MATH 503 (Mathematics for Data Science)
T251 – Final Exam

Name:________________________________ **ID:**________________________

1. Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 1 & -1 & 1 \\ 4 & -8 & 3 & -3 & 1 \\ -2 & 4 & -2 & -1 & 4 \\ 1 & -2 & 0 & -3 & 4 \end{pmatrix}$$

with eigenvalues of $\mathbf{A}^T\mathbf{A}$: $132.87, 44.93, 0.21, 0$ and matrices $\mathbf{U}, \mathbf{V}$ given below.

$$\mathbf{U} = \begin{pmatrix} -0.23 & -0.12 & 0.85 & 0.50 \\ -0.86 & -0.06 & 0.05 & -0.50 \\ 0.35 & -0.73 & 0.29 & -0.50 \\ -0.27 & -0.66 & -0.47 & 0.50 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} -0.40 & 0.81 & -0.31 & 0.28 & -0.07 \\ 0.07 & -0.13 & 0.17 & 0.45 & -0.86 \\ -0.09 & 0.17 & 0.85 & 0.35 & 0.32 \\ 0.89 & 0.45 & 0.004 & -0.008 & -0.004 \\ 0.16 & -0.30 & -0.38 & 0.77 & 0.38 \end{pmatrix}.$$

a. What are the singular values of $\mathbf{A}$?

b. Write the singular value matrix $\mathbf{S}$ in the SVD of $\mathbf{A}$

c. What is the smallest possible error (in spectral norm) in approximating **A** by a rank-2 matrix?

d. Write down the economical SVD of A.

e. What will be the 2-norm condition number of $A^T A$.

2. Suppose the matrix

$$\mathbf{A} = \begin{pmatrix} 6 & -1 \\ -1 & 6 \end{pmatrix}$$

has the diagonalization $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ with

$$\mathbf{D} = \begin{pmatrix} 5 & 0 \\ 0 & 7 \end{pmatrix}.$$

a. Find a matrix $\mathbf{P}$.

b. Write out the matrices $\mathbf{D}$ and $\mathbf{P}$ in the diagonalization of $\mathbf{A}^3$.

3. Consider fitting the plane:

$$d = m_0 + m_1 x + m_2 y$$

to the data:

x	-1	0	1	0
y	-1	2	-1	0
d	2	-1	4	1

a. Write down the loss function ℓ_2. (in terms of m_is)

b. Compute the gradient of the loss function

c. Find the first iterate $\mathbf{m}^{(1)}$ by using gradient descent with $\mathbf{m}^{(0)} = [1, 0, -1]^T$, using $\alpha = \frac{1}{4}$.

d. Find the third iterate $\mathbf{m}^{(2)}$ using stochastic gradient descent with $\alpha = \frac{1}{2}$ and sample order: row 2, row 3.

Useful Formulas:

$$L(\mathbf{m}) = \|\mathbf{A}\mathbf{m} - \mathbf{d}\|^2$$

$$\nabla L(\mathbf{m}) = 2\mathbf{A}^T(\mathbf{A}\mathbf{m} - \mathbf{d})$$

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha_k \mathbf{p}^{(k)}$$

$$\mathbf{p}^{(k)} = -\mathbf{A}^T\mathbf{r}^{(k)}$$

$$\mathbf{r}^{(k)} = \mathbf{d} - \mathbf{A}\mathbf{m}^{(k)}$$

$$\alpha_k = \frac{\mathbf{p}^{(k)T}\mathbf{p}^{(k)}}{(\mathbf{A}\mathbf{p}^{(k)})^T(\mathbf{A}\mathbf{p}^{(k)})}$$

$$\alpha_k \leq \frac{2}{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}$$

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha(\mathbf{A}_i^T\mathbf{A}_i\mathbf{m}^{(k)} - \mathbf{A}_i^T d_i)$$

A_i is the ith row of A; d_i is the corresponding entry of d.

Good luck :)