Math 446, Fall 2024, USC — Instructor: Steven Heilman

Name: ______________ USC ID: ______________ Date: ______________

Signature: ______________. Discussion Section: ______________
(By signing here, I certify that I have taken this test while refraining from cheating.)

Exam 1

This exam contains 8 pages (including this cover page) and 5 problems. Enter all requested information on the top of this page.

You may *not* use your books, notes, or any calculator on this exam.

You are required to show your work on each problem on this exam. The following rules apply:

- You have 50 minutes to complete the exam, starting at the beginning of class.
- **Organize your work**, in a reasonably neat and coherent way, in the space provided. Work scattered all over the page without a clear ordering will receive very little credit.
- **Mysterious or unsupported answers will not receive full credit**. A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.
- If you need more space, use the back of the pages; clearly indicate when you have done this. Scratch paper appears at the end of the document.

Problem	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
Total:	50	

Do not write in the table to the right. Good luck![a]

1. Label the following statements as TRUE or FALSE. If the statement is true, **explain your reasoning**. If the statement is false, **provide a counterexample or explain your reasoning**.

   (a) (2 points) In Python, the output of the code

   ```
   print(range(5))
   ```

   is

   ```
   [1, 2, 3, 4, 5]
   ```

   TRUE FALSE (circle one)

   (b) (2 points) The largest integer in Python is $(2 - 2^{-52})2^{1023}$.

   TRUE FALSE (circle one)

(c) (2 points) Let A be a 10×10 real symmetric matrix (so $A = A^T$). From the Spectral Theorem, there exists a 10×10 orthogonal matrix Q (so $QQ^T = Q^TQ = I$) and a 10×10 diagonal matrix D with real entries such that

$$A = QDQ^T.$$

Moreover, Q and D are both unique. That is, if $A = \widetilde{Q}\widetilde{D}\widetilde{Q}^T$ for some other 10×10 diagonal $\widetilde{D}$ and 10×10 real orthogonal matrix $\widetilde{Q}$, then $Q = \widetilde{Q}$ and $D = \widetilde{D}$.

TRUE FALSE (circle one)

(d) (2 points) If we enter the following command into the Python

```
strings = ["a", "as", "bat", "car", "dove", "python"]
[x for x in strings if len(x) > 2]
```

Python outputs

```
['bat', 'car', 'dove', 'python']
```

TRUE FALSE (circle one)

(e) (2 points) Python's implementation of k-means clustering is deterministic. That is, if I use a dataset and ask Python to perform k-means clustering on that dataset, the output of the `KMeans` function from `sklearn.cluster` will be the same, regardless of how many different times I ask for an output, and regardless of any random seed that is provided to Python.

TRUE FALSE (circle one)

2. (10 points) Describe the output of the following program.

```
x = 1
while x != 0:
    x = x / 2
    print(x)
```

Describe in detail what the program does, how many iterations the while loop performs, and why exactly the while loop terminates in that many iterations.

3. (10 points) Recall that two real-valued random variables $X, Y : \Omega \to \mathbf{R}$ with nonzero variance have correlation defined by

$$\mathrm{Corr}(X, Y) := \mathbf{E}\Big(\frac{X - \mathbf{E}X}{\sqrt{\mathrm{Var}(X)}}\frac{Y - \mathbf{E}Y}{\sqrt{\mathrm{Var}(Y)}}\Big),$$

where $\mathbf{E}$ denotes expected value and $\mathrm{Var}(X) := \mathbf{E}(X - \mathbf{E}X)^2$ denotes variance.

In this problem we will consider a sample space $\Omega = \{1, \ldots, n\}$ with n points and with the uniform probability law on Ω, so that $\mathbf{E}X = \frac{1}{n}\sum_{i=1}^{n} X(i)$.

Write a function in Python whose input is two Numpy arrays `X` and `Y` of the same length both with nonzero variance, and whose output is $\mathrm{Corr}(X, Y)$. The name of your function should be `my_correlation`. (You have to compute the correlation yourself, you cannot just write a one-line program that uses some built in Python correlation computation.)

Hint: you can use the following Numpy built-in functions: `np.mean` and `np.sqrt`

(You can and should assume we already ran the command `import numpy as np` .)

(Your program can freely assume that X and Y have nonzero variance, i.e. you do not have to check whether or not their variances are zero.)

4. (10 points) Suppose the singular value decomposition

$$A = UDV$$

of a real 100 by 100 matrix A has singular values $6, 5$ and 4, and the remaining 97 singular values are all at most $1/1000$. That is, D is a diagonal matrix with $D_{11} = 6$, $D_{22} = 5$, $D_{33} = 4$, and $0 \leq D_{ii} \leq 1/1000$ for all $3 < i \leq 100$. Explain in detail why PCA with $q = 3$ principal components outputs a matrix that closely approximates A. That is, justify why A is close to

$$U \begin{pmatrix} 6 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 5 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 4 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} V$$

5. (10 points) It is observed in practice that sometimes subtracting the average value of a matrix before applying Principal Component Analysis can improve the performance of PCA. With this heuristic in mind, consider the matrix

$$A = \begin{pmatrix} 90 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

and also consider the mean subtracted matrix

$$\widetilde{A} = A - 10 = \begin{pmatrix} 80 & -10 & -10 \\ -10 & -10 & -10 \\ -10 & -10 & -10 \end{pmatrix}.$$

Are the eigenvectors of A and $\widetilde{A}$ the same? Explain.

Now instead consider the matrix

$$B = \begin{pmatrix} 10 & -5 & -2 \\ -5 & 10 & -2 \\ -2 & -2 & 7 \end{pmatrix}.$$

and also consider the mean subtracted matrix

$$\widetilde{B} = B - 1 = \begin{pmatrix} 9 & -6 & -3 \\ -6 & 9 & -3 \\ -3 & -3 & 6 \end{pmatrix}.$$

Are the eigenvectors and eigenvalues of B and $\widetilde{B}$ the same? Explain.

(Scratch paper)