Machine Learning Homework 10

Author: Jesus Arturo Sol Navarro

1 Problem 3

For the following , assume that X is already centered and $\sum_X = \frac{1}{N}X^TX$

1.1 a)

$$\Sigma_{Y_1} = \frac{1}{N}(I\lambda X)^T(I\lambda X) = \frac{1}{N}\lambda X^T\lambda X = \lambda^2\Sigma_X$$

$$\frac{\sum_{i=1}^{K}\lambda^2\lambda_i}{\sum_{i=1}^{D}\lambda^2\lambda_i}$$

Since λ^2 cancels out, the fraction of variance preserved is 70%, just like in X .

1.2 b)

$$\Sigma_{Y_2} = \frac{1}{N}(XR)^T(XR) = R^T\Sigma_X R$$

Since the rotation matrices perform a similarity transformation,the eigenvalues remain the same and the fraction of variance is 70%, just like in X .

1.3 c)

$$\Sigma_{Y_3} = \frac{1}{N}(XP)^T(XP) = P\Sigma_X P$$

Multiplying by P uniformly scales and possibly flips signs in certain coordinates, but that does not alter the eigenvalues . The fraction of variance preserved is 70%.

1.4 d)

$$\Sigma_{Y_4} = \frac{1}{N}(XQ)^T(XQ) = Q^T\Sigma_X Q$$

In comparison to c), multiplying by Q does not scales uniformly. This means that additional information is needed to provide a justification.

1.5 e)

Each row of X is shifted by the same vector. This means that the relative distance between points will be the same. Thus, $\Sigma_{Y_5} = \Sigma_X$. Same covariance, means same eigenvalues and same variance preserved after projecting the data.

1.6 f)

Since rank(A)=5, A reduces X onto a subspace of maximum 5 dimensions. 5 non-zero eigenvalues and the rest become zero. However, more information is needed to provide information.

2 Problem 4

2.1 a)

- Center the Data

$$\bar{\mathbf{X}} = \frac{1}{N}\cdot\mathbf{X}^T\cdot\mathbf{1}_4 = \begin{bmatrix}2\\1\\1\end{bmatrix}$$

$$\tilde{\mathbf{X}} = \mathbf{X} - \bar{\mathbf{X}} = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & -3 \\ 2 & -2 & 1 \\ -4 & 0 & 1 \end{bmatrix}$$

- Compute the covariance matrix

$$\Sigma_{\tilde{\mathbf{X}}} = \frac{1}{4}\tilde{\mathbf{X}}^T\tilde{\mathbf{X}} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

- Use the Eigenvector Decomposition to transform the coordinate system

$$\Sigma_{\tilde{X}} = \Gamma \cdot \Lambda \cdot \Gamma^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^T$$

2.2 b)

$$\mathbf{Y} = \tilde{\mathbf{X}} \cdot \Gamma_2 = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & -3 \\ 2 & -2 & 1 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & -3 \\ 2 & 1 \\ -4 & 1 \end{bmatrix}$$

$$\text{Fraction of variance preserved} = \frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{D} \lambda_i} = 0.8181\%$$

2.3 c)

$$X_5 = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}$$

If X_5 is the same as the mean of the original dataset the same principal components remain.

3 Problem 5

$$L = [0, 3, 0, 0, 4]$$

$$LV\Sigma^{-1} = (U\Sigma V^{\top})V\Sigma^{-1} = U = L_{concept}$$

$$\begin{bmatrix} 0 & 3 & 0 & 0 & 4 \end{bmatrix} \begin{pmatrix} 0.58 & 0 \\ 0.58 & 0 \\ 0.58 & 0 \\ 0 & 0.71 \\ 0 & 0.71 \end{pmatrix} \begin{pmatrix} \frac{1}{12.4} & 0 \\ 0 & \frac{1}{9.5} \end{pmatrix} = \begin{bmatrix} 0.14 & 0.30 \end{bmatrix}$$

Score on second latent dimension (0.30) is higher than first (0.14) She is more likely to enjoy Casablanca more than Star Wars or Matrix, still se might like sci-fi movies, a bit less though than drama.

4 Problem 6

$$w^* = (X^TX)^{-1}X^Ty = (V\Sigma^2V^T)^{-1}V\Sigma U^Ty = (V(\Sigma^2)^{-1}V^T)V\Sigma U^Ty$$

$$w^* = V\Sigma^{-1}U^Ty$$

- $a = U^Ty$ has complexity $O(ND)$
- $b = \Sigma^{-1}a$ has complexity $O(D)$
- Vb has complexity $O(D^2)$
- Total complexity is $O(ND) + O(D) + O(D^2) = O(ND)$, since $N > D$