 4(5) = 30 + 20 = \boxed{50}$$

Partial derivative with respect to y:

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(5x^2 + 4xy + 2y^2 + 5) = 4x + 4y$$

$$\text{At } (x = 3, y = 5): \quad 4(3) + 4(5) = 12 + 20 = \boxed{32}$$

Q15. Find the partial derivatives of the function $f(x, y) = e^x + 4\log(y)$ with respect to x and y, at the point $(x = 3, y = 5)$.

Solution:

Partial derivative with respect to x:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(e^x + 4\log(y)) = e^x$$

$$\text{At } x = 3: \quad e^3 \approx \boxed{20.0855}$$

Partial derivative with respect to y:

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(e^x + 4\log(y)) = \frac{4}{y}$$

$$\text{At } y = 5: \quad \frac{4}{5} = \boxed{0.8}$$

2 Programming Background [2]

External Resources

[Introductory tutorial to Python and NumPy]

General Guidelines

- I strongly recommend working in a Linux-like environment and becoming proficient in terminal usage. You should not rely on Jupyter notebooks or autocomplete tools in IDEs like VSCode too much. For Windows systems, consider setting up WSL.
- Set up a Conda environment with **Python 3.11**, using this requirements.txt file. You will be using this throughout the semester.
- **For each question below, submit a) a screenshot of your code, b) the final answer, and c) your program's runtime.**
- The `time` python library will be useful in this process.

Q16. A Pythagorean triplet is a set of three natural numbers $a < b < c$, for which $a^2 + b^2 = c^2$. For example, $3^2 + 4^2 = 5^2$. There exists exactly one Pythagorean triplet for which $a + b + c = 1000$. Find the product abc.

```
(base) shreyas@Shreyass-MacBook-Pro AI % python pythag.py
Triplet: a=200, b=375, c=425
Product abc = 31875000
Runtime: 0.022395 seconds
```

```
import time

start_time = time.time()

for a in range(1, 1000):
    for b in range(a + 1, 1000 - a):
        c = 1000 - a - b
        if a * a + b * b == c * c:
            print(f"Triplet: a={a}, b={b}, c={c}")
            print(f"Product abc = {a * b * c}")
            end_time = time.time()
            print(f"Runtime: {end_time - start_time:.6f} seconds")
            exit()
```

Q17. It can be seen that the number 125874, and its double 251748, contain exactly the same digits, but in a different order. Find the smallest positive integer, x, such that $2x$, $3x$, $4x$, $5x$, and $6x$, contain the same digits.

```
(base) shreyas@Shreyass-MacBook-Pro AI % python small_positive.py
Smallest x: 142857
Runtime: 0.092440 seconds
(base) shreyas@Shreyass-MacBook-Pro AI %
```

```
import time

start_time = time.time()

def has_same_digits(x):
    digits = sorted(str(x))
    return all(sorted(str(x * i)) == digits for i in range(2, 7))

x = 1
while True:
    if has_same_digits(x):
        print("Smallest x:", x)
        break
    x += 1

end_time = time.time()
print(f"Runtime: {end_time - start_time:.6f} seconds")
```

[2] Questions from Project Euler.

Q18. A googol (10^{100}) is a massive number: one followed by one hundred zeros; 100^{100} is almost unimaginably large: one followed by two hundred zeros. Despite their size, the sum of the digits in each number is only 1. Considering natural numbers of the form, a^b, where $a, b < 100$, what is the maximum digital sum?

```python
import time

start_time = time.time()

max_sum = 0
max_a = max_b = 0

for a in range(1, 100):
    for b in range(1, 100):
        digit_sum = sum(map(int, str(a ** b)))
        if digit_sum > max_sum:
            max_sum = digit_sum
            max_a, max_b = a, b

end_time = time.time()

print(f"Maximum digital sum: {max_sum} (a={max_a}, b={max_b})")
print(f"Runtime: {end_time - start_time:.6f} seconds")
```

```
(base) shreyas@Shreyass-MacBook-Pro AI % python googl.py
Maximum digital sum: 972 (a=99, b=95)
Runtime: 0.036273 seconds
```

Q19. By starting at the top of the triangle below and moving to adjacent numbers on the row below, the maximum total from top to bottom is 23.

```
(base) shreyas@Shreyass-MacBook-Pro AI % python triangle.py
Max path sum: 7273
Runtime: 0.001055 seconds
```

```python
import time

def fetch_triangle(filepath):
    with open(filepath, 'r') as file:
        return [list(map(int, line.split())) for line in file.readlines()]

def max_path_sum(triangle):
    for row in range(len(triangle) - 2, -1, -1):
        for col in range(len(triangle[row])):
            triangle[row][col] += max(triangle[row + 1][col], triangle[row + 1][col + 1])
    return triangle[0][0]

start = time.time()
triangle_data = fetch_triangle("triangle.txt")
if triangle_data:
    result = max_path_sum(triangle_data)
    print(f"Max path sum: {result}")
    print(f"Runtime: {time.time() - start:.6f} seconds")
```

3
7 4
2 **4** 6
8 5 **9** 3

That is, $3 + 7 + 4 + 9 = 23$. Find the maximum total from top to bottom in the following text file: triangle.txt, a 15K text file containing a triangle with one hundred rows.

NOTE: It is not possible to try every route to solve this problem, as there are 2^{99} altogether! If you could check one trillion (10^{12}) routes every second, it would take over twenty billion years to check them all. Your goal is to find an efficient algorithm to solve it.

3 General AI Knowledge - Optional (just trying)

Answer in no more than 2-3 sentences per question.

Q20. What classical machine learning approach is the output neuron in a binary-classification neural network mathematically equivalent to?
Answer: Not sure

Q21. What is the purpose of hidden layers? Mathematically, why do we need them for many real-world problems?
Answer: Not sure

Q22. When training a deep reinforcement learning agent to drive a car, you observe loss decreasing steadily over time, and converging to some value. However, when using this agent to drive, the car keeps running off the road in one direction. What may have gone wrong during training?
Answer: During training, the AI model may have experienced more situations where it had to recover or move toward that side of the road, causing it to become biased in that direction.

Q23. The word 'goofy' is generally used as an adjective. However, in the sentence, 'Goofy, the dog . . . ', the same word is a proper noun. Which AI technique would you apply to figure out which part of speech a word is in a given sentence?
Answer: I think we would have to give the model context. For example, 'Goofy' is almost always an adjective, but in this case it's a name. While I don't know the exact technique name, I would train a model where I would label 'Goofy' in different contexts so it can learn to classify it correctly based on surrounding words and capitalization.

Q24. How does PyTorch achieve fast gradient computations? When might one compute the gradient of a loss function with respect to a model input?
Answer: Not sure

Q25. How would you implement a photo sharpening tool using convolution operations?
Answer: Not sure