1 Prerequisites

- **Course**: COMPSCI 589 Machine Learning, Spring, 2021
- **Instructor:** Justin Domke
- **Assignment**: 1
- **Group work policy:** No group work is permitted for this assignment.
- **Due date:** Feb 10th, 5:00 PM
- **Submission instructions**:
  - For this assignment, you can prepare your solutions in any manner you like. You may use Markdown, LaTeX, handwrite them, or a combination of these. Regardless of how they are prepared, you must produce a .pdf file that you upload to Gradescope.
  - When you submit to Gradescope, you must must mark page numbers for the different questions. We hate to do it, but we will penalize anyone who does not do this, as it creates a great deal of difficulty for the graders.
  - If you do not have access to the course on Gradescope, please send an email to Chris Nota (cnota@cs.umass.edu).
  - If the submission is late, please include a written justification as part of your assignment PDF.

Note: We have added an FAQ for small clarifications.

Section 1: Matrix Computations

Question 1 (28 points) Take the following objects:

- $c \in \mathbb{R}$
- $x \in \mathbb{R}^N$
- $y \in \mathbb{R}^M$
- $A \in \mathbb{R}^{O \times P}$
- $B \in \mathbb{R}^{Q \times R}$
- $C \in \mathbb{R}^{S \times T}$

Take the following quantities:

1. cx
2. cA
3. Ax
4. $x^\top Ay$
5. $xy^\top$
6. AB
7. ABC

For each of these, please provide:

1. What conditions must hold on $N, M, \cdots, T$ in order for the quantity to make sense. You can always assume these are finite integers no less than one, and do not need to state this.
   - If the *quantity* cannot makes sense, explain why and skip the remaining parts of the question.
2. In what space the quantity lies.
3. A short Python function that computes the quantity, assuming that c is a `float`, x, y are 1-D `numpy` arrays, and A, B, C are 2-D `numpy` arrays. You should use no Numpy functions other than (a) to create an array for your answer (like `array=numpy.zeros(...)`) and (b) indexing (like `x[i]` or `A[i,j]`). You may *optionally* use `array.shape` to inspect the sizes of input arrays, but this is not required.
4. The time complexity of your function.

To get you started, if you were given the quantity

$$x^\top y,$$

a set of correct answers would be

1) $N = M$

2) $\mathbb{R}$

3)

```
def x_dot_y(x,y): a = 0.0 N = x.shape[0] for i in range(N): a += x[i]*y[i] return a
```

4) $\mathcal{O}(N)$

FAQ

Section 2: Linear Algebra

Question 1 (18 points) Consider the matrix

$$\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}.$$

What are the eigenvectors and corresponding eigenvalues of this matrix? Make sure the eigenvectors are linearly independent. Please also give a short argument in at most five sentences.

Give two vectors of length 2, each with a corresponding scalar. For example, you might (incorrectly) say

> [3, 2] is an eigenvector with eigenvalue 21 because of blahblahblah.[– 21, 1] is an eigenvector with eigenvalue π because of blahblahblah.

Question 2 (10 points) Consider the matrix

$$\begin{bmatrix} 1 & .5 & -1 \\ 2 & 2 & -1 \\ -2 & -1 & 2 \end{bmatrix}.$$

What is this matrix's rank? Give an integer, and explain why in at most two sentences.

FAQ

Section 3: Probability

Suppose that you roll two fair, independent six-sided dice. Let the results be the random variables X and Y. Then, let

$$A = X + Y$$

be the sum of *X* and *Y*. Also let *B* = 1 if *A* is an even number and *B* = 0 otherwise.

Question 1 (3 points). What is $\mathbb{E}[B]$.

Question 2 (3 points). What is $\mathbb{E}[B|X = 2]$?

Question 3 (3 points). What is $\mathbb{E}[B|X = 1, Y = 3]$?

Question 4 (5 points). Is *X* independent of *Y*? Explain why in at most 2 sentences.

Question 5 (5 points). Is *X* independent of *A*? Explain why in at most 2 sentences.

Question 6 (5 points). Is *X* independent of *B*? Explain why in at most 2 sentences.

Section 4: Optimization

Suppose that $a_1, a_2, \cdots, a_N \in \mathbb{R}$. Give a solutions for each of the following, and a short proof that your answer is correct. You may assume that *N* is odd.

Question 1 (5 points) $\operatorname{argmin}_x (x - a_1)^2$ and $\min_x (x - a_1)^2$.

Question 2 (5 points) $\operatorname{argmin}_x (x - a_1)^2 + (x - a_2)^2$ and $\min_x (x - a_1)^2 + (x - a_2)^2$.

Question 3 (5 points) $\operatorname{argmin}_x \sum_{n=1}^{N} (x - a_n)^2$ and $\min_x \sum_{n=1}^{N} (x - a_n)^2$.

Question 4 (5 points) $\operatorname{argmin}_x |x - a_1|$ and $\min_x |x - a_1|$.

Question 5 (5 points) $\operatorname{argmin}_x |x - a_1| + |x - a_2|$ and $\min_x |x - a_1| + |x - a_2|$

Question 6 (5 points) $\operatorname{argmin}_x \sum_{n=1}^{N} |x - a_n|$

In some cases, the argmin might not be unique. If so, you can just give any minima.

FAQ