Exercise Sheet 3

Due: 15.11.2023, 10:00

Download the file **poly.csv** from ISIS. The file consists of two columns. The first column contains input examples x_i and the second column contains the corresponding outputs y_i.

Exercise 3.1

Assume we are given a training set $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^n \times \mathbb{R}$. Let $h_w(x) = w^\top \tilde{x}$, with $\tilde{x} = (1, x) \in \mathbb{R}^{n+1}$ and $w = (w_0, w_1, \ldots, w_n) \in \mathbb{R}^{n+1}$. Consider the following empirical error function:

$$E_m[h_w, \lambda] = \frac{1}{2m} \sum_{i=1}^{m} (h_w(x_i) - y_i)^2 + \frac{\lambda}{2m} \sum_{j=1}^{n} w_j^2$$

Write the error function in matrix notation. Derive the update rule (in matrix notation) for the gradient descent method with respect to the parameter w.

Exercise 3.2

Implement the gradient descent method for polynomial regression in one variable with L_2-regularization. Use matrix and vector operations instead of loops where possible.

Exercise 3.3

Apply the polynomial regression implemented in Exercise 3.2 to the poly-data. Dispense with the regularization here. Fit polynomials of order k = 0, 1, ..., 6 and print the learned weights. Plot the MSE in dependence of the order *k.*

Exercise 3.4

Apply the polynomial regression with L_2-regularization to the poly-data. For this, fit a polynomial of order k =6 and use the following regularization parameters:

$$\lambda \in \{0, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10\}.$$

Plot the MSE in dependence of the regularization parameter. For the sake of clarity, the x-axis should be scaled logarithmically. Create a plot showing the data points and for each $\lambda \in \{0, 0.01, 0.1, 1, 10\}$ the fitted polynomial. Moreover, print the learned weights for $\lambda \in \{0, 0.01, 0.1, 1, 10\}$. Discuss your results.

Exercise 3.5* (1 Bonus point)

Let $X \in \mathbb{R}^{m \times n+1}$ be an augmented data matrix and $y \in \mathbb{R}^m$ the output vector.

- Show that Multiple Linear Regression has a unique solution if $\text{rank}(X) = n + 1$.
- State a necessary condition on m for having a unique solution.
- What if $\text{rank}(X) < n + 1$? Does a solution exist? Is the solution unique? Does it depend on y? Prove your claims.

To solve this task, you can use the following basic results from linear algebra:

1) $\text{rank}(X) = \text{rank}(X^\top X) = \text{rank}(X^\top)$
2) $\text{im}(X^\top) = \text{im}(X^\top X)$
3) A system of linear equations $Ax = b$ with $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n$ has
   a) a unique solution if $\text{rank}(A) = \text{rank}([A|b]) = \text{n}$
   b) infinitely many solutions if $\text{rank}(A) = \text{rank}([A|b]) < n$
   c) no solution if $\text{rank}(A) < \text{rank}([A|b])$,

   where $[A|b] \in \mathbb{R}^{n \times n+1}$ is the concatenated matrix composed of A and b.