CS/ECE/ME532 Period 10 Activity

Estimated Time: 20 mins for P1, 25 mins for P2, 10 mins for P3

1. Let $\boldsymbol{A} = \begin{bmatrix} 3 & 3 & 3 & -1 & -1 & -1 \\ 1 & 1 & 1 & -3 & -3 & -3 \\ 1 & 1 & 1 & -3 & -3 & -3 \\ 3 & 3 & 3 & -1 & -1 & -1 \end{bmatrix}$. Use the provided script to help you complete the problem.

a) Use the K-means algorithm to represent the columns of $\boldsymbol{A}$ with a single cluster.

i. Express the rank-1 approximation to $\boldsymbol{A}$ based on this cluster as $\boldsymbol{T}\boldsymbol{W}^T$ where the column $\boldsymbol{T}$ contains the cluster and $\boldsymbol{W}$ is a vector of ones and zeros.

ii. Compare the rank-1 clustering approximation to the original matrix.

b) Use the K-means algorithm to represent the columns of $\boldsymbol{A}$ with two clusters.

i. Express the rank-2 approximation to $\boldsymbol{A}$ based on this cluster as $\boldsymbol{T}\boldsymbol{W}^T$ where columns of $\boldsymbol{T}$ contain the clusters and $\boldsymbol{W}$ is a matrix of ones and zeros.

ii. Compare the rank-2 clustering approximation to the original matrix.

2. Again let $\boldsymbol{A} = \begin{bmatrix} 3 & 3 & 3 & -1 & -1 & -1 \\ 1 & 1 & 1 & -3 & -3 & -3 \\ 1 & 1 & 1 & -3 & -3 & -3 \\ 3 & 3 & 3 & -1 & -1 & -1 \end{bmatrix}$. Now consider the singular value decomposition (SVD) $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$.

a) If the full SVD is computed, find the dimensions of $\boldsymbol{U}$, $\boldsymbol{S}$, and $\boldsymbol{V}$.

b) Find the dimensions of $\boldsymbol{U}$, $\boldsymbol{S}$, and $\boldsymbol{V}$ in the economy or skinny SVD of $\boldsymbol{A}$.

c) The Python and NumPy command `U,s,VT = np.linalg.svd(A,full_matrices=True)` computes the singular value decomposition, $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ where $\boldsymbol{U}$ and $\boldsymbol{V}$ are matrices with orthonormal columns comprising the left and right singular vectors and $\boldsymbol{S}$ is a diagonal matrix of singular values. Use this command to find the SVD of $\boldsymbol{A}$.

i. Find $\boldsymbol{U}^T\boldsymbol{U}$ and $\boldsymbol{V}^T\boldsymbol{V}$. Are the columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ orthonormal? Why? *Hint:* compute $\boldsymbol{U}^T\boldsymbol{U}$.

ii. Find $\boldsymbol{U}\boldsymbol{U}^T$ and $\boldsymbol{V}\boldsymbol{V}^T$. Are the rows of $\boldsymbol{U}$ and $\boldsymbol{V}$ orthonormal? Why?

iii. Find the left and right singular vectors associated with the largest singular value.

iv. What is the rank of $\boldsymbol{A}$?

d) The command`U,s,VT = np.linalg.svd(A,full_matrices=False)` computes the economy or skinny singular value decomposition, $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ where $\boldsymbol{U}$ and $\boldsymbol{V}$ are matrices with orthonormal columns comprising the left and right singular vectors and $\boldsymbol{S}$ is a square diagonal matrix of singular values. Find the economy SVD of $\boldsymbol{A}$.

i. Find $\boldsymbol{U}^T\boldsymbol{U}$ and $\boldsymbol{V}^T\boldsymbol{V}$. Are the columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ orthonormal? Why?

ii. Find $\boldsymbol{U}\boldsymbol{U}^T$ and $\boldsymbol{V}\boldsymbol{V}^T$. Are the rows of $\boldsymbol{U}$ and $\boldsymbol{V}$ orthonormal? Why?

e) Compare the singular vectors and singular values of the economy and full SVD. How do they differ?

f) Identify an orthonormal basis for the space spanned by the columns of $\boldsymbol{A}$.

g) Identify an orthonormal basis for the space spanned by the rows of $\boldsymbol{A}$.

h) Define the rank-r approximation to $\boldsymbol{A}$ as

$$\boldsymbol{A}_r = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

where σ_i is the ith singular value with left singular vector $\boldsymbol{u}_i$ and right singular vector $\boldsymbol{v}_i$.

i. Find the rank-1 approximation $\boldsymbol{A}_1$. How does $\boldsymbol{A}_1$ compare to $\boldsymbol{A}$?

ii. Find the rank-2 approximation $\boldsymbol{A}_2$. How does $\boldsymbol{A}_2$ compare to $\boldsymbol{A}$?

i) The economy SVD is based on the dimension of the matrices and does not consider the rank of the matrix. What is the smallest economy SVD (minimum dimension of the square matrix $\boldsymbol{S}$) possible for the matrix $\boldsymbol{A}$? Find $\boldsymbol{U}, \boldsymbol{S}$, and $\boldsymbol{V}$ for this minimal economy SVD.

3. Consider expressing the SVD of a rank-r matrix $\boldsymbol{X}$ as

$$\boldsymbol{X} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

where σ_i is the ith singular value with left singular vector $\boldsymbol{u}_i$ and right singular vector $\boldsymbol{v}_i$. Is the sign of the singular vectors unique? Why or why not? *Hint:* Consider replacing $\boldsymbol{u}_1$ with $\tilde{\boldsymbol{u}}_1 = -\boldsymbol{u}_1$.