CS573 - Optimization Methods

Fall Semester 2023
Assignment 1

Papageridis Vasileios - 4710
csd4710@csd.uoc.gr

November 17, 2023

Introduction

This assignment commences with a theoretical foundation by applying SVD to a specified matrix, elucidating the intricacies of eigenvalues, singular values, and their respective vectors. Subsequently, the focus shifts to a practical application involving the MovieLens dataset, demonstrating the real-world implications and advantages of these decomposition methods.
The first section provides a meticulous breakdown of the SVD process applied to a given matrix A. The assignment navigates through eigenvalue computations, singular value determinations, and the construction of U, Σ, and V matrices.
In the second section, the assignment transitions from theory to practice by implementing SVD and PCA on the MovieLens dataset. The practical implications of these methodologies are explored, specifically in the reduction of dimensionality and the revelation of latent patterns within the dataset. The analysis extends to the projection of data onto principal components, offering valuable insights into the variance retained in reduced spaces.

1 Exercise 1: Singular Value Decomposition and Principal Component Analysis on the paper

1.1 Singular Value Decomposition (SVD)

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 1 & 1 \end{bmatrix}$$

a)

$$A^T = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 11 & 7 \\ 7 & 11 \end{bmatrix}$$

Solving for the eigenvalues λ of $A^T A$:

$$|A^T A - \lambda I| = 0$$

Solving for the eigenvalues λ will give us the singular values.
The characteristic equation is given by:

$$|A^T A - \lambda I| = 0$$

Solving for λ, we have:

$$\det\left(\begin{bmatrix} 11-\lambda & 7 \\ 7 & 11-\lambda \end{bmatrix}\right) = (11-\lambda)^2 - 7^2 = 0$$

Expanding and solving for λ:

$$\lambda^2 - 22\lambda + 72 = 0$$

Factoring, we get:

$$(\lambda - 18)(\lambda - 4) = 0$$

So, the eigenvalues λ are 18 and 4.
Now, let's find the corresponding unit-length eigenvectors.
For $\lambda_1 = 18$:

$$(A^T A - 18I)\mathbf{v_1} = \begin{bmatrix} -7 & 7 \\ 7 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

This system of equations gives us $x_1 = x_2$. Therefore, an eigenvector corresponding to $\lambda = 18$ is any vector of the form $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
For $\lambda_2 = 4$:

$$(A^T A - 6I)\mathbf{v_2} = \begin{bmatrix} 7 & 7 \\ 7 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

This system of equations gives us $x_1 = -x_2$. Therefore, an eigenvector corresponding to $\lambda = 4$ is any vector of the form $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
Finally, normalize the eigenvectors to unit length:
For $\lambda_1 = 18$:

$$\mathbf{v_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 4$:

$$\mathbf{v_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Results: So, the eigenvalues are 16 and 6, and the corresponding unit-length eigenvectors are:

$$\mathbf{v_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{v_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 10 & 6 & 4 \\ 6 & 10 & 4 \\ 4 & 4 & 2 \end{bmatrix}$$

Solving for the eigenvalues λ of AA^T:

$$|AA^T - \lambda I| = 0$$

Solving for the eigenvalues λ will give us the singular values.
The characteristic equation is given by:

$$|AA^T - \lambda I| = 0$$

Solving for λ, we have:

$$\det\left(\begin{bmatrix} 10-\lambda & 6 & 4 \\ 6 & 10-\lambda & 4 \\ 4 & 4 & 2-\lambda \end{bmatrix}\right) = 0$$

Solving for λ we get that the eigenvalues are:

$$\lambda_1 = 18, \lambda_2 = 4, \lambda_3 = 0$$

Now, let's find the corresponding unit-length eigenvectors.
For $\lambda_1 = 18$:

$$(AA^T - 18I)\mathbf{u_1} = \begin{bmatrix} -8 & 6 & 4 \\ 6 & -8 & 4 \\ 4 & 4 & -16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

Therefore, an eigenvector corresponding to $\lambda = 18$ is any vector of the form $\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$.

For $\lambda_2 = 4$:

$$(AA^T - 4I)\mathbf{u_2} = \begin{bmatrix} 6 & 6 & 4 \\ 6 & 6 & 4 \\ 4 & 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

Therefore, an eigenvector corresponding to $\lambda = 4$ is any vector of the form $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

For $\lambda_3 = 0$:

$$(AA^T - 0I)\mathbf{u_3} = \begin{bmatrix} 10 & 6 & 4 \\ 6 & 10 & 4 \\ 4 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

Therefore, an eigenvector corresponding to $\lambda = 0$ is any vector of the form $\begin{bmatrix} -1/4 \\ -1/4 \\ 1 \end{bmatrix}$.

Finally, normalize the eigenvectors to unit length:
For $\lambda_1 = 18$:

$$\mathbf{u_1} = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 4$:

$$\mathbf{u_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

For $\lambda_3 = 0$:

$$\mathbf{u_3} = \frac{1}{3\sqrt{2}}\begin{bmatrix} -1/4 \\ -1/4 \\ 1 \end{bmatrix}$$

Results: So, the eigenvalues are 0, 4 and 18, and the corresponding unit-length eigenvectors are:

$$\mathbf{u_1} = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{u_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u_3} = \frac{1}{3\sqrt{2}}\begin{bmatrix} -1/4 \\ -1/4 \\ 1 \end{bmatrix}$$

The singular values σ are the square roots of the eigenvalues of $A^T A$ and AA^T. So,

- $\sigma_1 = \sqrt{18} \Rightarrow \sigma_1 = 3\sqrt{2}$
- $\sigma_2 = \sqrt{4} \Rightarrow \sigma_2 = 2$
- $\sigma_3 = 0$

Singular values tell us about the importance or significance of each singular vector in the decomposition. Larger singular values correspond to more significant contributions to the overall matrix transformation, while smaller singular values represent less significant contributions.

The rank of matrix A is the number of non-zero singular values. So, the rank of A is equal to 2.
A matrix rank cannot be greater than its dimensions. So, the rank is always limited by the smaller of the two dimensions of the matrix. It cannot exceed this value.

b)

Left Singular Vectors: The left singular vectors of A are the eigenvectors of AA^T.
These vectors form the columns of the U matrix.

$$U = \begin{bmatrix} 2/3 & -1/\sqrt{2} & -1/12\sqrt{2} \\ 2/3 & 1/\sqrt{2} & -1/12\sqrt{2} \\ 1/3 & 0 & 1/3\sqrt{2} \end{bmatrix}$$

Singular Values: The singular values σ that we calculated in the earlier stage are placed on the diagonal of the Σ matrix.

$$\Sigma = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$$

Right Singular Vectors: The right singular vectors of A are the eigenvectors of $A^T A$.
These vectors form the columns of the V matrix.

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

So the V^T matrix is:

$$V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Result:

$$A = U\Sigma V^T = \begin{bmatrix} 2/3 & -1/\sqrt{2} & -1/12\sqrt{2} \\ 2/3 & 1/\sqrt{2} & -1/12\sqrt{2} \\ 1/3 & 0 & 1/3\sqrt{2} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

c)

We know that in the *Singular Value Decomposition*, the singular values in Σ represent the importance of the corresponding singular vectors in U and V. Larger singular values correspond to more significant contributions to the overall matrix transformation.
Now, for dimensionality reduction we can choose to reduce the matrix A rank by 1. The reduced matrices will be:

$$U_r = \begin{bmatrix} 2/3 & -1/\sqrt{2} \\ 2/3 & 1/\sqrt{2} \\ 1/3 & 0 \end{bmatrix}$$

$$\Sigma_r = \begin{bmatrix} 3\sqrt{2} \\ 0 \end{bmatrix}$$

$$V_r = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$V_r^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

The matrix A can now be expressed as:

$$A = U_r \Sigma_r V_r^T = \begin{bmatrix} 2/3 & -1/\sqrt{2} \\ 2/3 & 1/\sqrt{2} \\ 1/3 & 0 \end{bmatrix} \begin{bmatrix} 3\sqrt{2} \\ 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \\ 1 & 1 \end{bmatrix}$$

1.2 Principal Components Analysis (PCA)

a)

The principal components of a matrix are the columns of matrix V in its
Singular Value Decomposition (SVD).
In our case, for the matrix A is given as:

$$A = U\Sigma V^T = \begin{bmatrix} 2/3 & -1/\sqrt{2} & -1/12\sqrt{2} \\ 2/3 & 1/\sqrt{2} & -1/12\sqrt{2} \\ 1/3 & 0 & 1/3\sqrt{2} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

So the principal components of matrix A are:

$$\mathbf{PC_1} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\mathbf{PC_2} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Relationship between singular values and principal components:
The singular values determine the spread or "length" of the principal components. Larger singular values indicate more significant contributions to the overall variation in the data. So, the singular values indicate the importance or significance of the corresponding principal components, and they are used to scale these components in the reconstruction of the original matrix.

1.2.1 b)

To project the data into 1 dimension, we calculate the following:

$$PC_1 \cdot A^T = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} & 2\sqrt{2} & \sqrt{2} \end{bmatrix}$$

Exercise 2: Interpretation and Reduction of the MovieLens Dataset Using SVD and PCA

In this section, a Jupyter notebook was utilized to conduct a comprehensive analysis of the MovieLens dataset employing Singular Value Decomposition (SVD) and Principal Component Analysis (PCA). The code implementation includes the following key steps:

Data Loading and Preprocessing:

- Importing the necessary libraries.
- Loading the MovieLens dataset.
- Preprocessing steps such as handling missing values.

Singular Value Decomposition (SVD):

- Creating an implementation of SVD.
- Applying SVD to decompose the user-movie rating matrix.
- Assessing the significance of singular values in capturing data variance.

Principal Component Analysis (PCA):

- Implementing PCA to reduce dimensionality.
- Analyzing the principal components.

Nearest Neighbors in Low-Dimensional Space:

- Finding the nearest neighbors for selected validation users within the PCA space.
- Calculating similarity scores based on common movie preferences.

Visualization:

- Plotting movies and their ratings for the validation user and nearest neighbors.
- Visualizing the distribution of ratings for enhanced understanding.

Discussion and Insights:

- Reflecting on the effectiveness of PCA in capturing user similarity.
- Analyzing pairings of validation users and their nearest neighbors.
- Assessing the impact of dimensionality reduction on user similarity detection.

Conclusion

SVD provided a robust technique for decomposing the user-movie rating matrix, uncovering latent patterns within the dataset. The resulting singular values and vectors offered insights into the significance of different components in capturing data variance. PCA, as an extension of SVD, facilitated dimensionality reduction, allowing for a more concise representation of user preferences.

The examination of nearest neighbors in the low-dimensional space further highlighted the effectiveness of SVD and PCA in capturing user similarity. By calculating similarity scores and analyzing common movie preferences, the assignment shed light on the practical implications of these optimization methods in the context of recommendation systems.