Q6 - HW2: Pattern Recognition

Vahid Maleki
Student ID: 40313004

October 18, 2025

1 Question 6

Let $\mathbf{x}$ be a random vector in an n-dimensional space with zero mean and covariance matrix Σ. In general, it is possible to find an orthonormal (unitary) transformation matrix Φ, and define an n-dimensional variable $\mathbf{z}$ as follows:

$$\mathbf{z} = \Phi^T \mathbf{x}$$

In the above relation, the transformation matrix Φ consists of the eigenvectors of the covariance matrix Σ, and we have:

$$\Phi^T \Sigma \Phi = \Lambda$$

where Λ is a diagonal matrix containing the eigenvalues of Σ, arranged in descending order.

1.1 (a)

Since some eigenvalues are extremely small, dimensionality reduction is justified. Define a reduced transformation matrix Φ' using the eigenvectors of Σ (i.e., ϕ_i) such that the feature space is reduced to n' dimensions ($n' < n$). Express the resulting n'-dimensional variable $\mathbf{z}'$ in terms of $\mathbf{x}$ and Φ'.

1.1.1 Solution

To reduce the dimensionality of the feature space from n to n' (where $n' < n$), we select the n' eigenvectors of the covariance matrix Σ that correspond to the n' largest eigenvalues. Since the eigenvalues λ_i are arranged in descending order, we choose the first n' eigenvectors, $\phi_1, \phi_2, \ldots, \phi_{n'}$.

The reduced transformation matrix, denoted as Φ', is an $n \times n'$ matrix whose columns are these selected eigenvectors.

$$\Phi' = [\phi_1, \phi_2, \ldots, \phi_{n'}]$$

The resulting n'-dimensional feature variable, denoted as $\mathbf{z}'$, is obtained by projecting the original n-dimensional vector $\mathbf{x}$ onto the subspace spanned by the columns of Φ'. This is achieved by the following transformation:

$$\mathbf{z}' = (\Phi')^T \mathbf{x}$$

Here, $\mathbf{z}'$ is an $n' \times 1$ column vector representing the original data in the reduced-dimensional space.

1.2 (b)

Show that the expected value of the reduced n'-dimensional vector $\mathbf{z}'$ is zero:

$$E\{\mathbf{z}'\} = \mathbf{0}$$

1.2.1 Solution

We want to show that the expected value of the reduced vector $\mathbf{z}'$ is a zero vector. We are given that the original vector $\mathbf{x}$ has a zero mean, so $E\{\mathbf{x}\} = \mathbf{0}$.

The expected value of $\mathbf{z}'$ is:

$$E\{\mathbf{z}'\} = E\{(\Phi')^T \mathbf{x}\}$$

Since the expectation operator $E\{\cdot\}$ is linear and the transformation matrix Φ' is a constant matrix of coefficients, we can move it outside the expectation.

$$E\{\mathbf{z}'\} = (\Phi')^T E\{\mathbf{x}\}$$

Substituting the given condition $E\{\mathbf{x}\} = \mathbf{0}$:

$$E\{\mathbf{z}'\} = (\Phi')^T \mathbf{0} = \mathbf{0}$$

This shows that the resulting n'-dimensional vector $\mathbf{z}'$ also has a zero mean.

1.3 (c)

Show that the variance of the i-th component of $\mathbf{z}'$ (for $i \leq n'$) equals λ_i, the i-th eigenvalue of Σ. That is, prove that:

$$E\{(z_i')^2\} = \lambda_i$$

Hint: ($z_i' = \phi_i^T \mathbf{x}$)

1.3.1 Solution

We need to prove that the variance of the i-th component of $\mathbf{z}'$ is equal to the i-th eigenvalue of Σ, i.e., $E\{(z_i')^2\} = \lambda_i$. The i-th component of $\mathbf{z}'$ is given by $z_i' = \phi_i^T \mathbf{x}$, where ϕ_i is the i-th eigenvector.

The variance of z_i' is $\mathrm{Var}(z_i') = E\{(z_i' - E\{z_i'\})^2\}$. From part (b), we know that $E\{z_i'\} = 0$, so the variance simplifies to $E\{(z_i')^2\}$.

Let's compute $E\{(z_i')^2\}$:

$$E\{(z_i')^2\} = E\{(\phi_i^T \mathbf{x})^2\} = E\{(\phi_i^T \mathbf{x})(\phi_i^T \mathbf{x})\}$$

Since $\phi_i^T \mathbf{x}$ is a scalar, we can write its square as $(\phi_i^T \mathbf{x})^T (\phi_i^T \mathbf{x}) = \mathbf{x}^T \phi_i \phi_i^T \mathbf{x}$. However, a more direct approach is:

$$E\{(z_i')^2\} = E\{\phi_i^T \mathbf{x}\mathbf{x}^T \phi_i\}$$

Because ϕ_i is a constant vector, we can move it outside the expectation:

$$E\{(z_i')^2\} = \phi_i^T E\{\mathbf{x}\mathbf{x}^T\} \phi_i$$

The covariance matrix Σ is defined as $\Sigma = E\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^T\}$. Since $E\{\mathbf{x}\} = \mathbf{0}$, this simplifies to $\Sigma = E\{\mathbf{x}\mathbf{x}^T\}$. Substituting this into our equation gives:

$$E\{(z_i')^2\} = \phi_i^T \Sigma \phi_i$$

By the definition of an eigenvector, $\Sigma \phi_i = \lambda_i \phi_i$. Substituting this relationship:

$$E\{(z_i')^2\} = \phi_i^T (\lambda_i \phi_i) = \lambda_i (\phi_i^T \phi_i)$$

The matrix Φ is an orthonormal transformation matrix, which means its column vectors ϕ_i are orthonormal. Therefore, the inner product $\phi_i^T \phi_i = 1$.

$$E\{(z_i')^2\} = \lambda_i (1) = \lambda_i$$

This proves that the variance of the i-th component of the transformed vector $\mathbf{z}'$ is equal to the i-th largest eigenvalue of the covariance matrix Σ.

1.4 (d)

Given $\mathbf{z}'$, and the matrices Φ and Φ', write the relations used to reconstruct the original vector in the n-dimensional space (denoted by $\mathbf{x}'$).

Hint: Add zeros to the end of $\mathbf{z}'$ to expand its dimension from n' to n, and call the resulting vector $\mathbf{z}'_{\mathbf{e}}$. Then apply the inverse transformation to $\mathbf{z}'_{\mathbf{e}}$ to reconstruct $\mathbf{x}'$.

$$\mathbf{z}'_{\mathbf{e}} = \begin{bmatrix} \mathbf{z}' \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{x}' = \Phi \mathbf{z}'_{\mathbf{e}}$$

1.4.1 Solution

To reconstruct the original vector $\mathbf{x}$ from the reduced n'-dimensional vector $\mathbf{z}'$, we first need to map $\mathbf{z}'$ back to the n-dimensional space. This is done by creating an n-dimensional vector $\mathbf{z}'_{\mathbf{e}}$ by appending $n - n'$ zeros to $\mathbf{z}'$.

$$\mathbf{z}'_{\mathbf{e}} = \begin{bmatrix} \mathbf{z}' \\ \mathbf{0} \end{bmatrix}$$

Here, $\mathbf{0}$ is a zero vector of size $(n - n') \times 1$.

The original transformation from $\mathbf{x}$ to the full n-dimensional vector $\mathbf{z}$ is $\mathbf{z} = \Phi^T \mathbf{x}$. Since Φ is an orthonormal matrix, its inverse is its transpose, $\Phi^{-1} = \Phi^T$. Therefore, the inverse transformation to recover $\mathbf{x}$ from $\mathbf{z}$ is $\mathbf{x} = \Phi \mathbf{z}$.

To obtain the reconstructed vector $\mathbf{x}'$ from the padded vector $\mathbf{z}'_{\mathbf{e}}$, we apply this inverse transformation:

$$\mathbf{x}' = \Phi \mathbf{z}'_{\mathbf{e}}$$

This projects the reduced representation back into the original n-dimensional space. Note that $\mathbf{x}'$ is an approximation of $\mathbf{x}$ because information was lost during the dimensionality reduction.

1.5 (e)

The mean squared error (MSE) caused by dimensionality reduction — i.e., by reducing $\mathbf{x}$ to the n'-dimensional vector $\mathbf{z}'$ and then reconstructing it as $\mathbf{x}'$ in the original n-dimensional space — is defined as:

$$e_{ms} = E\{|\mathbf{x} - \mathbf{x}'|^2\}$$

Show that this error equals:

$$e_{ms} = \sum_{i=n'+1}^{n} \lambda_i$$

where λ_i are the eigenvalues of the covariance matrix Σ corresponding to the eliminated dimensions.

Answer:

$$\begin{aligned} e_{ms} &= E\{|\mathbf{x} - \mathbf{x}'|^2\} = E\{(\mathbf{x} - \mathbf{x}')^T(\mathbf{x} - \mathbf{x}')\} \\ &= E\{(\Phi\mathbf{z} - \Phi\mathbf{z}'_{\mathbf{e}})^T(\Phi\mathbf{z} - \Phi\mathbf{z}'_{\mathbf{e}})\} \\ &= \Phi^T E\{(\mathbf{z} - \mathbf{z}'_{\mathbf{e}})^T(\mathbf{z} - \mathbf{z}'_{\mathbf{e}})\}\Phi = \sum_{i=n'+1}^{n} \lambda_i \end{aligned}$$

1.5.1 Solution

The mean squared error (MSE) is defined as $e_{ms} = E\{|\mathbf{x} - \mathbf{x}'|^2\}$. We can express this as the expected value of the squared Euclidean norm, which is $E\{(\mathbf{x} - \mathbf{x}')^T(\mathbf{x} - \mathbf{x}')\}$.

We substitute $\mathbf{x} = \Phi\mathbf{z}$ and $\mathbf{x}' = \Phi\mathbf{z}'_{\mathbf{e}}$:

$$e_{ms} = E\{(\Phi\mathbf{z} - \Phi\mathbf{z}'_{\mathbf{e}})^T(\Phi\mathbf{z} - \Phi\mathbf{z}'_{\mathbf{e}})\}$$

Factor out Φ from the expression:

$$e_{ms} = E\{(\Phi(\mathbf{z} - \mathbf{z}'_{\mathbf{e}}))^T(\Phi(\mathbf{z} - \mathbf{z}'_{\mathbf{e}}))\}$$

Using the transpose property $(AB)^T = B^T A^T$:

$$e_{ms} = E\{(\mathbf{z} - \mathbf{z}'_{\mathbf{e}})^T \Phi^T \Phi (\mathbf{z} - \mathbf{z}'_{\mathbf{e}})\}$$

Since Φ is orthonormal, $\Phi^T \Phi = I$, where I is the identity matrix.

$$e_{ms} = E\{(\mathbf{z} - \mathbf{z}'_{\mathbf{e}})^T I (\mathbf{z} - \mathbf{z}'_{\mathbf{e}})\} = E\{(\mathbf{z} - \mathbf{z}'_{\mathbf{e}})^T (\mathbf{z} - \mathbf{z}'_{\mathbf{e}})\}$$

This is the expected squared norm of the vector $\mathbf{z} - \mathbf{z}'_{\mathbf{e}}$. Let's examine this vector:

$$\mathbf{z} - \mathbf{z}'_{\mathbf{e}} = \begin{bmatrix} z_1 \\ \vdots \\ z_{n'} \\ z_{n'+1} \\ \vdots \\ z_n \end{bmatrix} - \begin{bmatrix} z_1 \\ \vdots \\ z_{n'} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ z_{n'+1} \\ \vdots \\ z_n \end{bmatrix}$$

The squared norm of this vector is the sum of the squares of its components:

$$(\mathbf{z} - \mathbf{z}'_{\mathbf{e}})^T (\mathbf{z} - \mathbf{z}'_{\mathbf{e}}) = \sum_{i=n'+1}^{n} (z_i)^2$$

Now, we take the expectation:

$$e_{ms} = E\left\{ \sum_{i=n'+1}^{n} (z_i)^2 \right\}$$

By the linearity of expectation:

$$e_{ms} = \sum_{i=n'+1}^{n} E\{(z_i)^2\}$$

From part (c), we know that $E\{(z_i)^2\} = \lambda_i$. Therefore:

$$e_{ms} = \sum_{i=n'+1}^{n} \lambda_i$$

The mean squared error from dimensionality reduction is the sum of the eigenvalues corresponding to the dimensions that were discarded.

1.6 (f)

Based on the obtained error, propose a criterion for selecting the reduced dimension n'.

1.6.1 Solution

The result from part (e) provides a direct way to quantify the error introduced by dimensionality reduction. The total variance of the original data $\mathbf{x}$ is the trace of its covariance matrix, which is also equal to the sum of all its eigenvalues:

$$\text{Total Variance} = \text{Tr}(\Sigma) = \sum_{i=1}^{n} \lambda_i$$

The MSE, $e_{ms} = \sum_{i=n'+1}^{n} \lambda_i$, represents the amount of variance (information) lost during the reduction. The amount of variance retained is $\sum_{i=1}^{n'} \lambda_i$.

A common criterion for choosing the reduced dimension n' is to retain a certain percentage of the total variance. For instance, we might want to keep 95% or 99% of the original variance. This leads to the following criterion:

Choose the smallest integer n' such that the ratio of the retained variance to the total variance is greater than or equal to a specified threshold T (e.g., $T = 0.95$).

$$\frac{\sum_{i=1}^{n'} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \geq T$$

To apply this criterion, one would compute the eigenvalues of the covariance matrix Σ, sort them in descending order, and then calculate the cumulative sum of these eigenvalues. The value of n' is chosen as the point where this cumulative sum first exceeds the desired percentage of the total sum.