UNIVERSITY OF CALIFORNIA, BERKELEY
DEPARTMENT OF STATISTICS

Homework #3

Student name: *Shizhe Zhang*

Course: *EECS 227AT* – Professor: *Gireeja Ranade*
Date: *February 8, 2026*

Problem 1. Norms

(a) Show that the following inequalities hold for any vector $\vec{x} \in \mathbb{R}^n$:

$$\frac{1}{\sqrt{n}}\|\vec{x}\|_2 \leq \|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1 \leq \sqrt{n}\|\vec{x}\|_2 \leq n\|\vec{x}\|_\infty. \tag{1}$$

NOTE: We can interpret different norms as different ways of computing distance between two points $\vec{x}, \vec{y} \in \mathbb{R}^2$. The ℓ_2 norm is the distance as the crow flies (i.e. point-to-point distance), the ℓ_1 norm, also known as the Manhattan distance, is the distance you would have to cover if you were to navigate from $\vec{x}$ to $\vec{y}$ via a rectangular street grid, and the ℓ_∞ norm is the maximum distance that you have to travel in either the north–south or the east–west direction.

(b) We define the sparsity of the vector $\vec{x}$ as the number of non-zero elements in $\vec{x}$. This is also commonly known as the ℓ_0 norm of the vector $\vec{x}$, denoted by $\|\vec{x}\|_0$. Show that for any non-zero vector $\vec{x}$,

$$\|\vec{x}\|_0 \geq \frac{\|\vec{x}\|_1^2}{\|\vec{x}\|_2^2}. \tag{2}$$

Find all vectors $\vec{x}$ for which the lower bound is attained.

Answer.

(a)

Proof. We show the inequalities one by one. First, we have

$$\|\vec{x}\|_\infty = \max_i |x_i| \leq \sqrt{\sum_{i=1}^n x_i^2} = \|\vec{x}\|_2. \tag{3}$$

This complets the first and last inequalities. Next, we have

$$\|\vec{x}\|_\infty = \max_i |x_i| \leq \sqrt{\sum_{i=1}^n x_i^2} = \|\vec{x}\|_2 \tag{4}$$

Next, W.L.O.G., we assume the entries of $\vec{x}$ are non-negative and $\|\vec{x}\|_2^2 = 1$.

$$\|\vec{x}\|_2^2 = 1 = (\sum_{i=1}^{n} x_i^2) \leq (\sum_{i=1}^{n} x_i)^2 = \|\vec{x}\|_1^2. \tag{5}$$

Finally, consider the vector $\tilde{\mathbf{1}} = [1, 1, \ldots, 1]^\top \in \mathbb{R}^n$. By Cauchy–Schwarz inequality, we have

$$\|\vec{x}\|_1 = \tilde{\mathbf{1}}^\top \vec{x} \leq \|\tilde{\mathbf{1}}\|_2 \|\vec{x}\|_2 = \sqrt{n}\|\vec{x}\|_2. \tag{6}$$

For $\vec{x}$ with negative entries, we can apply the above argument to $|\vec{x}|$ and get the same result.

□

(b)

Proof. Since the L_1 and L_2 norms remains unchanged if we delete all the zeros in $\vec{x}$ and only keep the non-zero entries.

Denote the number of non-zero entries in $\vec{x}$ as k. Then we can apply the Cauchy–Schwarz inequality to the vector of non-zero entries and get

By Cauchy–Schwarz inequality, we have

$$\|\vec{x}\|_1^2 = (\sum_{i=1}^{k} |x_i|)^2 \leq (\sum_{i=1}^{k} 1^2)(\sum_{i=1}^{k} x_i^2) = \|\vec{x}\|_0 \|\vec{x}\|_2^2. \tag{7}$$

The lower bound is attained when the ratio of non-zero entries of $\vec{x}$ to $\vec{\mathbf{1}}$ are equal in magnitude. i.e., the absolute value of the non-zero entries of $\vec{x}$ are all the same.

□

Problem 2. Diagonalization and Singular Value Decomposition

Let matrix
$$A = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

(a) Compute the eigenvalues and associated eigenvectors of A.

(b) Express A as $P\Lambda P^{-1}$, where Λ is a diagonal matrix and $PP^{-1} = I$. State P, Λ, and P^{-1} explicitly.

(c) Compute $\lim_{k\to\infty} A^k$.

(d) Give the singular values σ_1 and σ_2 of A.

Answer.

(a) The characteristic polynomial is

$$\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ \frac{1}{2} & \frac{1}{2} - \lambda \end{bmatrix} = (-\lambda)\left(\frac{1}{2} - \lambda\right) - \frac{1}{2} = \lambda^2 - \frac{1}{2}\lambda - \frac{1}{2}.$$

So the eigenvalues are
$$\lambda_1 = 1, \qquad \lambda_2 = -\frac{1}{2}.$$

For $\lambda_1 = 1$,

$$\begin{bmatrix} -1 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \implies -v_1 + v_2 = 0 \implies v_2 = v_1,$$

so an eigenvector is
$$v^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

For $\lambda_2 = -\frac{1}{2}$,

$$\begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \implies \frac{1}{2}v_1 + v_2 = 0 \implies v_2 = -\frac{1}{2}v_1,$$

so an eigenvector is
$$v^{(2)} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

(b) Let
$$P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}.$$

Then the columns of P are eigenvectors of A, so $AP = P\Lambda$ and hence $A = P\Lambda P^{-1}$.

For 2×2 matrix, the inverse is swapping the diagonal entries, negating the off-diagonal entries, and dividing by the determinant.

Compute
$$\det(P) = 1 \cdot 1 - (-2) \cdot 1 = 3,$$

so

$$P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}.$$

Therefore,

$$A = P\Lambda P^{-1} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}\frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}.$$

(c) Using diagonalization,

$$A^k = P\Lambda^k P^{-1} = P\begin{bmatrix} 1^k & 0 \\ 0 & \left(-\frac{1}{2}\right)^k \end{bmatrix}P^{-1}.$$

Since $\left(-\frac{1}{2}\right)^k \to 0$ as $k \to \infty$, we get

$$\lim_{k\to\infty} A^k = P\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}P^{-1}.$$

Compute this product: so

$$\lim_{k\to\infty} A^k = P\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}P^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\cdot\frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}.$$

Hence,

$$\boxed{\lim_{k\to\infty} A^k = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.}$$

(d) The singular values are the square roots of the eigenvalues of $A^\top A$. We compute

$$A^\top A = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{4} \end{bmatrix}.$$

The characteristic polynomial of $A^\top A$ is

$$\det\left(A^\top A - \mu I\right) = \det\begin{bmatrix} \frac{1}{4} - \mu & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{4} - \mu \end{bmatrix} = \left(\frac{1}{4} - \mu\right)\left(\frac{5}{4} - \mu\right) - \frac{1}{16} = \mu^2 - \frac{3}{2}\mu + \frac{1}{4}.$$

Thus,

$$\mu = \frac{\frac{3}{2} \pm \sqrt{\left(\frac{3}{2}\right)^2 - 1}}{2} = \frac{\frac{3}{2} \pm \sqrt{\frac{5}{4}}}{2} = \frac{3 \pm \sqrt{5}}{4}.$$

Therefore the singular values (ordered decreasingly) are

$$\sigma_1 = \sqrt{\frac{3+\sqrt{5}}{4}} = \frac{1}{2}\sqrt{3+\sqrt{5}}, \qquad \sigma_2 = \sqrt{\frac{3-\sqrt{5}}{4}} = \frac{1}{2}\sqrt{3-\sqrt{5}}.$$

Problem 3. Interpreting the Data Matrix

Suppose we have n data points, each with d features, arranged in a data matrix $X \in \mathbb{R}^{n \times d}$, which can be written equivalently as

$$X = \begin{bmatrix} \leftarrow \vec{x}_1^\top \rightarrow \\ \leftarrow \vec{x}_2^\top \rightarrow \\ \vdots \\ \leftarrow \vec{x}_n^\top \rightarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{f}_1 & \vec{f}_2 & \cdots & \vec{f}_d \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}. \tag{3}$$

Here $\vec{x}_i \in \mathbb{R}^d$ denotes the i-th data point and $\vec{f}_j \in \mathbb{R}^n$ denotes the j-th feature vector. For subproblems that require Python, assume X is stored as a NumPy array.

(a) Let $k \geq 1$ and define $\vec{1}$ as the vector of all ones. The empirical mean of $\vec{y} \in \mathbb{R}^k$ is defined as

$$\mu(\vec{y}) := \frac{1}{k}\vec{1}^\top \vec{y} = \frac{1}{k}\sum_{i=1}^{k} y_i. \tag{4}$$

What is the length of the vector of empirical feature means? Which of the following Python commands will generate this vector?

i. `mu = numpy.mean(X, axis = 0)`

ii. `mu = numpy.mean(X, axis = 1)`

(b) The empirical variance of $\vec{y} \in \mathbb{R}^k$ is

$$\sigma^2(\vec{y}) := \frac{1}{k}\|\vec{y} - \mu(\vec{y})\vec{1}\|_2^2 = \frac{1}{k}\sum_{i=1}^{k}(y_i - \mu(\vec{y}))^2. \tag{5}$$

The empirical standard deviation is

$$\sigma(\vec{y}) := \sqrt{\sigma^2(\vec{y})}. \tag{6}$$

What is the length of the vector of empirical standard deviations? Which of the following Python commands will generate this vector?

i. `sigma = numpy.std(X, axis = 0)`

ii. `sigma = numpy.std(X, axis = 1)`

(c) Suppose we want to modify X so that each feature vector is centered. How would you achieve this using Python code?

(d) Suppose we want to modify X so that each feature vector is standardized to have zero mean and unit variance. How would you achieve this using Python code?

Answer.

(a) The vector of empirical means contains $\mu(\vec{f}_1), \ldots, \mu(\vec{f}_d)$, one mean per feature, so its length is d.

```
mu = numpy.mean(X, axis = 0).
```

(b) Similarly, the vector of empirical standard deviations contains $\sigma(\vec{f}_1), \ldots, \sigma(\vec{f}_d)$, so its length is d. The correct command is:

```
sigma = numpy.std(X, axis = 0).
```

(c) Compute the feature means and subtract them from each row using broadcasting:

```
mu = numpy.mean(X, axis=0)    X_centered = X - mu.
```

(d) Compute feature means and standard deviations:

```
mu = numpy.mean(X, axis=0)
sigma = numpy.std(X, axis=0)
X_std = (X - mu) / sigma.
```

Problem 3 Cont'd

(e) The empirical covariance of $\vec{w}, \vec{y} \in \mathbb{R}^k$ is

$$\sigma(\vec{w}, \vec{y}) := \frac{1}{k}(\vec{w} - \mu(\vec{w})\vec{1})^\top (\vec{y} - \mu(\vec{y})\vec{1}) = \frac{1}{k}\sum_{i=1}^{k}(w_i - \mu(\vec{w}))(y_i - \mu(\vec{y})). \quad (7)$$

What is $\sigma(\vec{y}, \vec{y})$ in terms of the previously defined statistics?

(f) Assume X is centered. Let $\Sigma(X) \in \mathbb{R}^{d\times d}$ denote the empirical covariance matrix with entries

$$\Sigma(X)_{i,j} := \sigma(\vec{f}_i, \vec{f}_j). \quad (8)$$

Show that

$$\Sigma(X) = \frac{1}{n}X^\top X. \quad (9)$$

Then show that

$$\frac{1}{n}X^\top X = \frac{1}{n}\sum_{i=1}^{n}\vec{x}_i\vec{x}_i^\top. \quad (10)$$

(g) Let $\vec{b}$ be a unit vector in $\mathbb{R}^n$. Define vector, scalar, and projection length of $\vec{a}$ onto $\vec{b}$. Show that the vector of scalar projections satisfies

$$\vec{p} = X\vec{w}. \quad (11)$$

(h) Let $p_i := \vec{x}_i^\top \vec{w}$ and $\vec{p} = [p_1, \ldots, p_n]^\top$. Show that

$$\sigma^2(\vec{p}) = \frac{1}{n}\vec{w}^\top X^\top X\vec{w} = \vec{w}^\top \Sigma(X)\vec{w}. \quad (12)$$

Answer.

(e) It is the empirical variance of $\vec{y}$.

(f) Assume X is centered, so $\mu(\vec{f}_j) = 0$ for all j.

We first show $\Sigma(X) = \frac{1}{n}X^\top X$.

For any $i, j \in \{1, \ldots, d\}$, the (i, j) entry of $X^\top X$ is

$$(X^\top X)_{i,j} = \vec{f}_i^\top \vec{f}_j.$$

Since the data is centered, the empirical covariance is

$$\Sigma(X)_{i,j} = \sigma(\vec{f}_i, \vec{f}_j) = \frac{1}{n}\vec{f}_i^\top \vec{f}_j.$$

Therefore,

$$\Sigma(X)_{i,j} = \left(\frac{1}{n}X^\top X\right)_{i,j} \quad \text{for all } i, j,$$

Then, write X by rows:

$$X = \begin{bmatrix} \vec{x}_1^\top \\ \vec{x}_2^\top \\ \vdots \\ \vec{x}_n^\top \end{bmatrix}.$$

Then

$$X^\top X = \sum_{i=1}^{n} \vec{x}_i \vec{x}_i^\top$$

(g) Let $\vec{w} \in \mathbb{R}^d$ be a unit vector direction in feature-space. Define $\vec{p} \in \mathbb{R}^n$ by

$$p_i := \vec{x}_i^\top \vec{w}, \quad i = 1, \ldots, n, \qquad \vec{p} := [p_1, \ldots, p_n]^\top.$$

The i-th entry is the dot product of row i with $\vec{w}$,

$$(X\vec{w})_i = \vec{x}_i^\top \vec{w} = p_i.$$

Concatenating these entries, we get

$$\vec{p} = X\vec{w}. \tag{11}$$

(h) Since X is centered, the empirical mean of $\vec{p} = X\vec{w}$ is zero:

$$\mu(\vec{p}) = \frac{1}{n}\vec{1}^\top \vec{p} = \frac{1}{n}\vec{1}^\top X\vec{w} = \left(\frac{1}{n}\vec{1}^\top X\right)\vec{w} = 0,$$

Therefore,

$$\sigma^2(\vec{p}) = \frac{1}{n}\|\vec{p}\|_2^2 = \frac{1}{n}(X\vec{w})^\top(X\vec{w}) = \frac{1}{n}\vec{w}^\top X^\top X\vec{w}.$$

Using part (f), $\Sigma(X) = \frac{1}{n}X^\top X$, so

$$\sigma^2(\vec{p}) = \vec{w}^\top \Sigma(X)\vec{w}. \tag{12}$$

Problem 4. Understanding Ellipses

Consider the Euclidean space $\mathbb{R}^2$ with the orthogonal basis $\{\vec{e}_1, \vec{e}_2\}$. In this exercise, we study the ellipse

$$E = \left\{ x_1\vec{e}_1 + x_2\vec{e}_2 \;\middle|\; x_1, x_2 \in \mathbb{R}, \left(\sqrt{5}x_1 - \frac{3}{\sqrt{5}}x_2\right)^2 + \left(\frac{4}{\sqrt{5}}x_2\right)^2 \le 8 \right\}. \tag{13}$$

(a) Show that we can express the ellipse as

$$E = \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}^\top A\vec{x} \le 1\}$$

for symmetric positive definite A, where

$$A = \frac{1}{8}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{14}$$

(b) Show that the ellipse E can be viewed as a linear transformation of the unit disk by finding B such that

$$E = \{B\vec{v} \mid \|\vec{v}\|_2 \le 1\}.$$

Is this B unique?

(c) Relate the length and direction of the semi-major and semi-minor axes of E to the singular values of B (or eigenvalues of A).

(d) Compute the area of E.

Answer.

(a)

$$\left(\sqrt{5}x_1 - \frac{3}{\sqrt{5}}x_2\right)^2 + \left(\frac{4}{\sqrt{5}}x_2\right)^2 \le 8.$$

$$5x_1^2 - 6x_1x_2 + \left(\frac{9}{5} + \frac{16}{5}\right)x_2^2 = 5x_1^2 - 6x_1x_2 + 5x_2^2.$$

The original constraint is equivalent to

$$5x_1^2 - 6x_1x_2 + 5x_2^2 \le 8.$$

Write this as a quadratic form:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le 8,$$

i.e.

$$\vec{x}^\top\left(\frac{1}{8}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}\right)\vec{x} \le 1.$$

(b) Since $A \succ 0$, it has a (unique) symmetric positive definite square root $A^{1/2}$. Then

$$\vec{x}^\top A\vec{x} \le 1 \iff \|A^{1/2}\vec{x}\|_2^2 \le 1 \iff \|A^{1/2}\vec{x}\|_2 \le 1.$$

Let $\vec{v} = A^{1/2}\vec{x}$. Then $\|\vec{v}\|_2 \le 1$ and

$$\vec{x} = A^{-1/2}\vec{v}.$$

Therefore,

$$E = \{B\vec{v} \mid \|\vec{v}\|_2 \le 1\} \quad \text{with} \quad B = A^{-1/2}.$$

We can compute B explicitly by orthogonal diagonalization of A.

$$A = \frac{1}{8}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}.$$

The characteristic polynomial of $A = \frac{1}{8}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$ is

$$\det(A - \alpha I) = \det\begin{bmatrix} \frac{5}{8} - \alpha & -\frac{3}{8} \\ -\frac{3}{8} & \frac{5}{8} - \alpha \end{bmatrix} = \left(\frac{5}{8} - \alpha\right)^2 - \left(-\frac{3}{8}\right)^2 = \alpha^2 - \frac{5}{4}\alpha + \frac{1}{4}.$$

So the eigenvalues of A are 1 and $\frac{1}{4}$ with corresponding orthonormal eigenvectors

$$q_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad q_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Let $Q = [q_1\ q_2]$. Then

$$A = Q\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} Q^\top.$$

Thus

$$B = A^{-1/2} = Q\begin{bmatrix} 1^{-1/2} & 0 \\ 0 & (\frac{1}{4})^{-1/2} \end{bmatrix} Q^\top = Q\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^\top.$$

Carrying out the multiplication yields

$$B = A^{-1/2} = \frac{1}{2}\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Consider U as any 2×2 orthogonal matrix, then

$$\{BU\vec{v} \mid \|\vec{v}\|_2 \le 1\} = \{B\vec{w} \mid \|\vec{w}\|_2 \le 1\}$$

Thus, B is not unique.

(c) The semi-axis lengths are the singular values of B, which is the same as the eigenvalues of B since B is symmetric positive definite.

$$\sigma_1 = \frac{1}{\sqrt{1/4}} = 2, \qquad \sigma_2 = \frac{1}{\sqrt{1}} = 1$$

The semi-major axis length is 2, and its direction is $q_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

The semi-minor axis length is 1, and its direction is $q_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(d) The unit disk in $\mathbb{R}^2$ has area π. After the projection by B, the area is scaled by $|\det(B)|$. Since the singular values of B are 2 and 1, we have

$$|\det(B)| = \sigma_1\sigma_2 = 2 \cdot 1 = 2.$$

Hence

$$\text{Area}(E) = 2\pi.$$

Problem 5. SVD Part 2

Consider A to be the 4×3 matrix

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix}. \tag{15}$$

Here $\vec{a}_i$ for $i \in \{1, 2, 3\}$ form a set of orthogonal vectors satisfying

$$\|\vec{a}_1\|_2 = 3, \qquad \|\vec{a}_2\|_2 = 2, \qquad \|\vec{a}_3\|_2 = 1.$$

(a) What is the compact SVD of A? Express it as

$$A = U\Sigma V^\top,$$

with Σ the diagonal matrix of singular values ordered in decreasing fashion, and explicitly describe U and V.

(b) What is the dimension of the null space, $\dim(\mathcal{N}(A))$?

(c) What is the rank of A, $\operatorname{rank}(A)$? Provide an orthonormal basis for the range of A.

(d) Let I_3 denote the 3×3 identity matrix. Consider the matrix

$$\tilde{A} = \begin{bmatrix} A \\ I_3 \end{bmatrix} \in \mathbb{R}^{7\times 3}.$$

What are the singular values of $\tilde{A}$ (in terms of the singular values of A)?

Answer.

$$\vec{a}_i^\top \vec{a}_j = 0 \quad (i \neq j), \qquad \|\vec{a}_1\|_2 = 3,\ \|\vec{a}_2\|_2 = 2,\ \|\vec{a}_3\|_2 = 1.$$

Define normalized vectors

$$u_1 = \frac{\vec{a}_1}{\|\vec{a}_1\|_2} = \frac{\vec{a}_1}{3}, \qquad u_2 = \frac{\vec{a}_2}{\|\vec{a}_2\|_2} = \frac{\vec{a}_2}{2}, \qquad u_3 = \frac{\vec{a}_3}{\|\vec{a}_3\|_2} = \vec{a}_3.$$

Since the $\vec{a}_i$ are orthogonal and nonzero, u_1, u_2, u_3 are orthonormal in $\mathbb{R}^4$.

(a) We have

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, one valid compact SVD is

$$A = U\Sigma V^\top,$$

where

$$U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \in \mathbb{R}^{4\times 3}, \qquad \Sigma = \operatorname{diag}(3, 2, 1) \in \mathbb{R}^{3\times 3}, \qquad V = I_3 \in \mathbb{R}^{3\times 3}.$$

(b) The three columns are nonzero and orthogonal $\to$ linearly independent. By rank-nullity,

$$\dim(\mathcal{N}(A)) = 3 - \text{rank}(A) = 3 - 3 = 0.$$

(c) As argued above, $\text{rank}(A) = 3$. And the range of A is the span of its columns:

$$\mathcal{R}(A) = \text{span}(\vec{a}_1, \vec{a}_2, \vec{a}_3).$$

An orthonormal basis is given by the normalized columns:

$$\left\{ \frac{\vec{a}_1}{3}, \frac{\vec{a}_2}{2}, \vec{a}_3 \right\} = \{u_1, u_2, u_3\}.$$

(d) Compute

$$\vec{A}^\top \vec{A} = \begin{bmatrix} A^\top & I_3 \end{bmatrix} \begin{bmatrix} A \\ I_3 \end{bmatrix} = A^\top A + I_3.$$

Since the columns of A are orthogonal with norms $3, 2, 1$, we have

$$A^\top A = \begin{bmatrix} \vec{a}_1^\top \vec{a}_1 & \vec{a}_1^\top \vec{a}_2 & \vec{a}_1^\top \vec{a}_3 \\ \vec{a}_2^\top \vec{a}_1 & \vec{a}_2^\top \vec{a}_2 & \vec{a}_2^\top \vec{a}_3 \\ \vec{a}_3^\top \vec{a}_1 & \vec{a}_3^\top \vec{a}_2 & \vec{a}_3^\top \vec{a}_3 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore

$$\vec{A}^\top \vec{A} = A^\top A + I_3 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

$$\tilde{\sigma}_1 = \sqrt{10}, \qquad \tilde{\sigma}_2 = \sqrt{5}, \qquad \tilde{\sigma}_3 = \sqrt{2}.$$

More generally, if A has singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$, then $\tilde{A}$ has singular values

$$\tilde{\sigma}_i = \sqrt{\sigma_i^2 + 1}, \quad i = 1, 2, 3.$$

Problem 6. Homework Process

With whom did you work on this homework? List names and SIDs.
If you did not work with anyone, write "none".

Answer. none