

PCA Explanation



Principal Components Analysis---explain in the format of writtern in exam

Here's a concise **exam-style explanation** of **Principal Components Analysis (PCA)** suitable for written answers:

Principal Components Analysis (PCA)

Definition:

Principal Components Analysis (PCA) is a statistical technique used for **dimensionality reduction**. It transforms a large set of variables into a smaller one that still contains most of the information in the original dataset.

Objectives of PCA:

- To reduce the number of variables (dimensions) in a dataset.
- To retain as much variance (information) as possible.
- To identify new uncorrelated variables (called **principal components**) that are linear combinations of the original variables.

Steps in PCA:

1. Standardization:

Standardize the dataset to have a mean of 0 and standard deviation of 1, especially if variables are on different scales.

2. Covariance Matrix Computation:

Calculate the covariance matrix to understand the relationships between variables.

3. Eigenvalues and Eigenvectors:

Compute the eigenvalues and eigenvectors of the covariance matrix.

- **Eigenvectors** determine the direction of the new feature space.
- **Eigenvalues** determine the magnitude (i.e., how much variance is captured).

4. Sort and Select Principal Components:

- Sort eigenvalues in descending order.
- Select the top **k** eigenvectors (with the highest eigenvalues) to form a **feature vector**.

5. Form the New Dataset:

Multiply the original dataset by the feature vector to get the new reduced dataset.

Key Points:

- The first principal component captures the maximum variance.
 - Principal components are orthogonal (uncorrelated).
 - PCA is unsupervised (does not use class labels).
 - Useful in noise reduction, data visualization, and improving algorithm performance.
-

Applications:

- Image compression
 - Pattern recognition
 - Finance (e.g., reducing correlated financial indicators)
 - Genetics (e.g., genome-wide association studies)
-

Let me know if you want a version with a diagram or example included too!