– Nov. 26, 2023

Modern Optimization Techniques,
Prof. Lars Schmidt-Thieme, Nourhan Ahmed
Exercise Sheet 5

1/1

Deadline: **Sun Dec 03, 2023, 8:00 am**   Submit single unzipped PDF file on learn-web course "SoSe 2021: 3104 Modern Optimization Techniques

Instructions

Please following these instructions for solving and submitting the exercise sheet.

1. Student should clearly write his/her name, matriculation number and tutorial group number (i.e. "Group 1: Tuesday Tutorial", "Group 2: Wednesday Tutorial").

2. The submission should be made before the deadline, only through learnweb to your group submission link.

3. Should be submitted as a single unzipped PDF file on learn-web course "SoSe 2023: 3104 Modern Optimization Techniques".

4. Each student must submit an individual solution in-order to be eligible for bonus points.

5. Group submission are acceptable but will not contribute towards bonus points.

1 Newton Method (10P)   (10 points)

a) In your own words, describe what is the intuition of using the Newton step for function minimization. You can use a sketch. (2pts)

b) For the following equations, compute their derivatives and second derivatives, write down the Newton Update Formula and execute 4 iterations of the Newton Method. Discuss what is happening.

$$f_1 : \mathbb{R} \longrightarrow \mathbb{R}$$
$$f_1(x) = x^3 - 2x - 5 \qquad \text{for an initial } x = 8 \text{ and } x = -10 \text{ (3pts)}$$

$$f_2 : \mathbb{R} \longrightarrow \mathbb{R}$$
$$f_2(x) = 3x^{1/3} \qquad \text{for an initial } x = -0.5 \text{ and } x = 1 \text{ (3pts)}$$

c) In which cases the Newton step can overshoot? (2pts)

2 Newton Method for ML problems (10P)   (10 points)

a) The loss function of the linear Regression has the following form:

$$\mathcal{L}(X, \beta, Y) = \sum_{i=1}^{m} (x_i\beta - y_i)^2$$

  – Compute the Hessian $\nabla_\beta^2 L(X, \beta, Y)$. (2pts)
  – Does it make sense to use the Newton minimization algorithm here? Discuss your answer. (2pts)

b) The loss function of the logistic regression has the following form:

$$\mathcal{L}(X, \beta, Y) = -\sum_{i=1}^{m} y_i log(\sigma(x_i\beta)) + (1 - y_i)log(1 - \sigma(x_i\beta))$$

  – Compute the Hessian $\nabla_\beta^2 L(X, \beta, Y)$. (4pts)
  – Does it make sense to use the Newton minimization algorithm here? Discuss your answer. (2pts)