Cognitive Algorithms - Exercise Sheet 5

Unsupervised Learning

Department of Machine Learning - TU Berlin

Disclaimer

For each exercise, but particularly for exercises involving calculations:

Show your work or you will not receive (full) credit!

Furthermore: Exercises marked with an asterisk * are not required and don't contribute towards the credit you receive, but you are welcome to do them because they are fun.

Task 1 - Eigenvalues and Eigenvectors of Special Matrices [1.5 points]

Recall an eigenvector of a square matrix $A \in \mathbb{R}^{d\times d}$ is defined as a non-zero vector $\mathbf{v} \in \mathbb{R}^d$ such that $A\mathbf{v} = \lambda\mathbf{v}$ where $\lambda \in \mathbb{C}$ is called the eigenvalue of A corresponding to $\mathbf{v}$.

1. Which $d \times d$ matrices have only real eigenvalues? **[0.5 points]**
2. Consider an arbitrary **triangular** matrix.
   What are the corresponding **eigenvalues** of this matrix? **[0.5 points]**
3. Consider an arbitrary **diagonal** matrix.
   What are the corresponding **eigenvalues and eigenvectors** of this matrix? **[0.5 points]**

Task 2 - A PCA example [4 points]

Consider a data set with two data points: $X = \begin{bmatrix} -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{d\times n} = \mathbb{R}^{3\times 2}$.

1. Use **standard PCA** to project the data onto a 1-dimensional subspace. Recall you have to do the following steps:
   (a) Compute the empirical covariance matrix $\hat{\Sigma} \in \mathbb{R}^{3\times 3}$ of the data. **[0.5 points]**
   (b) The first principal direction $\mathbf{w} \in \mathbb{R}^3, \|\mathbf{w}\| = 1$ is given by the eigenvector of S corresponding to the largest eigenvalue. **[1 point]**
   (c) Compute the projected data $H = \mathbf{w}^\top X$. **[0.5 points]**
2. Use **linear Kernel PCA** to obtain the same result. Recall you have to do the following steps:
   (a) Compute the kernel matrix $\mathbb{R}^{2\times 2} \ni K = X^\top X$ of the data. **[0.5 points]**

(b) Compute a linear combination of the data points $\alpha \in \mathbb{R}^2$ as the eigenvector of K corresponding to the largest eigenvalue. **[1 point]**

(c) Compute $\mathbf{w} = X\alpha$. You should obtain a scaled version of the standard PCA result. **[0.5 points]**

Task 3 - Covariance matrix and eigenvalues [3 points]

For each of the three scenarios below, sketch two 2-dimensional Gaussian data sets (one sketch with uncorrelated and one sketch with correlated features) whose distribution's covariance matrix has the following two eigenvalues:

1. $\lambda_1 = 1, \lambda_2 = 1$ **[1 point]**
2. $\lambda_1 = 1, \lambda_2 = 5$ **[1 point]**
3. $\lambda_1 = 1, \lambda_2 = 0$ **[1 point]**

If it is impossible to sketch, explain why.

Hint: Remember the following fact from the lecture. The variance of a data set projected onto a direction $\mathbf{w} \in \mathbb{R}^D, ||\mathbf{w}|| = 1$ can be computed as $\mathbf{w}^\top \Sigma \mathbf{w}$, where Σ denotes the data covariance matrix. If $\mathbf{w}$ is an eigenvector of Σ with corresponding eigenvalue λ, then $\mathbf{w}^\top \Sigma \mathbf{w} = \mathbf{w}^\top \lambda \mathbf{w} = \lambda \mathbf{w}^\top \mathbf{w} = \lambda$.

Task 4 - Covariance matrix and eigenvalues II [1.5 points]

1. Consider a square matrix $S \in \mathbb{R}^{d \times d}$, and an eigenvector $\mathbf{v}$ of S with corresponding eigenvalue λ, i.e. $S\mathbf{v} = \lambda \mathbf{v}$.

   Prove that $\mathbf{v}$ is also an eigenvector of the scaled matrix $S_\alpha := \alpha S$ for any $\alpha \in \mathbb{R}$. **[1 point]**

2. Give an example for an orthogonal matrix that is not a rotation matrix. **[0.5 point]**

3. Consider a centered data set $X \in \mathbb{R}^{d \times n}$ with corresponding covariance matrix $\Sigma = \frac{1}{N} X X^\top$, and an eigenvector $\mathbf{v}$ of Σ with corresponding eigenvalue λ, i.e. $\Sigma \mathbf{v} = \lambda \mathbf{v}$. Suppose we rotate the data, $X \mapsto UX$ where $U \in \mathbb{R}^{d \times d}$ is a rotation matrix ($UU^T = U^T U = I$).

   Prove that λ is an eigenvalue of the covariance matrix of the rotated data set, $_U = \frac{1}{N} UX(UX)^\top$. Furthermore show that the eigenvectors of the covariance matrix of the rotated data are just the eigenvectors of the original data, but rotated by U, too. **[*]**

Task 5 - Non-negative matrix factorization [2 points]

You apply NMF to a dataset $X \in \mathbb{R}^{4 \times 3}$.

After training, your reconstruction $\tilde{X} = W \cdot H = [\tilde{X}_1, \tilde{X}_2, \tilde{X}_3]$ looks like this

$$\tilde{X}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \tilde{X}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \qquad \tilde{X}_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

where corresponding encoding $H = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$ has the following values

$$\mathbf{h}_1 = [1, 0, 0]^\top \qquad \mathbf{h}_2 = [0, 1, 0]^\top \qquad \mathbf{h}_3 = [0, 0, 1]^\top.$$

Compute W.