Exercise Sheet 1

May 2025

Exercise 1: Linear Decision Boundaries

Task 1

The binary linear classifier is defined by the function:

$$y = \text{sign}(w^\top x + b)$$

where:

$$w = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad b = -1$$

Let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Then the classifier becomes:

$$y = \text{sign}(1 \cdot x_1 - 2 \cdot x_2 + 3 \cdot x_3 - 1)$$

The decision boundary is defined by the set of points where the classifier output changes, i.e., where:

$$w^\top x + b = 0$$
$$x_1 - 2x_2 + 3x_3 - 1 = 0$$
$$x_1 - 2x_2 + 3x_3 = 1$$

In an n-dimensional space, a *hyperplane* is defined by a linear equation of the form:

$$w^\top x + b = 0$$

This equation defines a flat subspace of dimension $n - 1$.

In our case:

- The feature space is 3-dimensional: (x_1, x_2, x_3)
- The equation $x_1 - 2x_2 + 3x_3 = 1$ is linear in the variables

Therefore, it defines a 2D plane in 3D space — a **hyperplane** — which separates the space into two regions:

Regions where y = +1 and -1

Task 2

The normal vector to the decision boundary is:

$$\mathbf{n} = w = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$$

It has the following properties:

- It is **perpendicular** to the decision boundary (which is a plane in 3D space).
- It **points in the direction of increasing values** of the function $w^\top x + b$. That is:
  - Moving in the direction of $\mathbf{n}$ increases $w^\top x + b$.
  - Moving opposite to $\mathbf{n}$ decreases $w^\top x + b$.
- It determines the **orientation** of the decision boundary.

Task 3

a

$$y = \text{sign}(w^\top x + b)$$

$$d(x) = \frac{w^\top x + b}{\|w\|}$$

where:

$$w = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad b = -1 \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$w^\top x = (1)(1) + (-2)(1) + (3)(1) = 1 - 2 + 3 = 2$$

$$w^\top x + b = 2 + (-1) = 1$$

$$\|w\| = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{1 + 4 + 9} = \sqrt{14}$$

$$d(x) = \frac{1}{\sqrt{14}}$$

b

- **If** $d(x) > 0$**:**
  The point lies on the **positive side** of the hyperplane — the side toward which the normal vector w points.
  $\Rightarrow$ The classifier would predict: $y = +1$
- **If** $d(x) < 0$**:**
  The point lies on the **negative side** of the hyperplane — the side opposite to w.
  $\Rightarrow$ The classifier would predict: $y = -1$
- **If** $d(x) = 0$**:**
  The point lies **exactly on the decision boundary**.
  $\Rightarrow$ The classifier is undecided, since $\text{sign}(0)$ may be defined as 0 or may require special handling.

c

The further the point x is from the decision boundary (i.e., the larger $|d(x)|$), the more confident the model is in its prediction.

- A large positive distance $d(x) \gg 0$
  $\Rightarrow$ Strongly confident prediction of class $+1$.
- A large negative distance $d(x) \ll 0$
  $\Rightarrow$ Strongly confident prediction of class -1.
- A distance close to zero ($d(x) \approx 0$)
  $\Rightarrow$ The point lies near the decision boundary
  $\Rightarrow$ Low confidence, possibly due to noisy or ambiguous input.

d

The equation of the decision boundary is:

$$w^\top x + b = 0$$

The orthogonal projection of a point x onto the hyperplane is given by:

$$x_{\text{proj}} = x - \frac{w^\top x + b}{\|w\|^2} w$$

where:

$$w = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad b = -1, \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$w^\top x = (1)(1) + (-2)(1) + (3)(1) = 1 - 2 + 3 = 2$$

$$w^\top x + b = 2 + (-1) = 1$$

$$\|w\| = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{1+4+9} = \sqrt{14}$$

Now, we compute the orthogonal projection:

$$x_{\text{proj}} = x - \frac{w^\top x + b}{\|w\|^2} w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{14} \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$$

$$x_{\text{proj}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{14} \\ -\frac{2}{14} \\ \frac{3}{14} \end{pmatrix}$$

$$x_{\text{proj}} = \begin{pmatrix} 1 - \frac{1}{14} \\ 1 + \frac{2}{14} \\ 1 - \frac{3}{14} \end{pmatrix}$$

$$x_{\text{proj}} = \begin{pmatrix} \frac{13}{14} \\ \frac{16}{14} \\ \frac{11}{14} \end{pmatrix}$$

e

The orthogonal projection of a point helps us measure how close the point is to the decision boundary.

The support vectors are the points that are closest to the boundary and determine where the boundary should be placed to maximize the margin.

SVM aims to adjust the boundary so that the margin is as wide as possible, and this margin is defined by the orthogonal projection of the support vectors.