

Exercise 8: Gradient Descent

Lecture Information Processing and Communication

Jörn Anemüller, June 2022

Submit solutions until Tuesday 2022-06-14, 23:59h, by uploading to your group's exercise folder on cs.uol.de. You may submit your solutions in groups of at most two students.

Note that this exercise sheet is shorter than the others. You will receive the next programming exercise by Friday, June 10th, for which you will have time to finish until the week after.

1. Gradient descent in pseudocode

Explain the gradient descent algorithm by providing a description of it in pseudocode.

2. Simplified example of logistic regression

In a simplified special case of the logistic regression method, the following assumptions are made:

- Input data $x^{(i)}$ are one-dimensional (i.e., scalar).
- A number of N training examples $x^{(i)}$ with $i = 1, \dots, N$ are given.
- A bias-term (the “dummy-one”) is *not* necessary.

For this case, write down the equation of the logistic regression model and the cross-entropy loss function.

Derive the gradient of the loss function.

3. Loss functions

Four different proposed loss-functions, L_1, L_2, L_3, L_4 , are given below, that depend input variables x_1 and x_2 .

For each of the four functions, perform these steps:

- Depict the function with contour lines qualitatively (no beauty contest!).
- Mark minima and maxima of each function.
- Decide whether the function is suitable as a loss function for some learning algorithm.

1. $L_1 = x_1 + x_2$

2. $L_2 = |x_1| + |x_2|$

3. $L_3 = x_1^2 + x_2^2$

4. $L_4 = \cos(x_1) + \sin(x_2)$