ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Pattern Recognition and Machine Learning

Exercises Week 1

Lecturer: Matthias Seeger
Assistants: Abdallah Elguindy, Mathias Humbert, Young Jun Ko, Xinchao Wang

Exercise 1

Suppose that inputs into a binary linear classifier are $\boldsymbol{x} = [x_1, x_2]^T \in \mathbb{R}^2$. The decision boundary goes through $\boldsymbol{p}_1 = [1, 0]^T$ and $\boldsymbol{p}_2 = [5, 8]^T$.

1. Derive the slope a and offset b in the equation $x_2 = ax_1 + b$ which describes the decision boundary. Does the line go through the origin?

2. Determine a weight vector $\boldsymbol{w} = [w_1, w_2]^T$ so that the decision boundary is given by
$$\boldsymbol{w}^T\boldsymbol{x} + b = w_1x_1 + w_2x_2 + b = 0.$$
What is the relationship between $\boldsymbol{w}$ and $\boldsymbol{p}_2 - \boldsymbol{p}_1$?

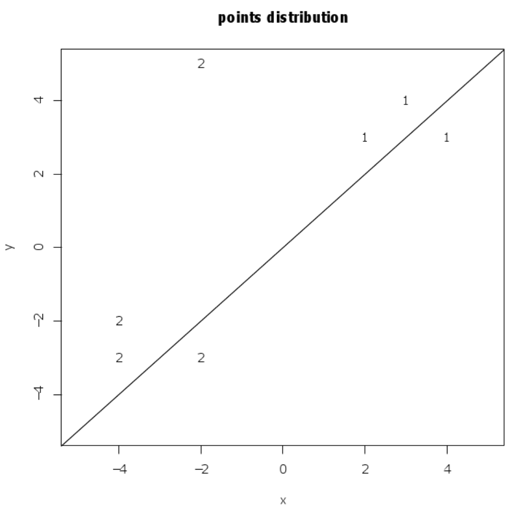
3. Give a formula for the linear binary classifier $f(\boldsymbol{x})$, mapping $\mathbb{R}^2$ to $\{-1, +1\}$, which has this decision boundary and classifies the origin $\mathbf{0} = [0, 0]^T$ as $+1$.
For each of the following points, determine whether $f(\boldsymbol{x})$ (a) classifies it as $+1$, (b) classifies it as -1, or (c) it lies on the decision boundary. $[5, 9]^T$, $[10, 18]^T$, $[0, -3]^T$.

4. **Bonus**: How would you construct the equation $\tilde{\boldsymbol{w}}^T\boldsymbol{x} + \tilde{b}$ for a line, so that the distance between the new line and each of $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ is 1? The answer is not unique, choose whatever is simplest for you.

Exercise 2

This question is about the perceptron algorithm, applied to the data given in the table and figure below. Use discriminant functions of the form
$$y(\boldsymbol{x}) = w_1x + w_2y,$$
without a bias term (all of them go through the origin), and (in constrast to the course notes) do not normalize the input vectors.

x	y	class
4	3	1
3	4	1
2	3	1
-2	-3	2
-4	-2	2
-4	-3	2
-2	5	2



1. Run the perceptron algorithm for at least 3 updates (unless it converges with less than 3 updates). Start with $\boldsymbol{w} = [1, -1]^T$. You may pick points in any order you like.

2. How would the algorithm behave if the point $[-2, 5]^T$ belonged to class 1?

3. What would happen if the point $[-4, -3]^T$ was labeled as 1?

Exercise 3

This exercise is about a linear discriminant based on a feature map $\boldsymbol{\phi}(\boldsymbol{x})$. Recall from the course notes that "linear" means *linear with respect to the weights*: the function can be nonlinear in the input $\boldsymbol{x} = [x_1, x_2]^T \in \mathbb{R}^2$. Consider the following discriminant function:

$$y(\boldsymbol{x}) = \frac{1}{4}x_1^2 + \frac{1}{4}x_2^2 - \frac{1}{2}x_1 - \frac{3}{2}x_2 + \frac{3}{2}.$$

1. Bring this function into the standard form $y(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}) + b$. What is $\boldsymbol{w}$, $\boldsymbol{\phi}(\boldsymbol{x})$, and b?

2. Draw the decision boundary and shade the decision regions $\mathcal{H}_{+1}$ and $\mathcal{H}_{-1}$ in your drawing.
   Hint: Try to bring $y(\boldsymbol{x})$ into the form $a\|\boldsymbol{x} - \boldsymbol{\mu}\|^2 + b$, where $\boldsymbol{\mu} = [\mu_1, \mu_2]^T$ is a vector.