Winter term 2024/25
U. von Luxburg
E. Günther/ K. Frohnapfel

Assignment 03

Mathematics for Machine Learning

Submission due Friday **08.11.24, 23:59** via Ilias

Justify all your claims.

Exercise 1 (**Symmetric matrices**, 1+1+3 points)**.**
Prove the following statements.

a) Any orthogonal, symmetric matrix $A \in \mathbb{R}^{n \times n}$ can only have the eigenvalues -1 or 1.

b) Consider a symmetric matrix $A \in \mathbb{R}^{n \times n}$ and an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and let $\tilde{A} = QAQ^T$. Then $\tilde{A}$ is again symmetric and has the same eigenvalues as A.

c) For any matrix $A \in \mathbb{R}^{n \times m}$ with $A^T A = I_m$ the matrix AA^T is the orthogonal projection onto the space range(A).

Exercise 2 (**Scalar products and norms**, 2+2+1 points)**.**
Consider a normed vector space $(V, \|\cdot\|)$. The parallelogram equality states that the norm is induced by a scalar product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ via $\|v\| = \sqrt{\langle v, v \rangle}$ if and only if for all $v, w \in V$ it holds that

$$\|v + w\|^2 + \|v - w\|^2 = 2 \left(\|v\|^2 + \|w\|^2 \right) \tag{1}$$

a) Prove the forward implication. That is, prove that Eq. (1) holds if the norm is given by a scalar product.

b) Let $p > 0$. Prove that there is a scalar product on $\mathbb{R}^2$ such that the associated norm is given by

$$\|x\| = (|x_1|^p + |x_2|^p)^{1/p}$$

for all $x \in \mathbb{R}^2$ if and only if $p = 2$.

c) Let $(V, \|\cdot\|)$ be a normed vector space, where the norm is given by a scalar product. Prove the Pythagorean theorem $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ for all orthogonal $u, v \in V$.

Exercise 3 (**Gram–Schmidt orthonormalization**, 3+2 points)**.**

a) Consider the subspace $V = \operatorname{span}\left(1, x, x^2, x^3\right) \subseteq \mathbb{R}^{\mathbb{R}}$ of polynomials with degree ≤ 3 with the scalar product $\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \mathrm{d}x$ for $f, g \in V$. Apply Gram-Schmidt to the basis $1, x, x^2, x^3$ in order to compute an orthonormal basis of V.
Hint: You do not need to show the steps of the computations of scalar products and norms, you can use a computer to compute the integrals. To check whether your solution is correct, simply verify whether your basis is orthonormal.

b) What happens if Gram-Schmidt is applied to a list of vectors $(v_1, \ldots, v_n)$ that is not linearly independent?

Exercise 4 (**Spectral clustering**, 1+1+1+2 points).
Let $W \in \mathbb{R}^{n \times n}$ be a symmetric matrix with non-negative entries and $D \in \mathbb{R}^{n \times n}$ the diagonal matrix which contains the row sums of W, that is, $d_{i,i} = \sum_{j=1}^{n} w_{i,j}$. Define $L := D - W$.

a) Prove that for all $x \in \mathbb{R}^n$ it holds

$$x^t L x = \frac{1}{2} \sum_{i,j=1}^{n} w_{i,j} (x_i - x_j)^2 \,.$$

b) Conclude that L is symmetric and positive semi-definite.

c) Show that the vector of constant ones $\mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$ is an eigenvector of L.

d) Solve the constrained minimization problem

$$\min_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} x^t L x \quad \text{subject to } \langle x, \mathbb{1} \rangle = 0 \,.$$

Note: This exercise introduces the basic principle behind spectral clustering. The matrix W is called the adjacency matrix and holds the edge weights of a graph as entries. D holds the node degrees on its diagonal and is therefore called the degree matrix and L is called the (unnormalized) graph Laplacian matrix. The minimization problem in d) is a relaxation of the min cut problem, which tries to find the minimal number of edges that you have to remove to separate a graph into two subgraphs.