Exam Optimization

Kasper Rosenkrands

Fall 2019

Contents

1 Introduction

This is a collection of the possible exam subjects at the optimization exam of the 2019 fall semester at Aalborg University.

2 Line search

Throughout this exercise we will utilize the `cars` dataset and we will refer to speed by s and distance by d.

2.1 Exercise 1: Gradient descent

We want to fit a straight line of the form $m(s) = a + b \cdot s$ to the data. We want to determine a and b. One way is to minimise the objective function given by

$$f(a, b) = \frac{1}{n} \sum_{i=1}^{n} f_i(a, b), \tag{1}$$

where

$$f_i(a, b) = (m(s_i) - d_i)^2. \tag{2}$$

2.1.1 What is the gradient of f?

We find the gradient by differentiating the function given by the following, first w.r.t. a and then b

$$f(a, b) = \frac{1}{n} \sum_{i=1}^{n} (m(s_i) - d_i)^2 \tag{3}$$

The gradient becomes

$$\nabla f(a, b) = \left[\frac{2}{n} \sum_{i=1}^{n} m(s_i) - d_i, \quad \frac{2}{n} \sum_{i=1}^{n} (m(s_i) - d_i) s_i \right] \tag{4}$$

2.1.2 Implement gradient descent and then use it to find the best straight line

2.1.2.1 What is meant by *the best* straight line in relation to the objective function above

2.1.2.2 Discuss different ways to determine the step sizes

2.1.3 Try with different ways to choose step sizes and illustrate it (including plotting the objective function and the iterates, $\{x_k\}_k$)

2.1.4 Show some iterates in a plot showing the data (e.g. `plot(dist ~ speed, cars)`)

2.2 Exercise 2: Stochastic gradient descent / incremental gradient descent

2.2.1 What is the difference between stochastic gradient descent and gradient descent?

2.2.2 How do you think the optimisation path (the path $(k, f(x_k))$) looks like for stochastic gradient descent compared to that of the gradient descent?

2.2.3 Optional: Implement stochastic gradient descent.

2.2.4 Optional: Illustrate the behaviour of the stochastic gradient descent, including:

2.2.4.1 Different ways to choose step sizes.

2.2.4.2 The total objective function with a discussion of how it differs from a similar plot from the gradient descent method.

2.2.4.3 Some iterates in a plot showing the data (e.g. `plot(dist ~ speed, cars)`).

3 Calculating derivatives

3.1 Exercise 1

3.2 Exercise 2

3.3 Exercise 3

4 Quasi Newton

5 Least Squares

6 Constrained Optimization