

# Solutions to Sheet 4

Friedrich May, 355487; Markus Moll, 406263; Mariem Mounir, 415862  
Group 57

June 14, 2020

## Exercise 1

From the given points the centered datamatrix

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 3.5 & 0 & 0.5 \\ 2.5 & 2 & 1.5 \\ -2.5 & -2 & -1.5 \\ -3.5 & 0 & -0.5 \end{bmatrix}$$

can be created. The covariance matrix

$$C = A^T A = \begin{bmatrix} 37 & 10 & 11 \\ 10 & 8 & 6 \\ 11 & 6 & 5 \end{bmatrix}$$

is derived from the datamatrix  $A$ . From the characteristic polynomial of  $C - \lambda^3 + 50\lambda^2 - 264\lambda$  the eigenvalues

$$\lambda_1 = 44, \lambda_2 = 6, \lambda_3 = 0$$

can be calculated. Using the corresponding eigenvectors

$$v_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} -1 \\ -4 \\ 7 \end{pmatrix}$$

the PCA-transformation

$$U = \begin{bmatrix} 3\sqrt{11} & -\sqrt{6} & -\sqrt{66} \\ \sqrt{11} & 2\sqrt{6} & -4\sqrt{66} \\ \sqrt{11} & \sqrt{6} & 7\sqrt{66} \end{bmatrix}$$

of  $A$  can be calculated. Therefore the first and second principal components of  $A$  are

$$\mathbb{P}_1 = \text{span} \left( \begin{pmatrix} 3\sqrt{11} \\ \sqrt{11} \\ \sqrt{11} \end{pmatrix} \right), \mathbb{P}_2 = \text{span} \left( \begin{pmatrix} -\sqrt{6} \\ 2\sqrt{6} \\ \sqrt{6} \end{pmatrix} \right)$$

. That leads to the best fit subspaces

$$\mathbb{U}_1 = \mathbb{P}_1, \mathbb{U}_2 = \text{span} \left( \sqrt{11} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \sqrt{6} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \right)$$

## Exercise 2

a)

$A \in \mathbb{R}^{n \times n}$  is a symmetric matrix. Let  $(a_i)_{i \in [n]}$  be the eigenvectors and  $(\lambda_i)_{\lambda \in [n]}$  the eigenvalues.

if  $\lambda_i \geq 0$ :

We have  $A = U \text{diag}(\lambda_i) U^T$  with  $U \in \mathbb{R}^{n \times n}$  orthogonal matrix. Let  $x \in \mathbb{R}^n$

$$x^T A x = x^T U \text{diag}(\lambda_i) U^T x = \sum_{i=1}^n \lambda_i ([U^T x]_i)^2 \geq 0$$

.

if  $x^T A x \geq 0$ :

for  $x = a_i$

$$a_i^T A a_i = \lambda_i a_i^T a_i \geq 0 \text{ and } a_i^T a_i \geq 0$$

. So  $\lambda_i \geq 0$ .

b)

For

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$x^T A x = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -3 < 0$$

## Exercise 3

a) If we compute the power iteration on  $M_1$  and  $M_2$  for 5 iterations, we find the following eigenvectors:

- for  $M_1$  : (-0.948, -0.318)
- for  $M_2$  : (0.410, 0.404, 0.817)

- b) if we compute three iterations of the power iteration procedure for  $M_3$ , we find : ( 0.577, 0.577, -0.577).

The Power Iteration will not converge, because if we calculate the eigenvalues of  $M_3$ , we will find 2 values -2 and 2. So we don't have the assumption 5.24. That's why the Power Iteration Algo fails.

- c)  $M_1$  have two eigenvalues  $\lambda_1 = -2$  and  $\lambda_2 = -1$

If we apply the power iteration algo on  $M_1^{-1}$ , we will find the eigenvector of  $M_1$  associated with  $\lambda_2$  We find : (0.970, 0.239)

- d) • for  $M_4$  it needs 3 iterations and the eigenvector is :  
(0.6401, 0.7682)  
• for  $M_5$  it needs 60 iterations and the eigenvector is :  
(0.5814, 0.8135)

## Exercise 4

a)

Compute  $S$ :

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

$$L = D - S$$

$$D = \begin{pmatrix} 2 & 0 & & \dots & & & & & 0 \\ 0 & 3 & 0 & & & & & & \\ & 0 & 2 & 0 & & & & & \\ & & 0 & 3 & 0 & & & & \\ \vdots & & & 0 & 2 & 0 & & & \vdots \\ & & & & 0 & 3 & 0 & & \\ & & & & & 0 & 3 & 0 & \\ & & & & & & 0 & 3 & 0 \\ 0 & & & \dots & & & 0 & 2 & \end{pmatrix}$$

$$L = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 3 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 2 \end{pmatrix}$$

**b)**

The calculation of the eigenvalues of  $L$  yielded the eigenvalues 0,0.6972,0.6972,3,3,3,4.3028,4.3028,5. The three smallest eigenvalues therefore are  $\lambda_1 = 0$ ,  $\lambda_2 = 0.6972$  and  $\lambda_3 = 0.6972$ . The corresponding eigenvectors are

$$e_1 = (-1/3 \ -1/3 \ -1/3 \ -1/3 \ -1/3 \ -1/3 \ -1/3 \ -1/3 \ -1/3)^T, \ e_2 = (-0.2331 \ -0.2605 \ -0.0432 \ -0.3234 \ -0.3507 \ -0.1335 \ 0.3940 \ 0.3666 \ 0.5838)^T, \ e_3 = (0.5396 \ 0.3046 \ 0.3984 \ -0.2367 \ -0.4717 \ -0.3778 \ 0.0733 \ -0.1617 \ -0.0679)^T$$

**c)**

From the eigenvectors the matrix  $U$  can be derived

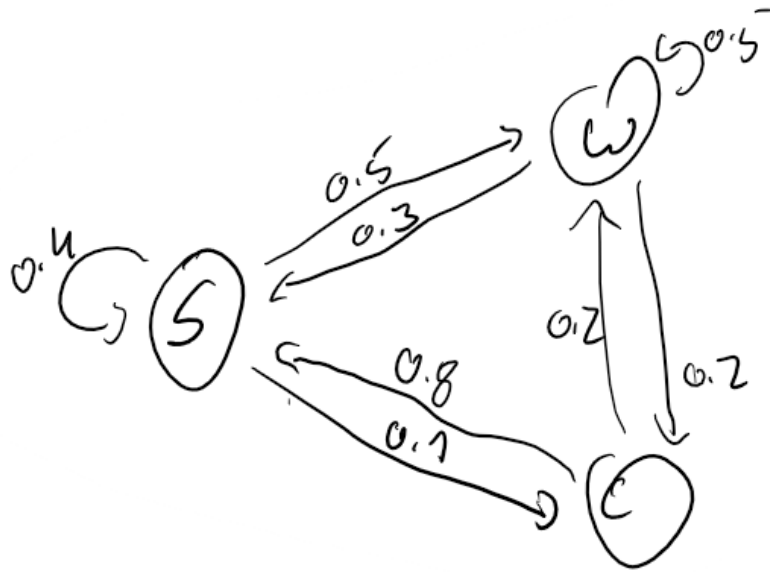
$$U = \begin{pmatrix} -0.3333 & -0.2331 & 0.5396 \\ -0.3333 & -0.2605 & 0.3046 \\ -0.3333 & -0.0432 & 0.3984 \\ -0.3333 & -0.3234 & -0.2367 \\ -0.3333 & -0.3507 & -0.4717 \\ -0.3333 & -0.1335 & -0.3778 \\ -0.3333 & 0.3940 & 0.0733 \\ -0.3333 & 0.3666 & -0.1617 \\ -0.3333 & 0.5838 & -0.0679 \end{pmatrix}$$

Splitting the rows of  $U$  using k-Means clustering yields the clusters  $\{U_1, U_2, U_3\}$ ,  $\{U_4, U_5, U_6\}$ ,  $\{U_7, U_8, U_9\}$ .

## Exercise 5

**a)**

The Markov Chain for the students behaviour can be described by the following graph:



**b)**

We can derive the transition matrix

$$Q = \begin{pmatrix} 0.4 & 0.5 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.8 & 0.2 & 0 \end{pmatrix}$$

The initial distribution is the distribution after leaving the state  $c$ . So  $p_0 = (0.8 \ 0.2 \ 0)$ . From this distribution the probability distribution after 2 hours can be derived:

$$p_2 = p_0 * Q^2 = (0.398 \ 0.464 \ 0.138)$$

Therefore the probability of the student starting to clean two hours later is 13.8%.

**c)**

To find the long-term fraction of time the student is studying we need to find the stationary distribution of the Markov Chain. The Markov Chain obviously is ergodic. Therefore we can find a stationary distribution  $\pi$  by calculating  $p_t = p_0 * Q^t$  for any starting distribution until  $p_t = p_{t+1}$ .

$t$	$p_t$
0	(0.398 0.464 0.138)
1	(0.4088 0.4586 0.1326)
2	(0.40718 0.46022 0.1326)
3	(0.407018 0.46022 0.1326762)
4	(0.407083 0.460171 0.1326746)
5	(0.407081 0.460176 0.1326743)
6	(0.407079 0.460177 0.1326743)
7	(0.407078 0.460177 0.1326743)
8	(0.407078 0.460177 0.1326743)

The fraction of time the student is studying in the long run is equivalent to the stationary propability of the student choosing to study wich is 40.708%.