Assignment

1. **Explain the Eigenvalue and eigenvector in detail along with some examples. Its role in PCA.**

**Ans:**

**Eigenvalues and Eigenvectors:** A Detailed Exploration and Their Pivotal Role in Principal Component Analysis (PCA)

Eigenvalues and eigenvectors are fundamental concepts in linear algebra that unlock a deeper understanding of linear transformations. In essence, they reveal the intrinsic properties of a matrix, showing directions in which a transformation acts simply by stretching or compressing. These concepts are not merely abstract mathematical ideas; they are the cornerstone of various applications, most notably in the dimensionality reduction technique known as Principal Component Analysis (PCA).

Unveiling Eigenvalues and Eigenvectors

Imagine a matrix as a function that transforms vectors in a space. When this transformation is applied, most vectors are knocked off their original span, pointing in new directions. However, certain special vectors, known as eigenvectors, are exceptional. When a matrix acts on an eigenvector, the resulting vector points in the exact same direction as the original eigenvector. The only change is in its magnitude; it is scaled by a factor. This scaling factor is called the eigenvalue.

Formally, for a square matrix A, a non-zero vector v is an eigenvector of A if it satisfies the following equation:

Av = λv

Where:

* A is the square matrix.
* v is the eigenvector.
* λ (lambda) is the eigenvalue, a scalar value.

An eigenvector represents a direction that is invariant under the transformation of the matrix. The corresponding eigenvalue indicates the extent of scaling (stretching or shrinking) along that direction. A positive eigenvalue means the eigenvector is stretched, a negative eigenvalue means it is stretched and its direction is reversed, and an eigenvalue between -1 and 1 signifies a compression.

Calculating Eigenvalues and Eigenvectors: A Step-by-Step Guide

The process of finding the eigenvalues and eigenvectors of a matrix involves a few key steps:

1. Form the Characteristic Equation: To find the eigenvalues, we rearrange the core equation:  
   Av - λv = 0  
   Av - λIv = 0 (where I is the identity matrix)  
   (A - λI)v = 0

For a non-zero eigenvector v, the matrix (A - λI) must be singular, meaning its determinant is zero. This leads to the characteristic equation:  
det(A - λI) = 0

1. Solve for Eigenvalues: Solving the characteristic equation for λ will yield the eigenvalues of the matrix. For an n x n matrix, this will be a polynomial of degree n, resulting in n eigenvalues (which may not all be distinct).
2. Find the Eigenvectors: For each eigenvalue found, substitute it back into the equation (A - λI)v = 0 and solve for the vector v. This will give the eigenvector(s) corresponding to that eigenvalue.

Examples

2x2 Matrix Example

Let's find the eigenvalues and eigenvectors of the following matrix:

A = [[2][3],[2][3]]

Step 1: Form the Characteristic Equation

First, we compute A - λI:  
A - λI = [[2][3],[2][3]] - λ[[3],[3]] = [[2-λ, 1], [1, 2-λ]]

Now, we find the determinant and set it to zero:  
det(A - λI) = (2-λ)(2-λ) - (1)(1) = 0  
4 - 4λ + λ² - 1 = 0  
λ² - 4λ + 3 = 0

Step 2: Solve for Eigenvalues

We solve this quadratic equation for λ:  
(λ - 3)(λ - 1) = 0  
The eigenvalues are λ₁ = 3 and λ₂ = 1.

Step 3: Find the Eigenvectors

* For λ₁ = 3:  
  We solve (A - 3I)v = 0:  
  [[2-3, 1], [1, 2-3]] [x, y] =  
  [[-1, 1], [1, -1]] [x, y] =

This gives us the equation -x + y = 0, or x = y. Any non-zero vector where the components are equal is an eigenvector. A simple choice is v₁ =[3][3].

* For λ₂ = 1:  
  We solve (A - 1I)v = 0:  
  [[2-1, 1], [1, 2-1]] [x, y] =  
  [[3][3],[3][3]] [x, y] =

This gives us the equation x + y = 0, or x = -y. A simple choice for the eigenvector is v₂ = [1, -1].

1. Use the **Students' Social Network Profile Clustering** dataset from below Kaggle link and create an end-to-end project on Jupyter/Colab.

<https://www.kaggle.com/datasets/zabihullah18/students-social-network-profile-clustering/data>

1. Download the dataset from above link and load it into your Python environment.
2. Perform the EDA and do the visualizations.
3. Check the distributions/skewness in the variables and do the transformations if required.
4. Check/Treat the outliers and do the feature scaling if required.
5. Create a ML model to segment the students with similar interests, demographic profiling, and trend analysis over time.
6. Try out all the 3 clustering methods (K-Mean, Hierarchical, DBSCAN) and compare their silhoutte scores.

Ans:

https://colab.research.google.com/drive/1IEpl3-KjUq22Jl1xG5rsEEldpXT0ozr9?usp=sharing

https://github.com/rr4323/data\_scientist\_mastry/blob/main/student\_profile\_clustering/student\_profile\_clustering.ipynb

1. Use the **Anime Recommendations** dataset from below Kaggle link and create an end-to-end project on Jupyter/Colab.

<https://www.kaggle.com/datasets/CooperUnion/anime-recommendations-database/data>

*Code for reference* - <https://www.kaggle.com/code/benroshan/content-collaborative-anime-recommendation>

1. Download the dataset from above link and load it into your Python environment.
2. Perform the EDA and do the visualizations.
3. Check the distributions/skewness in the variables and do the transformations if required.
4. Create a content based Recommender system

Ans:

https://github.com/rr4323/Sales-Data-Analysis/blob/main/anime\_recommendation/anime\_recommendation.ipynb

https://colab.research.google.com/drive/1VX7E2AzGgwg1\_eJjYSdPeioupkzQmfSJ?usp=sharing