Big Data Management and Analytics
WS 2017/18

Tutorial 9: High Dimensionality Data

Assignment 9-1   *PCA*

Consider the $X \in \mathbb{R}^{M \times N}$ matrix containing six data points $X_i \in \mathbb{R}^2$.

$$X = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 0 \\ 5 & 6 \\ 6 & 6 \\ 7 & 6 \end{pmatrix}$$

dim 1	1	2	3	5	6	7
dim 2	0	0	0	6	6	6

Conduct a PCA on the given data, i.e. project the data onto a one-dimensional space. Please state the eigenvectors, eigenvalues, covariance matrix and visualize the data before and after the PCA.

Assignment 9-2   *PCA Power Iteration*

Given the matrix M:

$$M = \begin{pmatrix} \frac{14}{3} & 6 \\ 6 & 9 \end{pmatrix}$$

1. Determine the strongest eigenvector of M using the Power Iteration method.

2. After how many iterations can a convergence be observed?

Assignment 9-3   *SVD Decomposition*

Given the matrix M:

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$$

1. Find the eigenpairs for matrix M

2. Find the SVD for the original matrix $M = U\Sigma V^T$

3. Compute the one-dimensional approximation of the matrix M