Exercise 7: Tikhonov Regularization for Logistic Regression

Lecture Information Processing and Communication

Jörn Anemüller, June 2022

We request that you submit your solution until Monday 2022-06-06, 23:59h, by uploading to your group's exercise folder on cs.uol.de. If you cannot make it by Monday, you may still submit your solution until Tuesday 2022-06-07, 23:59h. (We have some travel duties and hence would appreciate the Monday submission date.) You may submit your solutions in groups of at most two students. You are free to write your code in matlab or in python (but we provide the example functions in matlab only).

Summary of exercise 7

The sole part of this exercise is to implement the "bonus" question 9 of the previous exercise sheet (Ex05-06).

9. Add a regularization term to the logistic regression gradient and perform gradient descent optimization for different parameter values of the regularization parameter λ

Regularization based on the l2-norm of the weight vector $\mathbf{w}$ essentially adds the l2-norm of the weight vector to loss function $L(\mathbf{w})$ which results in the partial derivative (i.e., the gradient's components) of

$$\frac{\partial}{\partial w_i} L_{\text{reg}}(\mathbf{w}) = \sum_{n=1}^{N} (\hat{y}^{(n)} - y^{(n)})\, x_i^{(n)} + \frac{\lambda}{N} w_i,$$

where

$$L_{\text{reg}}(\mathbf{w}) = \frac{1}{2N} \sum_{n=1}^{N} (\hat{y}^{(n)} - y^{(n)})^2 + \frac{\lambda}{2N} \sum_{j=1}^{D} w_j.$$

Note that this expression should be used only for $i = 1, \ldots, D$ but not for the bias-term w_0. For hints, refer to, e.g., `https://towardsdatascience.com/implement-logistic-regression-with-l2-regularization-from-`

Rerun your logistic regression algorithm from above with regularization for several values of the regularization parameter λ.

Plot the "digit"-image-representation corresponding to the weight vector and observe how it changes across the different λ-values.