PCA Using Eigenvalue

PCA and Eigenvalues

1. **Covariance Matrix:** PCA starts with the covariance matrix, which summarizes the relationships between features in your data.
2. **Eigenanalysis:** We perform eigenanalysis on this matrix. This mathematical operation reveals the eigenvectors and eigenvalues that hold the key to dimensionality reduction.
3. **Eigenvalues for Dimensionality Reduction:** PCA prioritizes eigenvectors based on their corresponding eigenvalues. Here's how:
   - The eigenvector with the **largest eigenvalue** captures the direction of **maximum variance** in the data. This is considered the most informative direction.
   - Subsequent eigenvectors, with progressively **smaller eigenvalues**, represent directions of decreasing importance in terms of variance.

Step-By-Step Explanation of PCA (Principal Component Analysis)

Step 1: Standardization

First, we need to standardize our dataset to ensure that each variable has a mean of 0 and a standard deviation of 1.

Step2: Covariance Matrix Computation

measures the strength of joint variability between two or more variables, indicating how much they change in relation to each other. To find the covariance we can use the formula:

$$cov(x1, x2) = \frac{\sum_{i=1}^{n}(x1_i - x1)(x2_i - x2)}{n-1}$$

The value of covariance can be positive, negative, or zeros.

- Positive: As the x1 increases x2 also increases.
- Negative: As the x1 increases x2 also decreases.
- Zeros: No direct relation

Step 3: Compute Eigenvalues and Eigenvectors of Covariance Matrix to Identify Principal Components