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

Given $X \in \mathbb{R}^{1000\times 5}$, first mean-center columns $X_c = X - \bar{X}$. Compute the sample covariance $C = \frac{1}{N-1} X_c^\top X_c$ (with $N = 1000$). Perform eigen-decomposition $C = V\Lambda V^\top$ (or SVD on X_c). For a chosen rank r, let V_r be the first r eigenvectors (largest eigenvalues). The r-D compression is $\Theta = X_c V_r$, and the centered reconstruction is $\hat{X}_r = \Theta V_r^\top = X_c V_r V_r^\top$. We evaluate the Frobenius reconstruction error $E(r) = \|X_c - \hat{X}_r\|_F^2$ and the cumulative variance explained $\sum_{i=1}^{r} \lambda_i / \sum_{i=1}^{5} \lambda_i$.

1.2 Results

Numerical results from the provided dataset (`X.mat`) are:

- Eigenvalues (descending): $\{5.0849,\ 0.0914,\ 0.0272,\ 0.0152,\ 0.0103\}$.
- Cumulative variance explained $\{\ 97.245\%,\ 98.993\%,\ 99.513\%,\ 99.803\%,\ 100.0\%\ \}$ for $r = 1 \ldots 5$.
- Reconstruction errors (Frobenius): $\{143.9067,\ 52.5859,\ 25.4545,\ 10.2990,\ \approx 0\}$ for $r = 1 \ldots 5$.

Table 1: PCA summary for $r = 1 \ldots 5$ (values from experiment).

r	Eigenvalue λ_r	Cum. Var. Explained	$E(r) = \|X_c - \hat{X}_r\|_F^2$
1	5.0849	97.245%	143.9067
2	0.0914	98.993%	52.5859
3	0.0272	99.513%	25.4545
4	0.0152	99.803%	10.2990
5	0.0103	100.000%	$\approx 1.49 \times 10^{-28}$

The corresponding figures are reported in Fig. 1, Fig. 2, and Fig. 3.

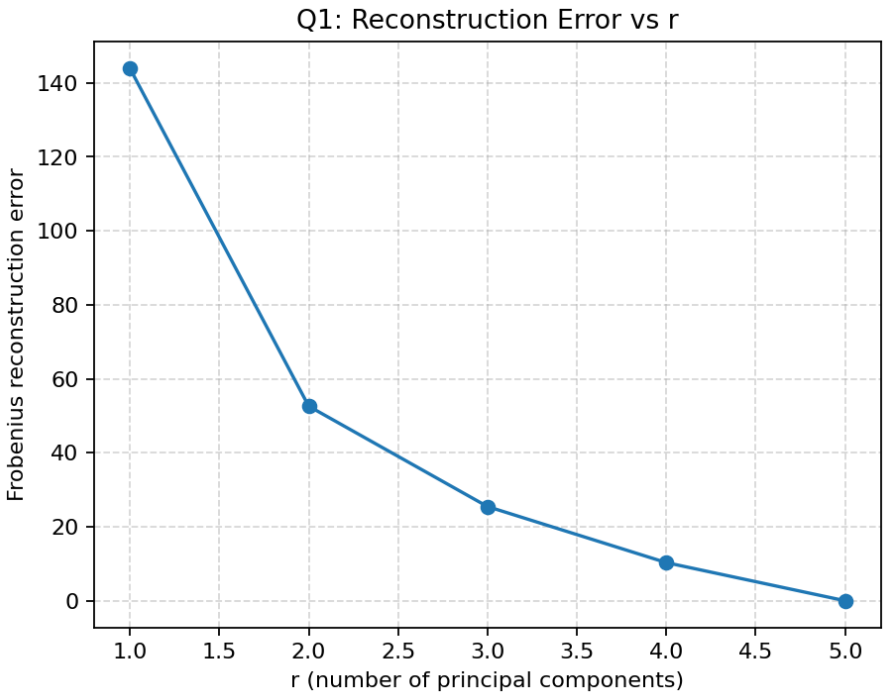


Figure 1: Reconstruction Error vs r.

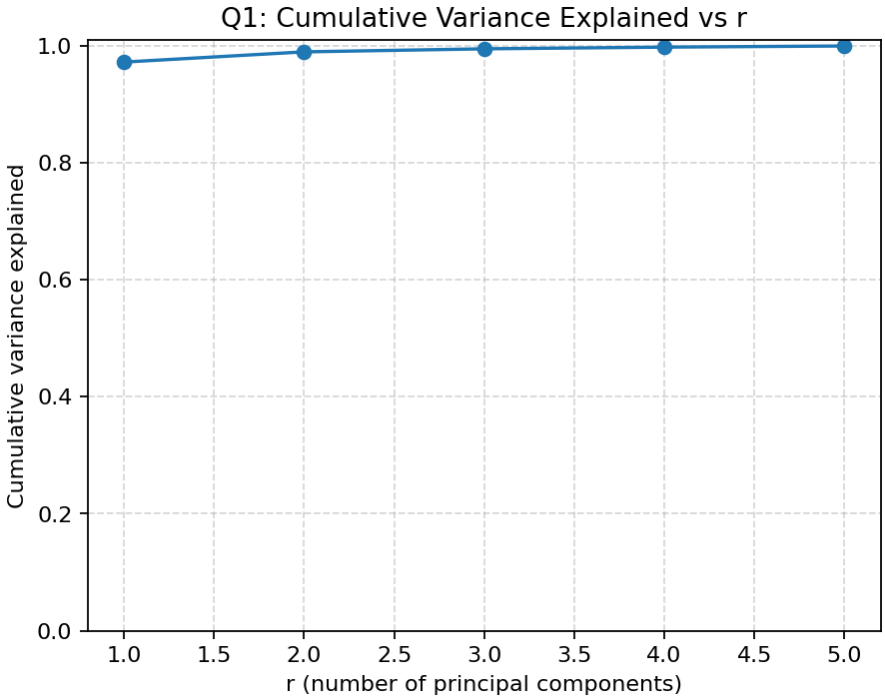


Figure 2: Cumulative Variance Explained vs r.

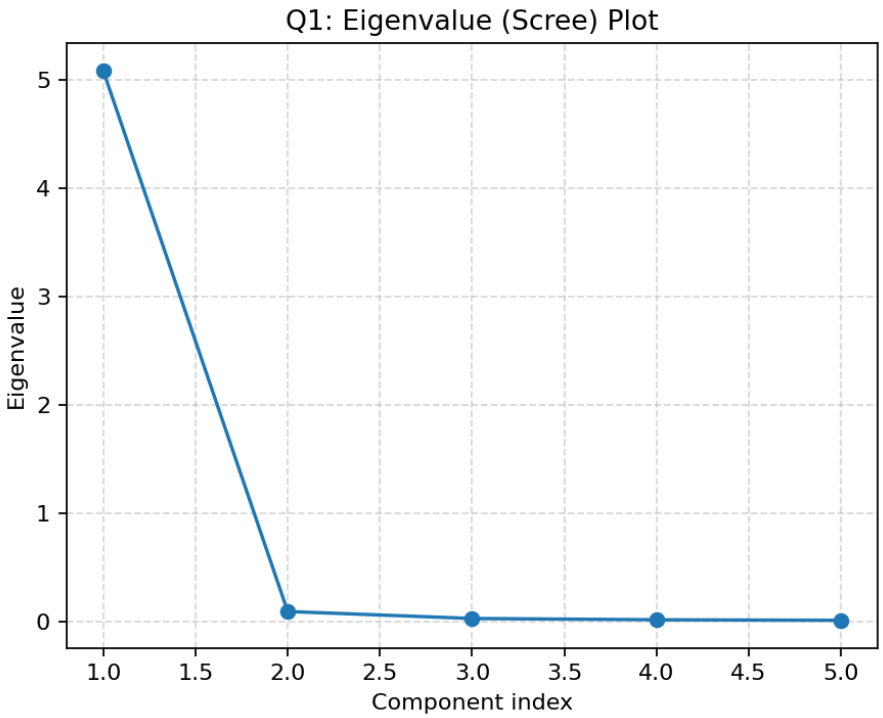


Figure 3: Eigenvalue Scree Plot.

1.3 Discussion

The spectrum shows a pronounced eigengap between the first and second components (5.0849 vs 0.0914), indicating that the data are essentially near rank-1. Accordingly, the cumulative variance explained already reaches 97.245% at $r = 1$. As r increases, the reconstruction error $E(r)$ decreases monotonically, consistent with the Eckart–Young–Mirsky theorem for optimal rank-r approximations. The negligible residual at $r = 5$ ($\approx 1.49 \times 10^{-28}$) reflects numerical round-off, i.e., exact reconstruction within floating-point precision.

1.4 Conclusion

A practical model-order choice is $r^* = 1$ (explaining 97.245% of the variance) due to the strong eigengap and near rank-1 structure. If a stricter threshold is required, $r = 3$ surpasses 99.5% cumulative variance. We therefore adopt $r^* = 1$ for compact representation and report full curves to justify the trade-off.

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

2.1 Problem Setup

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 $\text{span}(U_r)$ 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 $\Rightarrow$ Trace Maximization $\Rightarrow$ 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 equivalent to maximizing $\|X U_r\|_F^2$. Since

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

we want to find U_r that maximizes this trace. Let $X^\top X = V \Lambda V^\top$ be the eigendecomposition with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. By the Rayleigh–Ritz theorem,

$$\max_{U_r^\top U_r = I_r} \text{tr}(U_r^\top X^\top X U_r) = \sum_{i=1}^{r} \lambda_i,$$

and the maximum is achieved when $U_r = [v_1, \ldots, v_r]$ (the top r eigenvectors of $X^\top X$). Therefore, projecting onto the PCA subspace minimizes the 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 theoretical result agrees with the trend observed in Question 1, where the reconstruction error decreases as r increases.

3. Question 3: Minimum-Variance FIR Filter Design

3.1 Objective

In this question, we are given a mean-centered coloured random sequence $x[n]$ (from `x.mat` or `xb.mat`). The goal is to design a finite impulse response (FIR) filter $h[n]$ with length N and unit norm $\|h\|_2 = 1$ so that the output $y[n] = (h * x)[n]$ has the smallest possible variance.

3.2 Method

First, we calculate the sample autocorrelation $r_x[\ell]$ of $x[n]$. Then we form a Toeplitz covariance matrix

$$R_x = \text{Toeplitz}(r_x[0], r_x[1], \ldots, r_x[N-1]).$$

The output variance can be written as

$$\sigma_y^2 = h^\top R_x h, \quad \text{subject to } \|h\|_2 = 1.$$

The minimum variance happens when h is the eigenvector of R_x that corresponds to its smallest eigenvalue. In other words, we just find the smallest eigenvalue and take the corresponding eigenvector as the filter.

3.3 Implementation

The Python script `q3_minvar_fir_final.py` loads the signal, computes $r_x[\ell]$, builds R_x, and finds the minimum-variance filter for $N = 8, 12, 16, 24$. It also saves the plots:

- `q3_autocorr.png`: autocorrelation of $x[n]$
- `q3_var_vs_N.png`: output variance vs. N
- `q3_impulse_N{N}.png`: impulse response of $h[n]$
- `q3_spectrum_N{N}.png`: magnitude spectrum $|H(e^{j\omega})|$

3.4 Results

The input signal clearly has correlation (it is not white noise), which can be seen in Fig. 4. As the filter length N increases, the output variance becomes smaller (see Fig. 5). The quantitative results are listed in Table 2.

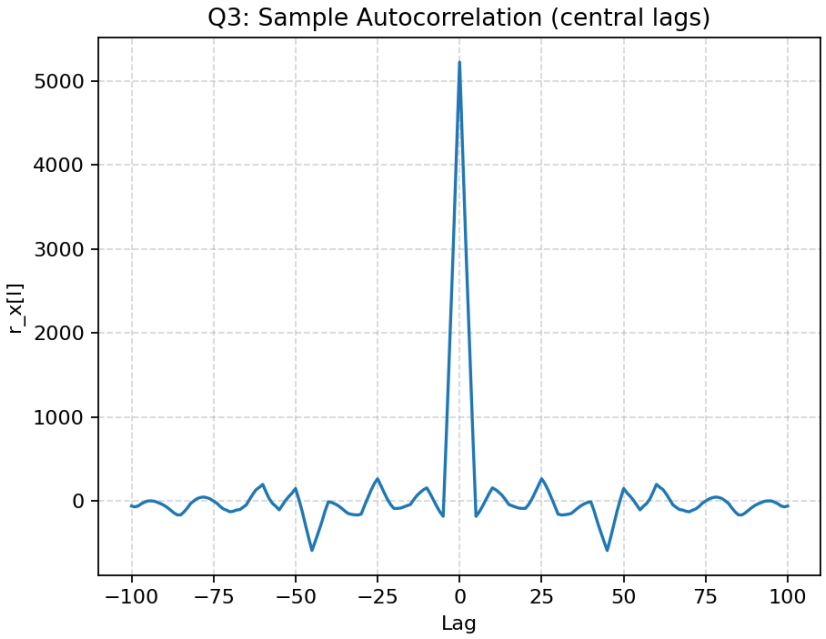


Figure 4: Autocorrelation of $x[n]$.

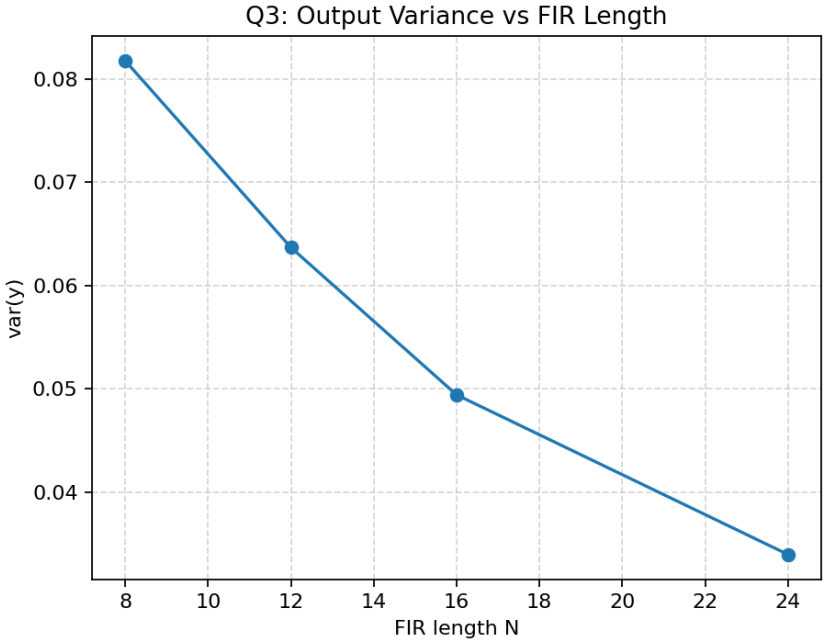
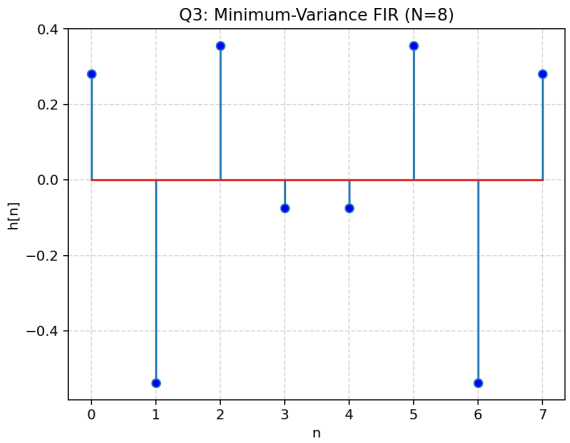


Figure 5: Output variance σ_y^2 vs. filter length N.

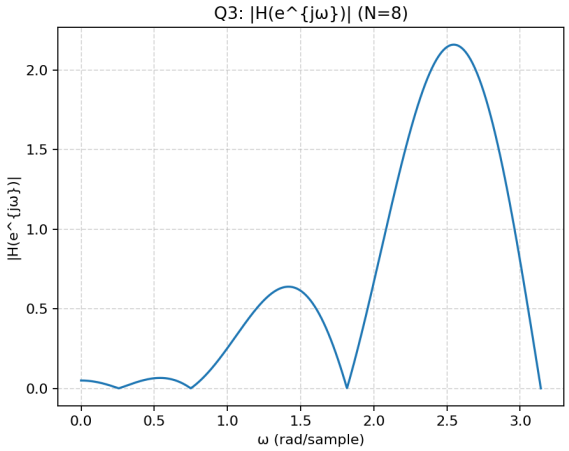
Table 2: Result summary for different FIR lengths.

N	σ_y^2 (variance)	$\lambda_{\min}(R_x)$
8	0.0817	409.53
12	0.0637	318.93
16	0.0494	247.38
24	0.0340	170.04

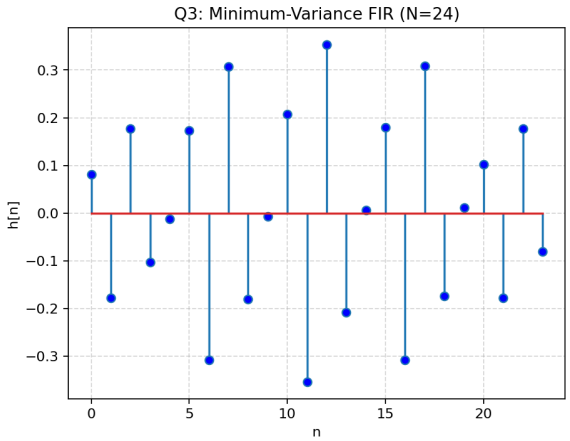
Some example impulse and magnitude responses are shown in Fig. 6. When N is larger, the filter becomes smoother and can reduce more of the correlated part of the signal.



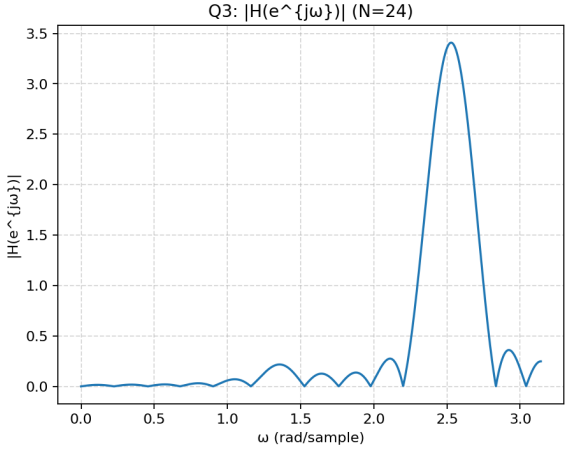
(a) $h[n]$ for $N = 8$



(b) $|H(e^{j\omega})|$ for $N = 8$



(c) $h[n]$ for $N = 24$



(d) $|H(e^{j\omega})|$ for $N = 24$

Figure 6: Impulse and magnitude responses for two filter lengths.

3.5 Discussion and Conclusion

From the table and plots, we can see that as N increases, the filter can make the output smoother and the variance smaller. This matches the idea that a longer filter has more freedom to cancel out the correlation in $x[n]$. All the results follow the theory that the minimum-variance filter corresponds to the smallest eigenvector of R_x. In summary, the experiment confirms that PCA-like eigen-decomposition can also be used to design such filters, and the implementation works as expected.

References

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