Big Data Management and Analytics
Assignment 11

Assignment 11-1

Given the matrix M:

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

1. Find the eigenpairs for matrix M

i. Compute: $M^T M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$

ii. Find eigenvalues:

$$\det(M^T M - \lambda \cdot I_{2x2}) = 0$$
$$\lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2)$$
$$Eigenvalue\ \lambda_1 = 4 \ \rightarrow singular\ value\ \sigma_1 = \sqrt{\lambda_1} = 2$$
$$Eigenvalue\ \lambda_2 = 2 \ \rightarrow singular\ value\ \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}$$

Assignment 11-1

iii. Find eigenvectors:

$$1^{st} eigenvector\ v_1: (M^T M - \lambda_1 \cdot I_{2x2})v_1 = 0 \rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} v_1 = 0$$

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{normalize} v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$2^{nd} eigenvector\ v_2: (M^T M - \lambda_2 \cdot I_{2x2})v_2 = 0 \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v_2 = 0$$

$$v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xrightarrow{normalize} v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

iv. Eigenpairs (eigenvalue, eigenvector):

$$(\lambda_1, v_1) = \left(4, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\right), (\lambda_2, v_2) = \left(2, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\right)$$

Assignment 11-1

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

From the results of (1.) we know:

$$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \, and \, V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Assignment 11-1

- How can we find now U?
  Multiply the SVD $A = U\Sigma V^T$ with V on each side yields: $AV = U\Sigma$

$$U \cdot \Sigma = (u_1\, u_2\; \ldots\; u_m) \cdot \begin{pmatrix} \sigma_1 & 0 & \cdots \\ 0 & \sigma_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$
$$= (\sigma_1 \cdot u_1 \quad \sigma_2 \cdot u_2 \quad \ldots\; \sigma_r \cdot u_r \;\; 0 \ldots 0)$$
$$= (A \cdot v_1 \quad A \cdot v_2 \quad \ldots A \cdot v_r \;\; 0 \ldots 0)$$

Assignment 11-1

$$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \qquad A = M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$$

- Compute

$$u_1 = \frac{1}{\sigma_1} \cdot A \cdot v_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \frac{2}{\sqrt{2}} \\ \frac{2}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$u_2 = \frac{1}{\sigma_2} \cdot A \cdot v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Assignment 11-1

$$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \qquad A = M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$$

- Having $u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and $u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ we could write now the SVD as follows:

$$M = U\Sigma V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & * \\ \frac{1}{\sqrt{2}} & 0 & * \\ 0 & 1 & * \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & * \\ \frac{1}{\sqrt{2}} & 0 & * \\ 0 & 1 & * \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 1 & -1 \\ 0 & 0 \end{pmatrix}$$

Assignment 11-1

$$M = U\Sigma V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & * \\ \frac{1}{\sqrt{2}} & 0 & * \\ 0 & 1 & * \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & * \\ \frac{1}{\sqrt{2}} & 0 & * \\ 0 & 1 & * \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ 1 & -1 \\ 0 & 0 \end{pmatrix}$$

- In the last matrix multiplication → entries in the last column of U get multiplied by 0

Assignment 11-1

- How do we compute now u_3 as a third orthonormal vector?
- $\{u_1, u_2\}$ is an orthonormal basis for a plane in $\mathbb{R}^3$
  - To extend $\{u_1, u_2\}$ to an orthonormal basis for all of $\mathbb{R}^3$ we need a third vector u_3 that is normal to this plane
  - How? Compute the cross product: $u_3 = u_1 \times u_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}$
- Now we have all components:

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{\sqrt{2}}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Assignment 11-1

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

The k-approximated representation is given by $M \approx U_k \Sigma_k V_k^T$. Set k=1:

$$M \approx U_k \Sigma_k V_k^T \approx \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \cdot (2) \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{\sqrt{2}} & \frac{2}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$$

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

Assignment 11-2

- Find the CUR-decomposition of the matrix, when we pick two „random“ rows and columns. The columns we pick are Alien and StarWars and the rows are the ones of Jack and Jill.

- From the lecture (Ch.7, Sl. 46) we know that the scaled column for Alien is: $[1.54, 4.63, 6.17, 7.72, 0, 0, 0]^T$. The second column for Star Wars is the same. We thus define C as follows:

$$C = \begin{pmatrix} 1.54 & 1.54 \\ 4.63 & 4.63 \\ 6.17 & 6.17 \\ 7.72 & 7.72 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Assignment 11-2

- The unscaled rows for R are:

$$R_{unscaled} = \begin{pmatrix} 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \end{pmatrix}$$

- The probability p_i with which we select row i is given by:

$$p_i = \sum_j m_{i,j}^2 \Big/ \|M\|_F^2$$

- The **square of the Frobenius norm for M** is $\|M\|_F^2 = 243$
- The **square of the Frobenius norm for Jack** is: $row_{jack} = \sum_j m_{3,j}^2 = 5^2 + 5^2 + 5^2 = 75$
- The **square of the Frobenius norm for Jill** is: $row_{jill} = \sum_j m_{4,j}^2 = 4^2 + 4^2 = 32$
- The **probability for selecting Jack** is: $p_{jack} = 75/243 = 0.309$
- The **probability for selecting Jill** is: $p_{jill} = 32/243 = 0.132$

Assignment 11-2

- The unscaled rows for R are:

$$R_{unscaled} = \begin{pmatrix} 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \end{pmatrix}$$

- Scaling the row for Jack, we divide all its row entries by:

$$\sqrt{r * p_{jack}} = \sqrt{2 * 0.309} = 0.786$$

- Scaling the row for Jill, we divide all its row entries by:

$$\sqrt{r * p_{jill}} = \sqrt{2 * 0.132} = 0.514$$

- This yields the scaled matrix R:

$$\mathrm{R} = \begin{pmatrix} 6.36 & 6.36 & 6.36 & 0 & 0 \\ 0 & 0 & 0 & 7.78 & 7.78 \end{pmatrix}$$

Assignment 11-2

- Now that we have C and R, we construct the middle matrix U:

- First construct a matrix W from the intersection of the selected rows from R and the columns from C:

$$W = \begin{pmatrix} 5 & 5 \\ 0 & 0 \end{pmatrix}$$

	Matrix	Alien	Star Wars	Casablanca	Titanic
Joe	1	1	1	0	0
Jim	3	3	3	0	0
John	4	4	4	0	0
Jack	5	5	5	0	0
Jill	0	0	0	4	4
Jenny	0	0	0	5	5
Jane	0	0	0	2	2

Assignment 11-2

- $W = \begin{pmatrix} 5 & 5 \\ 0 & 0 \end{pmatrix}$
- Take the SVD from W:

$$W = X\Sigma Y^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

- Taking the Moore-Penrose pseudoinverse of Σ leads to:

$$\Sigma^+ = \begin{pmatrix} 1/\sqrt{50} & 0 \\ 0 & 0 \end{pmatrix}$$

Assignment 11-2

- Now we can compute $U = Y(\Sigma^+)^2 X^T$

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \left(\begin{pmatrix} 1/\sqrt{50} & 0 \\ 0 & 0 \end{pmatrix} \right)^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/50 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1/50\sqrt{2} & 0 \\ 1/50\sqrt{2} & 0 \end{pmatrix}$$

Assignment 11-3

- Given the matrix $M = \begin{pmatrix} 14/3 & 6 \\ 6 & 9 \end{pmatrix}$
- Determine the strongest eigenvector of M using the Power Iteration method.

```
Input: dxd data matrix M
x_0 = random unit vector
while x_i/||x_i||- x_{i-1}/||x_{i-1}||> ε do
        x_i = M^i x_0
        i=i+1
return x_i/||x_i||
```

Assignment 11-3


```
iteration:  1
x_i:  [[ 10.66666667  15.        ]]
x_i-1:  [[ 1.  1.]]
x_i_norm:  [[ 0.57952379  0.81495532]]
x_i-1_norm:  [[ 0.70710678  0.70710678]]
delta:  [[-0.12758299  0.10784854]]
----------------
iteration:  2
x_i:  [[ 139.77777778  199.        ]]
x_i-1:  [[ 10.66666667  15.        ]]
x_i_norm:  [[ 0.57478017  0.81830786]]
x_i-1_norm:  [[ 0.57952379  0.81495532]]
delta:  [[-0.00474361  0.00335254]]
----------------
iteration:  3
x_i:  [[ 1846.2962963   2629.66666667]]
x_i-1:  [[ 139.77777778  199.        ]]
x_i_norm:  [[ 0.57461679  0.8184226 ]]
x_i-1_norm:  [[ 0.57478017  0.81830786]]
delta:  [[-0.00016339  0.00011474]]
----------------
iteration:  4
x_i:  [[ 24394.04938272  34744.77777778]]
x_i-1:  [[ 1846.2962963   2629.66666667]]
x_i_norm:  [[ 0.57461117  0.81842654]]
x_i-1_norm:  [[ 0.57461679  0.8184226 ]]
delta:  [[ -5.61592869e-06   3.94293042e-06]]
----------------
convergence reached:  [[ 0.57461117  0.81842654]]
```