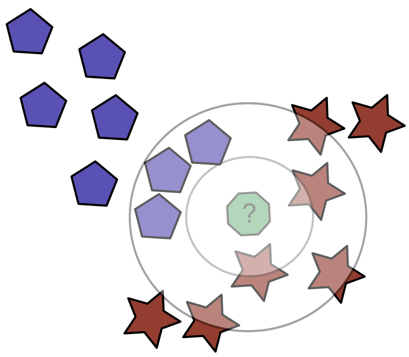
Mathematical foundations
for machine learning

Machine Learning
Institute of Software Engineering and
Theoretical Computer Science
Faculty IV, TU Berlin
Prof. Dr. Klaus-Robert Müller

Exercises

Exercise 1

A set of training data is given (red and blue). Using kNN to classify the previously unknown data point (green). Which label will be assigned for $k \in 1, 2, 5$?



Exercise 2

1. What is the dot product of the following vectors $\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix}$?

   ☐ 3 ☐ 5 ☐ 7

2. Which of the following mappings does not define an inner product on the vector space $L_2(\mathbb{R})$ the square-integrable functions on $\mathbb{R}$?

   ☐ $\langle f, g \rangle := \int_{-\infty}^{+\infty} f(x)g(x)dx$ for all $f, g \in L_2(\mathbb{R})$

   ☐ $\langle f, g \rangle := \int_{-\infty}^{+\infty} \exp(-x^2)f(x)g(x)dx$ for all $f, g \in L_2(\mathbb{R})$

   ☐ $\langle f, g \rangle := \int_{-\infty}^{+\infty} (f(x) + g(x))dx$ for all $f, g \in L_2(\mathbb{R})$

Exercise 3

Let $\langle \cdot, \cdot \rangle$ be the dot product on $\mathbb{R}^n$.

1. Show that the mapping

$$k(\cdot, \cdot) : \mathbb{R}^2 \times \mathbb{R}^2 \ni (\mathbf{x}, \mathbf{y}) \mapsto (\langle \mathbf{x}, \mathbf{y} \rangle)^2 \in \mathbb{R}$$

   does not define an inner product on $\mathbb{R}^2$.

2. Considering the mapping $\Phi : \mathbb{R}^2 \to \mathbb{R}^3$ with

$$\Phi : \mathbb{R}^2 \ni (x_1, x_2) \mapsto (x_1^2, x_2^2, \sqrt{2}x_1x_2) \in \mathbb{R}^3.$$

Show:

$$\langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle = k(\mathbf{x}, \mathbf{y}).$$

Exercise 4

The 1-Norm is often used in optimization to yield sparse solutions (many zero elements). In this exercise we want to illustrate this.
We are looking for a vector $\mathbf{w} = (x, y)^T$, that solves the optimization problem

$$\max_{\mathbf{w}} f(\mathbf{w}) \qquad s.t \quad \|\mathbf{w}\| = 1$$

We consider $f(x, y) = 0.5 \cdot x + y$ and want to compare the solutions to the optimization problem for the 1-Norm and the 2-Norm.

1. In the xy-plane sketch all points for which the 2-Norm is 1.

$$C_2 := \{(x, y)^T \in \mathbb{R}^2 | \; \|(x, y)^T\|_2 = x^2 + y^2 = 1\}.$$

2. Sketch all points for which the 1-Norm is 1.

$$C_1 := \{(x, y)^T \in \mathbb{R}^2 | \; \|(x, y)^T\|_1 = |x| + |y| = 1\}.$$

3. Sketch the contour lines $c = f(x, y)$ for $c = 0.5$, $c = 1$ and $c = 1.1$.

4. Geometrically, where does the solution of the optimization problem lie for the 1-Norm, and where for the 2-Norm?

Exercise 5

1. Let $A, B \in \mathbb{R}^{2\times 3}$ be 2×3- matrices, then
   - □ $A + B \in \mathbb{R}^{2\times 3}$
   - □ $A + B \in \mathbb{R}^{4\times 6}$
   - □ $A + B \in \mathbb{R}^{4\times 9}$

2. For $A \in \mathbb{R}^{m\times n}$:
   - □ A has m rows and n columns
   - □ A has n rows and m columns
   - □ The rows of A have the length m and the columns of A have length n.

3. For every square $n \times n$ Matrix A:
   - □ $\mathrm{rank}(A) = n \Rightarrow A$ is invertible, but there are invertible A with $\mathrm{rank}(A) \neq n$
   - □ A is invertible $\Rightarrow \mathrm{rank}(A) = n$, but there is an A with $\mathrm{rank}\, A = n$, which is not invertible.
   - □ $\mathrm{rank}\, A = n \Leftrightarrow A$ is invertible

4. The rank of the matrix $\begin{pmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{pmatrix}$ is

☐ 1 ☐ 3 ☐ 4

5. For every symmetric, invertible matrix $A \in \mathbb{R}^{n\times n}$ and vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

   ☐ $\langle A\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, A\mathbf{w}\rangle$
   ☐ $\langle A\mathbf{v}, A\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$
   ☐ $\langle A\mathbf{v}, \mathbf{w}\rangle = \left\langle \mathbf{v}, A^{-1}\mathbf{w}\right\rangle$

6. Which of the following matrices is orthogonal?

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

7. Let $\langle \cdot, \cdot\rangle$ be the dot product on $\mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$ an arbitrary square matrix with full rank. Which of the following mappings from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$ defines an inner product on $\mathbb{R}^n$?

   ☐ $f(\mathbf{x}, \mathbf{y}) := \langle A\mathbf{x}, \mathbf{y}\rangle$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
   ☐ $g(\mathbf{x}, \mathbf{y}) := \langle A\mathbf{x}, A\mathbf{y}\rangle$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
   ☐ $h(\mathbf{x}, \mathbf{y}) := \left\langle A\mathbf{x}, A^T\mathbf{y}\right\rangle$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

8. Let A be an invertible matrix and λ an eigenvalue of$A$.

   ☐ λ is an eigenvalue of A^{-1}
   ☐ $-\lambda$ is an eigenvalue of A^{-1}
   ☐ $1/\lambda$ is an eigenvalue of A^{-1}

9. Let $A \in \mathbb{R}^{n\times n}$ be a square matrix and $p \in \mathbb{N}$. Then A and A^p will have

   ☐ the same eigenvectors and eigenvalues
   ☐ the same eigenvalues, but not necessarily the same eigenvectors
   ☐ the same eigenvectors, but not necessarily the same eigenvalues.

10. Which of the following matrices are positive definite?

    ☐ $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ ☐ $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ☐ $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$

11. Let $A, B \in \mathbb{R}^{n\times n}$ and let $A = ZBZ^{-1}$ for an invertible matrix $Z \in \mathbb{R}^{n\times n}$. Then A and B will have

    ☐ the same eigenvectors and eigenvalues.
    ☐ the same eigenvalues, but not necessarily the same eigenvectors
    ☐ the same eigenvectors, but not necessarliy the same eigenvalues

Exercise 6

Let $R \in \mathbb{R}^{n\times n}$ be an orthogonal matrix that is to say $RR^T = R^TR = I$. Show that the dot product of two vectors is invariant to a multiplication with R which means that for all $\mathbf{x}, \mathbf{y} \in R^n$ it is the case that

$$\langle R\mathbf{x}, R\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle .$$

Exercise 7

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be eigenvectors of A to the eigenvalues $\lambda, \mu \in \mathbb{R}$, with $\lambda \neq \mu$. Show: $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, that is to say $\langle \mathbf{v}, \mathbf{w}\rangle = 0$.
hint: $\lambda \langle \mathbf{v}, \mathbf{w}\rangle = \ldots$.

Exercise 8

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Show that if A is positive semidefinite, then there is a real square matrix B with $A = BB$.
Hint: Let $A = U\Lambda U^T$ be the eigendecomposition of A. Construct B from U and Λ.

Exercise 9

We consider two functions f and g which transform an input vector $\mathbf{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$ into a scalar: $f(\mathbf{x}) = \mathbf{u}^\top \mathbf{x},\ \mathbf{u} = (u_1, \ldots, u_d)^\top \in \mathbb{R}^d$ and $g(\mathbf{x}) = \mathbf{x}^\top \mathbf{x}$.

1. Compute the partial derivative of f and g with respect to one entry x_j ($j \in \{1, 2, \ldots, d\}$)

   (a) $\frac{\partial f(\mathbf{x})}{\partial x_j} =$

   (b) $\frac{\partial g(\mathbf{x})}{\partial x_j} =$

2. Compute the gradient $\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_d}\right)^\top$ for f and g.

   (a) $\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_d}\right)^\top =$

   (b) $\nabla g(\mathbf{x}) = \left(\frac{\partial g(\mathbf{x})}{\partial x_1}, \ldots, \frac{\partial g(\mathbf{x})}{\partial x_d}\right)^\top =$