Solutions to Sheet 4

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

June 10, 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 characterisitc polynominal 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

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

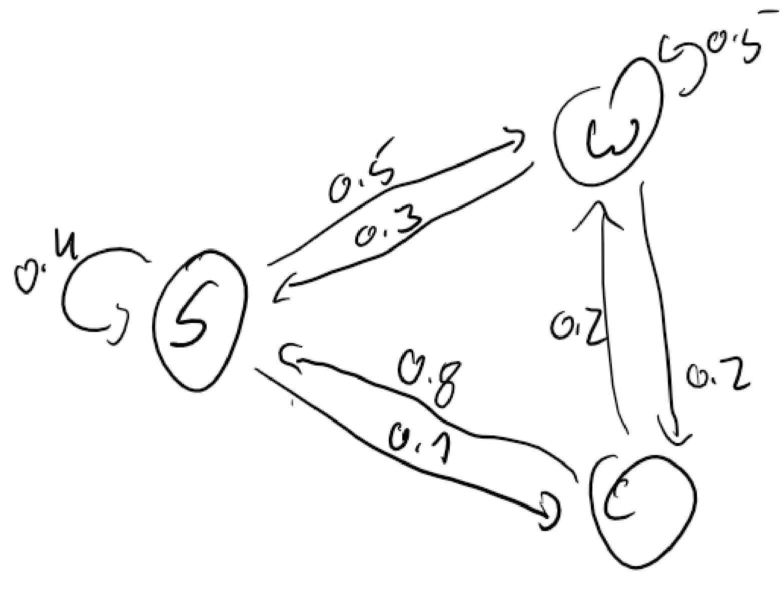
. That leads to the best fit subspaces

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

Exercise 5

a)

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



b)

We can derive the transition matrix

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

The initial distribution is the distribution after leaving the state c. So $p_0 = \begin{pmatrix} 0.8 & 0.2 & 0 \end{pmatrix}$. From this distribution the propability distribution after 2 hours can be dervied:

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

Therefore the proibaility 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 Martkov Chain obviously is ergodic. Therefore we can find a stationary distribution π by calculating $p_t = p_0 * \mathcal{Q}^t$ 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 eqivalent to the stationary propability of the student choosing to study wich is 40.708%.