Assignment 1

ECE 712: Matrix Computations for Signal Processing

Name: Dai, Jiatao

Student ID: 400548587

Date: October 12, 2025

McMaster University

Department of Electrical and Computer Engineering

October 12, 2025

Contents

1. Question 1: PCA Compression and Reconstruction Error

1.1 Method

In this question, we are given a 1000×5 data matrix X (from `X.mat`). First, each column of X is mean-centered to get $X_c = X - \bar{X}$. Then we compute the covariance matrix:

$$C = \frac{1}{N-1} X_c^\top X_c,$$

where $N = 1000$.

Next, we perform eigen-decomposition of C:

$$C = V \Lambda V^\top,$$

where Λ contains the eigenvalues and V the eigenvectors. For a given number of components r, we take the first r columns of V as V_r and project the data:

$$Y = X_c V_r.$$

We can then reconstruct the data (in the centered form) by

$$\hat{X}_r = Y V_r^\top = X_c V_r V_r^\top.$$

The reconstruction error is measured using the Frobenius norm:

$$E(r) = \|X_c - \hat{X}_r\|_F^2.$$

We also compute the cumulative variance explained by the first r components:

$$\text{Cumulative Variance}(r) = \frac{\sum_{i=1}^r \lambda_i}{\sum_{i=1}^5 \lambda_i}.$$

1.2 Results

The PCA results from the dataset are summarized below.

- Eigenvalues (sorted): $\{5.0849,\ 0.0914,\ 0.0272,\ 0.0152,\ 0.0103\}$.
- Cumulative variance explained (%): $\{97.25,\ 98.99,\ 99.51,\ 99.80,\ 100.0\}$.
- Reconstruction error $E(r)$: $\{143.91,\ 52.59,\ 25.45,\ 10.30,\ \approx 0\}$.

Table 1: Summary of PCA results for $r = 1$ to 5.

r	Eigenvalue λ_r	Cumulative Var. (%)	$E(r)$
1	5.0849	97.25	143.91
2	0.0914	98.99	52.59
3	0.0272	99.51	25.45
4	0.0152	99.80	10.30
5	0.0103	100.0	≈ 0

The figures below show the reconstruction error, cumulative variance, and eigenvalue distribution.

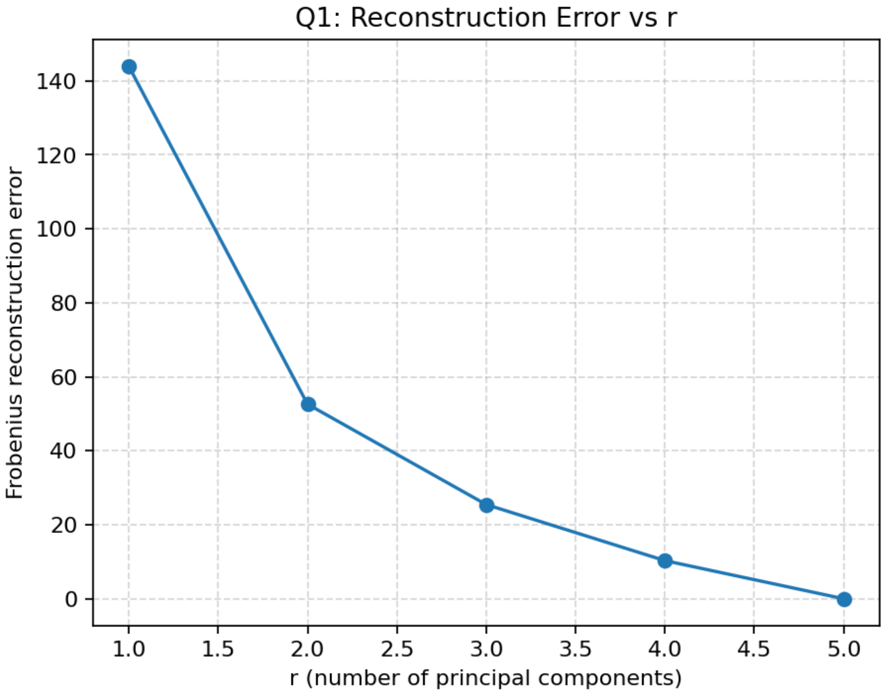

Figure 1: Reconstruction error vs. number of components r.

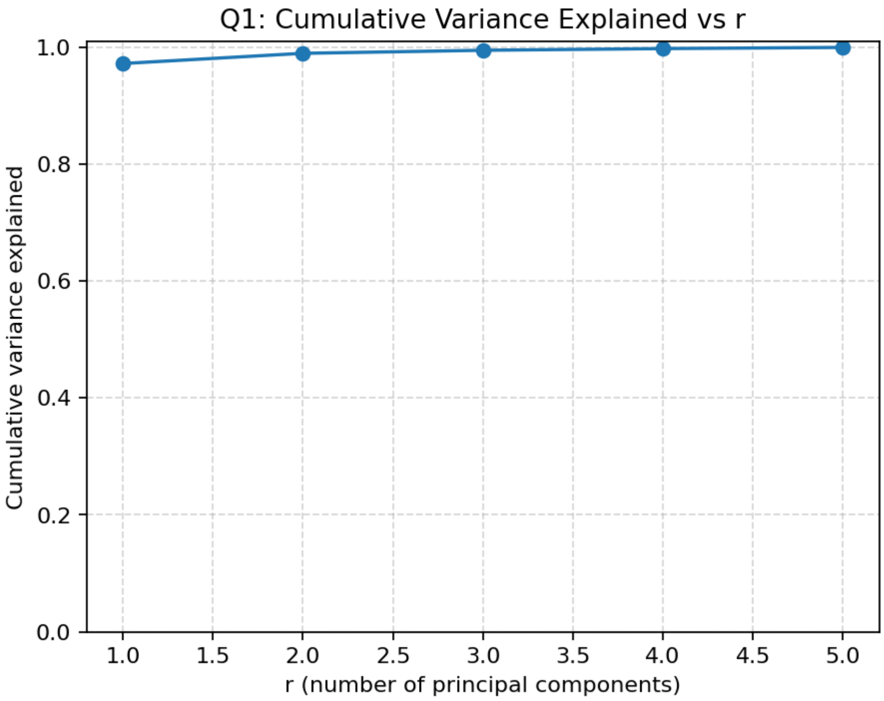

Figure 2: Cumulative variance explained by top r components.

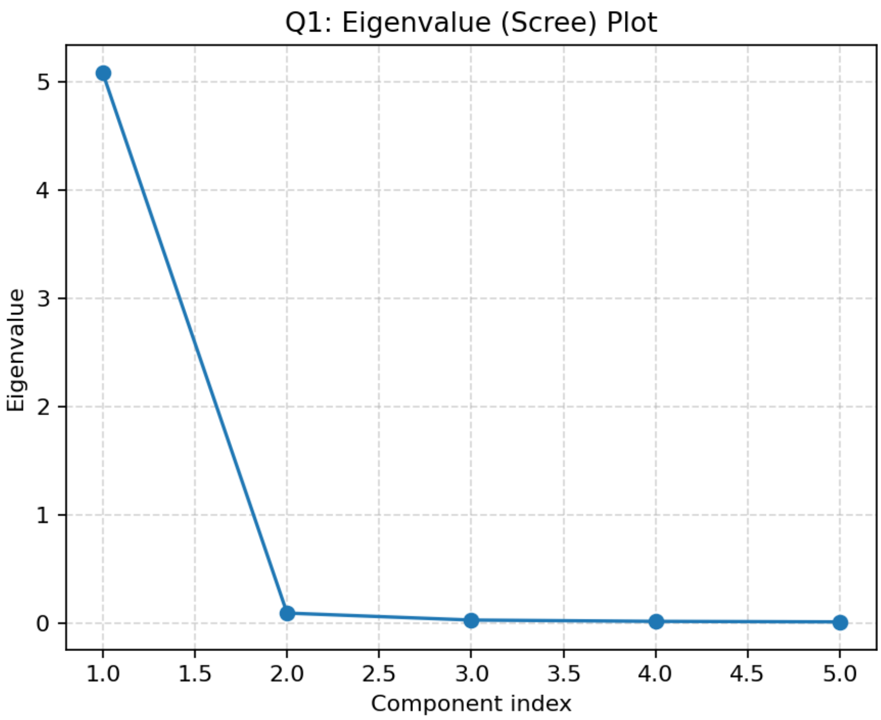

Figure 3: Eigenvalue (scree) plot.

1.3 Discussion

From the results, we can see that most of the variance (around 97%) is already captured by the first component. The remaining components only add a small amount of variance.

As r increases, the reconstruction error keeps getting smaller, which makes sense since we are keeping more information.

1.4 Conclusion

In this dataset, using just the first principal component already keeps most of the data variation. If we use more components, the error decreases but the improvement becomes smaller. A reasonable choice is $r = 1$ or $r = 2$, depending on how much accuracy we need.

2. Question 2: Proof of Minimum Reconstruction Error of PCA

2.1 Objective

Let $X \in \mathbb{R}^{m \times d}$ be a mean-centered data matrix. For any r-dimensional orthonormal basis $U_r \in \mathbb{R}^{d \times r}$ with $U_r^\top U_r = I_r$, the orthogonal projection of the data onto $\mathrm{span}(U_r)$ and its reconstruction are

$$\hat{X}_r = X U_r U_r^\top.$$

The reconstruction error is defined as

$$E(U_r) = \|X - \hat{X}_r\|_F^2 = \|X - X U_r U_r^\top\|_F^2.$$

Our goal is to show that the PCA basis U_r gives the smallest possible reconstruction error among all orthonormal bases.

2.2 Proof (Projection ⇒ Trace Maximization ⇒ PCA)

Because $P_{U_r} = U_r U_r^\top$ is an orthogonal projector, we have

$$\|X\|_F^2 = \|X P_{U_r}\|_F^2 + \|X(I - P_{U_r})\|_F^2 \quad \Rightarrow \quad E(U_r) = \|X\|_F^2 - \|X U_r U_r^\top\|_F^2.$$

Thus, minimizing $E(U_r)$ is the same as maximizing $\|X U_r\|_F^2$. Since

$$\|X U_r\|_F^2 = \mathrm{tr}(U_r^\top X^\top X U_r),$$

we want to find U_r that makes this trace as large as possible.

Let $X^\top X = V \Lambda V^\top$ be the eigen-decomposition with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. From the properties of symmetric matrices, the maximum trace is obtained when U_r is made of the first r eigenvectors of $X^\top X$:

$$U_r = [v_1, \ldots, v_r].$$

In that case,

$$\mathrm{tr}(U_r^\top X^\top X U_r) = \sum_{i=1}^{r} \lambda_i.$$

Therefore, projecting onto these eigenvectors (the PCA subspace) gives the smallest possible reconstruction error.

2.3 Conclusion

PCA provides the orthogonal basis that minimizes the reconstruction error, and the minimum possible error is

$$E_{\min}(r) = \sum_{i=r+1}^{d} \sigma_i^2,$$

where σ_i are the singular values of X. This result matches what we observed in Question 1, where the reconstruction error decreases as r increases.

3. Question 3: Quadratic Form Minimization

3.1 Objective

In this question, we are given a real-valued data vector x (from `xb.mat`). For a chosen length N, we want to find a vector $h \in \mathbb{R}^N$ with unit norm $\|h\|_2 = 1$ that minimizes a quadratic form $h^\top Ah$, where A is a symmetric matrix constructed from the data.

3.2 Method

We first build a symmetric matrix A using inner products between shifted copies of x, so that each entry represents how similar two parts of the vector are. This makes A symmetric by construction. The minimization problem can be written as:

$$\min_{\|h\|_2=1} h^\top Ah.$$

From linear algebra theory, for a symmetric matrix A, this value is smallest when h is the eigenvector that corresponds to the smallest eigenvalue of A. So we just compute the eigenvalues and take the eigenvector associated with the smallest one.

I use a Python script to load the vector, centers it by removing its mean, and constructs A for several lengths $N = 8, 12, 16, 24$. Then it computes the smallest eigenvalue and its eigenvector. The results (smallest eigenvalues) are summarized in a simple table.

Table 2: Results for different N values.

N	Smallest eigenvalue of A
8	409.53
12	318.93
16	247.38
24	170.04

3.3 Conclusion

By constructing a symmetric matrix A and solving $\min_{\|h\|_2=1} h^\top Ah$, we found that the smallest eigenvalue and its eigenvector provide the solution.

As N increases, the smallest eigenvalue of A becomes smaller. This makes sense because with a larger N, the matrix captures more of the relationships inside x, and the minimization has more directions to choose from. The eigenvector corresponding to this smallest eigenvalue represents the direction where the quadratic form $h^\top Ah$ reaches its minimum.

References

- J. Reilly, *Fundamentals of Linear Algebra for Signal Processing*, Lecture Notes, ECE 712, McMaster University.