CS/ECE/ME532 Period 11 Activity

Estimated time: 30 *mins for Q1 and* 30 *mins for Q2.*

1. We return to the movies rating problem considered previously. The movies and ratings from your friends on a scale of 1-10 are:

Movie	Jake	Jennifer	Julia	Justin	Jackson	James	Jasmine
Star Trek	4	7	2	8	7	4	2
Pride and Prejudice	9	3	5	6	10	5	5
The Martian	4	8	3	7	6	4	1
Sense and Sensibility	9	2	6	5	9	5	4
Star Wars: Empire Strikes	4	9	2	8	7	4	1

A script is available to help you complete this activity.

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

i. Express the rank-2 approximation to $\boldsymbol{X}$ based on this cluster as $\boldsymbol{T}\boldsymbol{W}^T$ where the columns of $\boldsymbol{T}$ contains the cluster centers and $\boldsymbol{W}$ is a vector of ones and zeros.

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

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

i. Express the rank-3 approximation to $\boldsymbol{X}$ based on this cluster as $\boldsymbol{T}\boldsymbol{W}^T$ where the column $\boldsymbol{T}$ contains the cluster centers and $\boldsymbol{W}$ is a vector of ones and zeros.

ii. Compare the rank-3 clustering approximation to the original matrix.

c) Use SVD to approximate $\boldsymbol{X}$ using r "tastes", the columns of $\boldsymbol{T}$, that is, $\boldsymbol{X} \approx \boldsymbol{T}\boldsymbol{W}$ where $\boldsymbol{T}$ is 5-by-r and has orthonormal columns.

i. Express $\boldsymbol{T}$ and $\boldsymbol{W}$ as a function of the matrices in the SVD $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$. (In equation form, not numbers.)

ii. Find $\boldsymbol{T}, \boldsymbol{W}$ and the rank-r approximation to $\boldsymbol{X}$ for $r = 1$. What aspects of the ratings does the first taste capture?

iii. Find $\boldsymbol{T}, \boldsymbol{W}$ and the rank-r approximation to $\boldsymbol{X}$ for $r = 2$. What aspects of the ratings does the second taste capture?

iv. Your friend Jon rates Star Trek 6 and Pride and Prejudice 4. You use the first two tastes of your other friends to predict his ratings of the remaining three movies.

A. Formulate a system of equations that can be solved to predict Jon's preferences for the remaining three movies.

B. Predict Jon's ratings for the remaining three movies.

2. Let a 4-by-2 matrix $\boldsymbol{X}$ have SVD $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ where $\boldsymbol{U} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\boldsymbol{S} = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}$, and $\boldsymbol{V} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

a) Express the solution to the least-squares problem $\min_{\boldsymbol{w}} ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2$ as a function of $\boldsymbol{U}$, $\boldsymbol{S}, \boldsymbol{V}$, and $\boldsymbol{y}$.

b) Let $\boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$. Find the weights $\boldsymbol{w}$ as a function of γ. Calculate $||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}||_2^2$ and $||\boldsymbol{w}||_2^2$ as a function of γ. What happens to $||\boldsymbol{w}||_2^2$ as $\gamma \to 0$?

c) Now consider a "low-rank" inverse. Instead of writing

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T = \sum_{i=1}^{p} \frac{1}{\sigma_i}\boldsymbol{v}_i\boldsymbol{u}_i^T$$

where p is the number of columns of $\boldsymbol{X}$ (assumed less than the number of rows), we approximate

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \approx \sum_{i=1}^{r} \frac{1}{\sigma_i}\boldsymbol{v}_i\boldsymbol{u}_i^T$$

In this approximation we only invert the largest r singular values, and ignore all of them smaller than σ_r. If $r = 1$, use the low-rank inverse to find $\boldsymbol{w}$, $||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}||_2^2$, and $||\boldsymbol{w}||_2^2$ when $\boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ as in part b). Compare $||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}||_2^2$, and $||\boldsymbol{w}||_2^2$ to the results for part b).