Machine Learning Group
Fakulty IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Week 1: Euclidean Vector Spaces

Exercise 1

1. The Cartesian product $A_1 \times \cdots \times A_n$ of sets $A_1, \ldots, A_n$ is defined as

   ☐ $A_1 \times \cdots \times A_n \coloneqq A_1 \setminus (A_2 \cup \ldots \cup A_n)$
   ☒ $A_1 \times \cdots \times A_n \coloneqq \{(a_1, \ldots, a_n) \mid a_1 \in A_1, \ldots, a_n \in A_n\}$ set of all ordered pairs
   ☐ $A_1 \times \cdots \times A_n \coloneqq \{a_1 \cdot a_2 \cdot \ldots \cdot a_n \mid a_1 \in A_1, \ldots, a_n \in A_n\}$

2. Let $n \in \mathbb{N} \setminus \{0\}$. Then $\mathbb{R}^n$ contains

   ☐ n real numbers ☒ n-tuples of real numbers ☐ n-tuples of vectors

3. Which of the following sets together with standard addition and scalar multiplication does not form a real vector space?

   ☒ The set of integers $\mathbb{Z}$
   ☐ The set of complex numbers $\mathbb{C}$
   ☐ The set of real-valued, continuous functions $\{f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ continous }\}$

   Solution: Integers not closed under scalar multiplication with reals.

4. Scalar multiplication in a real vector space $V = (\mathcal{V}, +, \cdot)$ is given by a mapping

   ☐ $\mathcal{V} \times \mathcal{V} \to \mathbb{R}$ ☒ $\mathbb{R} \times \mathcal{V} \to \mathcal{V}$ ☐ $\mathbb{R} \times \mathcal{V} \to \mathbb{R}$

5. How many vector subspaces does $\mathbb{R}^2$ have?

   ☐ two: $\{\mathbf{0}\}$ and $\mathbb{R}^2$
   ☐ four: $\{\mathbf{0}\}$, $\mathbb{R} \times \{0\}$, $\{0\} \times \mathbb{R}$ (the axes), $\mathbb{R}^2$
   ☒ infinitely many

6. Which of the following subsets of $\mathbb{R}^2$ is not a vector subspace?

   ☐ $\{\mathbf{0}\}$ ☐ $\{\boldsymbol{x} \in \mathbb{R}^2 \mid x_1 = 2x_2\}$ ☒ $\{\boldsymbol{x} \in \mathbb{R}^2 \mid x_1 = x_2 + 1\}$

   Solution: Because of scalar multiplication: $(x_1, x_2) \mapsto (\lambda x_1, \lambda x_2), \lambda x_1 = \lambda(x_2 + 1) \neq \lambda x_2 + 1$

7. For which of the following objects does it make sense to say that they are "linearly independent"?

   ☒ Elements of a vector space $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathcal{V}$
   ☐ Real numbers $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$
   ☐ The linear combination $\lambda_1 \boldsymbol{v}_1 + \ldots + \lambda_n \boldsymbol{v}_n$

8. Which of the following vectors form a basis of $\mathbb{R}^2$?

   ☒ $\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ ☐ $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ ☐ $\begin{bmatrix} 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

9. What is the scalar product of the vectors $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}$?

   ☒ 3 ☐ 5 ☐ 7

10. Which of the following statements is false?

    ☐ For any Euclidean vector space $(V, \langle\cdot,\cdot\rangle)$ the function $\|\cdot\|$ defined by $\|\boldsymbol{v}\| \coloneqq \sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}$ $(\boldsymbol{v} \in \mathcal{V})$ satisfies the properties of a norm.
    ☒ For any Euclidean vector space $(V, \langle\cdot,\cdot\rangle)$ it holds that $\langle \boldsymbol{v}, \boldsymbol{w}\rangle \geq 0$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.
    ☐ Let M be a subspace of an Euclidean vector space $(V, \langle\cdot,\cdot\rangle)$. Then

    $$\{\boldsymbol{v} \in \mathcal{V} \mid \forall \boldsymbol{u} \in M : \boldsymbol{v} \perp \boldsymbol{u}\}$$

    is also subspace of V.

Solution: Why 3 is true: If two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ are orthogonal to $\boldsymbol{u}$, so is their sum and any rescaling.

Exercise 2

Consider $(\mathbb{R} \setminus \{-1\}, \star)$, where $a \star b := ab + a + b$ with $a, b \in \mathbb{R} \setminus \{-1\}$.

1. *Show* that $(\mathbb{R} \setminus \{-1\}, \star)$ is an Abelian group.

   Solution:

   (a) associativity:

   $$\begin{aligned} a \star (b \star c) &= a \star (bc + b + c) = a(bc + b + c) + a + (bc + b + c) = abc + ab + ac + a + bc + b + c \\ (a \star b) \star c &= (ab + a + b) \star c = (ab + a + b)c + (ab + a + b) + c = abc + ac + bc + ab + a + b + c \end{aligned}$$

   (b) identity element:
   $$a \star e = a \Leftrightarrow ae + a + e = a \Leftrightarrow (a+1)e = 0 \Leftrightarrow e = 0$$

   (c) inverse element: (we only show $a \star a^{-1} = e$, the other condition follows from commutativity)
   $$a \star a^{-1} = e \Leftrightarrow aa^{-1} + a + a^{-1} = 0 \Leftrightarrow a^{-1} = -\frac{a}{a+1}$$

   (d) commutativity:
   $$a \star b = ab + a + b = ba + b + a = b \star a$$

2. *Solve* $3 \star x \star x = 15$ in the Abelian group $(\mathbb{R} \setminus \{-1\}, \star)$.

   Solution:

   $$3 \star x \star x = (3x + 3 + x) \star x = (3x + 3 + x)x + (3x + 3 + x) + x = 4x^2 + 8x + 3 = 15 \Leftrightarrow x = 1$$

3. Is $(\mathbb{R} \setminus \{-1\}, +, \star)$ a field? Justify!

   Solution: No, because it is not distributive: $a \star (b + c) \neq a \star b + a \star c$.

Exercise 3

Show: If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ form an orthonormal basis of a Euclidean vector space $(V, \langle \cdot, \cdot \rangle)$, the following holds for all $\boldsymbol{x} \in \mathcal{V}$:

$$\boldsymbol{x} = \sum_{i=1}^{n} \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle \, \boldsymbol{v}_i$$

Hint: Establish first that $\boldsymbol{x}$ can be expressed as $\boldsymbol{x} = \lambda_1 \boldsymbol{v}_1 + \ldots + \lambda_n \boldsymbol{v}_n$. Then show that $\langle \boldsymbol{x}, \boldsymbol{v}_i \rangle = \lambda_i$ for all $1 \leq i \leq n$.

Solution: For all $1 \leq i \leq n$ it holds that

$$\langle \boldsymbol{x}, \boldsymbol{v}_i \rangle = \langle \lambda_1 \boldsymbol{v}_1 + \ldots + \lambda_n \boldsymbol{v}_n, \boldsymbol{v}_i \rangle = \lambda_1 \langle \boldsymbol{v}_1, \boldsymbol{v}_i \rangle + \ldots + \lambda_n \langle \boldsymbol{v}_n, \boldsymbol{v}_i \rangle = \lambda_i \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle = \lambda_i$$

Exercise 4

Let $(V, \langle \cdot, \cdot \rangle)$ be a Euclidean vector space and U an r-dimensional vector subspace $\mathcal{U} \subseteq \mathcal{V}$ with orthonormal basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$. The orthogonal projection of a vector $\boldsymbol{v} \in \mathcal{V}$ onto $\mathcal{U}$ is given by

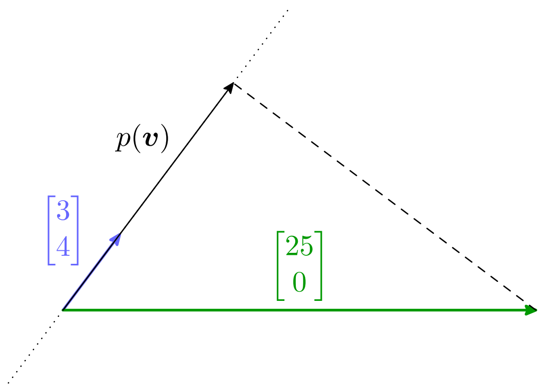
$$p(\boldsymbol{v}) := \sum_{i=1}^{r} \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle \, \boldsymbol{u}_i$$

1. *Compute* the orthogonal projection of the vector $(25, 0)^\top$ onto the subspace spanned by the vector $(3, 4)^\top$. Visualize the subspace and the projection in a drawing with both vectors.

2. Let $\boldsymbol{x} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$. *Show:*
   $$p(\lambda \boldsymbol{x}) = \lambda p(\boldsymbol{x}).$$

3. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$. *Show:*
   $$p(\boldsymbol{x} + \boldsymbol{y}) = p(\boldsymbol{x}) + p(\boldsymbol{y}).$$

Solution:



1. $\left\| \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\| = \sqrt{25} = 5$, therefore $\frac{1}{5} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ orthonormal basis. Then
$$\begin{bmatrix} 25 \\ 0 \end{bmatrix} \mapsto \left\langle \begin{bmatrix} 25 \\ 0 \end{bmatrix}, \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} \right\rangle \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} = 15 \cdot \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} = \begin{bmatrix} 9 \\ 12 \end{bmatrix}$$

2. $$p(\lambda \boldsymbol{x}) = \sum_{i=1}^{r} \langle \lambda \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i = \lambda \sum_{i=1}^{r} \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i = \lambda p(\boldsymbol{x}).$$

3. $$p(\boldsymbol{x} + \boldsymbol{y}) = \sum_{i=1}^{r} \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i = \sum_{i=1}^{r} \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i + \sum_{i=1}^{r} \langle \boldsymbol{y}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i = p(\boldsymbol{x}) + p(\boldsymbol{y}).$$

Exercise 5

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

1. *Show* that the mapping
$$\begin{aligned} k \colon \mathbb{R}^2 \times \mathbb{R}^2 &\to \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 \end{aligned}$$
does not define a scalar product on $\mathbb{R}^2$.

2. Consider the mapping $\Phi \colon \mathbb{R}^2 \to \mathbb{R}^3$ with
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2} \cdot x_1 x_2 \end{bmatrix}.$$
Show:
$$\langle \Phi(\boldsymbol{x}), \Phi(\boldsymbol{y}) \rangle = k(\boldsymbol{x}, \boldsymbol{y})$$

Solution:

1. Violation of bilinearity:
$$k(\lambda \boldsymbol{x}, \boldsymbol{y}) = \langle \lambda \boldsymbol{x}, \boldsymbol{y} \rangle^2 = \lambda^2 \langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 = \lambda^2 k(\boldsymbol{x}, \boldsymbol{y})$$

2. $$\begin{aligned} \langle \Phi(\boldsymbol{x}), \Phi(\boldsymbol{y}) \rangle = \left\langle \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2} \cdot x_1 x_2 \end{bmatrix}, \begin{bmatrix} y_1^2 \\ y_2^2 \\ \sqrt{2} \cdot y_1 y_2 \end{bmatrix} \right\rangle &= x_1^2 y_1^2 + x_2^2 y_2^2 + 2 x_1 x_2 y_1 y_2 \\ &= (x_1 y_1 + x_2 y_2)^2 = k(\boldsymbol{x}, \boldsymbol{y}) \end{aligned}$$

Exercise 6

The 1-norm is often used for finding sparse solutions to an optimization problem (vectors or matrices with many entries equal to zero). This will be demonstrated in the following exercise:
We are looking for a vector $\boldsymbol{w} = (x, y)^\top \in \mathbb{R}^2$, which solves the optimization problem

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

Consider $f(x, y) = \frac{1}{2} x + y$ and compare the solutions to this optimization problem for the 1- and the 2-norm.

1. Draw the set of all points on the x-y-plane that have 2-norm equal to 1 (i.e. the ℓ_2 unit circle).

$$C_2 := \left\{ \boldsymbol{w} \in \mathbb{R}^2 \mid \|\boldsymbol{w}\|_2 = \sqrt{x^2 + y^2} = 1 \right\}.$$

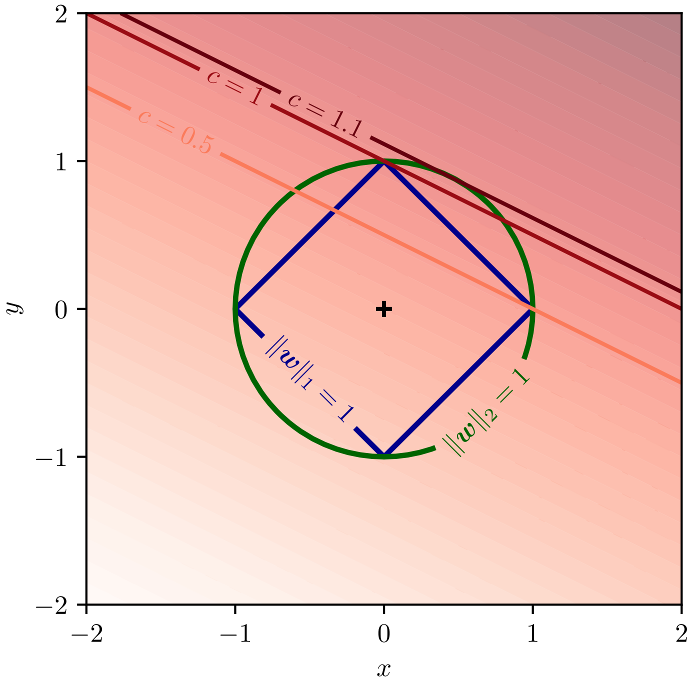
2. Draw the set of all points on the x-y-plane that have 1-norm equal to 1 (i.e. the ℓ_1 unit circle).

$$C_1 := \left\{ \boldsymbol{w} \in \mathbb{R}^2 \mid \|\boldsymbol{w}\|_1 = |x| + |y| = 1 \right\}.$$

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

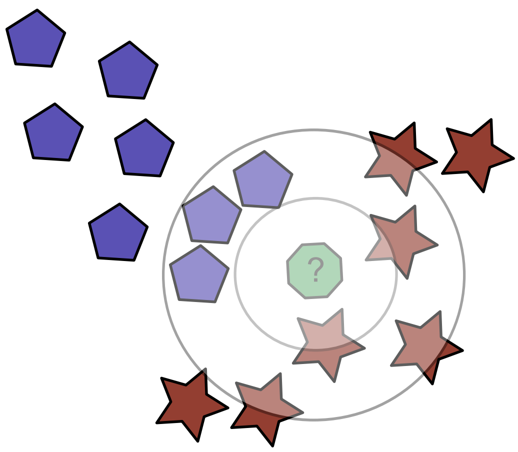
4. Where in your drawing does the solution to the optimization problem lie for the 1-norm? Where does it lie for the 2-norm?

Solution:



Exercise 7

Given training data (red and blue) a new data point (green) should be classified using the k-NN algorithm. Which label does the new data point receive for $k \in 1, 2, 5$?



Solution: 1: red
2: red
5: blue