**Principal Component Analysis (PCA) in Machine Learning**

**1. Introduction to PCA**

Principal Component Analysis (PCA) is a dimensionality reduction technique commonly used in Machine Learning and statistics. It helps to reduce the number of features while preserving as much variance (information) as possible.

**2. Why Use PCA?**

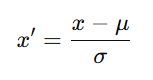
* **High-dimensional data** can lead to issues such as the **curse of dimensionality**, making models slow and prone to overfitting.
* PCA helps in **visualization** by reducing dimensions to 2D or 3D.
* It speeds up machine learning algorithms by reducing the number of input features.
* It removes redundancy and correlation among features.

**3. How PCA Works (Step-by-Step Process)**

**Step 1: Standardization of Data**

Since PCA is affected by the scale of data, it is necessary to standardize or normalize it.

* Formula for standardization:

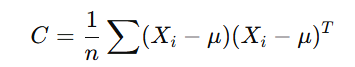


Where:

* + x = original feature value
  + μ = mean of the feature
  + σ = standard deviation of the feature

**Step 2: Compute the Covariance Matrix**

* Covariance measures the relationship between two variables.
* Covariance matrix:



* Where Xi ​ is the dataset matrix.

**Step 3: Compute Eigenvalues and Eigenvectors**

* Eigenvectors determine the direction of the new feature space.
* Eigenvalues indicate the magnitude (importance) of each eigenvector.
* These are calculated by solving



Where C is the covariance matrix, λ is an eigenvalue, and V is an eigenvector.

**Step 4: Select Principal Components**

* Sort eigenvalues in decreasing order and choose the top **k** components based on variance explained.
* The sum of eigenvalues represents the total variance.

**Step 5: Transform Data to New Feature Space**

* Multiply original dataset by the selected eigenvectors to get reduced features.

**4. Choosing the Right Number of Components**

* The explained variance ratio tells how much variance each principal component retains.
* A common method is to select the number of components where the cumulative explained variance is around **95%**.
* **Scree Plot**: A plot of eigenvalues helps in deciding the number of principal components.

**5. Advantages of PCA**

* Reduces computational cost.
* Removes multicollinearity among features.
* Enhances visualization in lower dimensions.

**6. Limitations of PCA**

* May lead to loss of interpretability.
* Assumes that principal components are linear combinations of original features.
* PCA is sensitive to outliers.

**7. Applications of PCA**

* Image Compression
* Face Recognition
* Feature Extraction in Machine Learning
* Noise Filtering

**Conclusion**

PCA is a powerful tool for dimensionality reduction that balances variance retention and computational efficiency. However, it should be applied carefully, especially when interpretability is essential.