

Exercises

§6.1: From Regression to Classification

Python: There is an optional starter script that supports Python-based problem(s) on this sheet. Use it as scaffolding: read the comments, complete the lines marked **TODO**, and compare the output with the mathematics requested in the problem(s).

Problem 1 Let $f(x) = -3 + 2x$. For $x = (0, 1, 2, 3)$ with labels $y = (-1, -1, +1, +1)$, compute the scores, predicted classes, functional margins, and classification error rate.

Problem 2 Consider the two-dimensional score $f(x_1, x_2) = -6 + 2x_1 + x_2$. (a) Find the decision boundary. (b) Give a normal vector to the boundary. (c) Classify the points $(1, 1)$, $(2, 3)$, $(4, 0)$.

Problem 3 Suppose $y = +1$. For scores $f = 0.2, 1, 4$, compute the zero-one loss and squared regression loss $(f - y)^2$. Explain why the two losses reflect different goals.

Problem 4 Suppose a classifier has $f(\mathbf{x}) = w_0 + \mathbf{v}^T \mathbf{x}$. Replace every coefficient by five times its value. What happens to (a) the decision boundary, (b) the predicted classes, (c) the functional margins $y_p f_p$, (d) the geometric margins $y_p f_p / \|\mathbf{v}\|$?

Problem 5 A classifier produces scores $(-2, 1, -0.5, 3)$ for observations with labels $(-1, +1, +1, +1)$. Compute the functional margins and the classification error rate. Which observation is misclassified?

Problem 6 Use Python and np to classify

```
scores = np.array([-1.2, 0.4, 2.0, -0.1, 0.8])  
y = np.array([-1, 1, 1, 1, -1])
```

using threshold 0. Compute the predicted labels and classification error rate.

Problem 7 Why is a score $f(\mathbf{x}) = 3.7$ not automatically a probability that the class is $+1$? List two properties a transformation of the score should have if it is to represent a probability.