Case Study 1:

Analyzing Market Trends with PCA

Background: You are a financial analyst at a consulting firm tasked with understanding the underlying market trends that influence the returns of a set of financial assets. Your goal is to identify the key factors that drive these returns and to provide insights into how these factors can be used for strategic decision-making.

Objective: Apply Principal Component Analysis (PCA) using Singular Value Decomposition (SVD) to a simplified dataset of asset returns to identify the principal components. Use these components to interpret the main market trends.

Data: Consider a simplified dataset with the following returns for three assets (A, B, and C) over three time periods (T1, T2, T3):

	Asset A	Asset B	Asset C
T1	0.10	0.20	0.15
T2	0.05	0.10	0.05
T3	0.15	0.25	0.20

Tasks:

1. Data Statistics:
    a. Calculate the mean return for each asset.
    b. Center the data by subtracting the mean return from each asset's returns.
    c. Compute the covariance matrix of the centered data.
2. Singular Value Decomposition (SVD):
    d. Perform SVD on the covariance matrix to find the eigenvalues and eigenvectors.
    e. Identify the principal components based on the eigenvectors.
    f. Calculate the proportion of total variance explained by each principal component.
3. Interpretation:
    g. Discuss the significance of the principal components in terms of market trends.
    h. Explain how these components can be used to understand the behavior of the assets.
    i. Provide recommendations on how the firm can use these insights for investment strategy or risk management.

Case Study 2:

Suppose that a set of (n) vectors $v_1, v_2, v_3, \dots, v_n \in R^m$ and a vector $(w \in R^m)$ is given. We seek reals $(x_1, x_2, \dots, x_n)$ that minimizes the following quantity:

$$L = ||\left(\sum_{i=1}^{n} x_i v_i\right) - w\, ||$$

Let (A) be an ($m \times n$) matrix whose (i) - th column is the vector (v_i) and let $(\vec{x})$ be an (n)-dimensional column vector whose (i) - th coordinate is (x_i). Then the above can be re-written as minimizing

$$L = |A\,\vec{x} - w|$$

(a) Show that the problem can be directly solved if A is a diagonal matrix.

(b) Suppose that the SVD of A is given as $(A = U\Sigma V^T)$. Then

$$A\vec{x} - w = U\Sigma V^T\vec{x} - w = U\Sigma V^T\vec{x} - UU^T w = U(\Sigma V^T\vec{x} - U^T w)$$

Use this equation to explain how the general problem reduces to the special case when A is a diagonal matrix.

(c) Suppose that we are given n data points $(x_1, y_1), \dots, (x_n, y_n)$ in R^2. Our goal is to fnd a linear equation $y = ax + b$ that 'best fits' the data. Formally, we would like to fnd reals a and b that minimizes

$$\sum_{i=1}^{n} (ax_i + b - y_i)^2$$

Explain how this problem is a special case of the problem given above. 2

(d) Extend (c) to describe the problem of fnding a quadratic equation $y = ax^2 + bx + c$ that 'best fits' the data points $(x_1, y_1), \dots, (x_n, y_n) \in R^2$ as a special case of the problem given above