Machine Learning **Universität Hamburg**
Spring 2024 Sören Laue

Exercise Sheet 3

Exercise 1

Consider linear least squares regression with a $\|.\|_2^2$-regularizer, i.e., we have the following regularized risk minimization problem

$$\min_w \frac{1}{2n} \|Xw - y\|_2^2 + \frac{\lambda}{2} \|w\|_2^2$$

This problem can be solved similarly to the linear least squares problem without the regularization term (of course, not exactly the same). The goal of this exercise is to come up with a way of solving it. For this, compute the gradient and set it to zero, similarly as we did in class for the case without the regularization term. What system of linear equations need to be solved?

Hint: We computed the gradient of $\|w\|_2^2$ already in class. Furthermore, remember that $w = \mathbb{I}w$ where $\mathbb{I}$ is the identity matrix.

Exercise 2

Implement a solver for the problem from Exercise 1. For this, extend the least squares regression solver from the first exercise sheet.

If you had trouble solving Exercise 1, you can for this time also make use of the corresponding solver from scikit-learn.

Please note: The rest of scikit-learn is still not allowed on this exercise sheet. Please implement the corresponding methods yourself.

Exercise 3

Extend your solver to also solve polynomial regression.

Exercise 4

Implement k-fold cross validation and use this with polynomial regression (use a polynomial of degree 6) on the provided data set. The provided data set consists of a training set and a test set (`dataset_poly_train.npy` and `dataset_poly_test.npy`).

Exercise 5

Extend your solution to also compute the regularization path and plot the training error and the validation error. Vary the regularization parameter λ for the regularization path in the interval $[10^{-9}, 10^3]$. Use the empirical error $\frac{1}{2n} \|Xw - y\|_2^2$ for n data points. Note, the error is scaled by the number of data points for which the error is computed such that the error is comparable for data sets of different size, like the training and the validation set.

Exercise 6

Pick the best λ^* from the regularization path and compute the corresponding best classifier w^* for the whole training set. Report training and test error.

Please turn in your solutions by Thursday, April 25th.