Q18 - HW2: Pattern Recognition

Vahid Maleki
Student ID: 40313004

October 18, 2025

1   Question 18

Given the following **covariance matrix**:

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 3 & 0 & 0 & 0 & 5 \end{bmatrix}$$

To ensure that **less than 30% of the total information (variance)** is lost, what is the **minimum number of eigenvectors** that must be retained in the **PCA (Principal Component Analysis)** algorithm?

1.1   Solution

To determine the minimum number of eigenvectors to retain in PCA while ensuring that less than 30% of the total variance is lost, we need to compute the eigenvalues of the covariance matrix Σ, calculate the total variance, and find the smallest number of eigenvalues (in descending order) that account for at least 70% of the total variance.

Step 1: Find the Eigenvalues

The covariance matrix Σ is block-diagonal with a 2×2 block in rows/columns 2 and 6, and the other diagonal entries are eigenvalues directly. The block is:

$$\begin{bmatrix} 4 & 3 \\ 3 & 5 \end{bmatrix}$$

Compute the eigenvalues of this block by solving:

$$\det \left(\begin{bmatrix} 4 - \lambda & 3 \\ 3 & 5 - \lambda \end{bmatrix} \right) = 0$$

$$(4 - \lambda)(5 - \lambda) - 9 = \lambda^2 - 9\lambda + 20 - 9 = \lambda^2 - 9\lambda + 11 = 0$$

$$\lambda = \frac{9 \pm \sqrt{81 - 44}}{2} = \frac{9 \pm \sqrt{37}}{2}$$

Since $\sqrt{37} \approx 6.08$, the eigenvalues are:

$$\lambda_1 \approx \frac{9 + 6.08}{2} \approx 7.54, \quad \lambda_2 \approx \frac{9 - 6.08}{2} \approx 1.46$$

The other diagonal entries correspond to eigenvalues for the remaining dimensions:
- Row 1: 1 - Row 3: 1 - Row 4: 2 - Row 5: 3
Thus, the full set of eigenvalues is:

$$1, 1, 2, 3, 7.54, 1.46$$

Step 2: Sort Eigenvalues

Sort the eigenvalues in descending order:

$$7.54, 3, 2, 1.46, 1, 1$$

Step 3: Compute Total Variance
The total variance is the sum of all eigenvalues:

$$7.54 + 3 + 2 + 1.46 + 1 + 1 = 16$$

Step 4: Determine Minimum Eigenvectors for 70% Variance
To ensure that less than 30% of the variance is lost, we need to retain at least 70% of the total variance:

$$70\% \text{ of } 16 = 0.7 \times 16 = 11.2$$

Add the eigenvalues in descending order until the cumulative sum is at least 11.2:
- First eigenvalue: 7.54 - Second eigenvalue: $7.54 + 3 = 10.54$ - Third eigenvalue: $10.54 + 2 = 12.54$
After three eigenvalues, the cumulative variance is 12.54, which exceeds 11.2.

Step 5: Conclusion
To ensure that at least 70% of the total variance is retained (i.e., less than 30% is lost), we need to keep the top three eigenvectors.

$$\boxed{3}$$