Universität Konstanz
Fachbereich Mathematik und Statistik
Prof. Dr. Stefan Volkwein
Sabrina Rogg

Sommersemester 2015

Optimierung

http://www.math.uni-konstanz.de/numerik/personen/rogg/de/teaching/

Sheet 6

Tutorial: 7th July

Exercise 17 (Scaled gradient method)
Consider the quadratic function $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x) = \frac{1}{2} x^\top \begin{pmatrix} 100 & -1 \\ -1 & 2 \end{pmatrix} x + \begin{pmatrix} 1 & 1 \end{pmatrix} x + 3.$$

Implement a modified version of the Gradient Method where the update is

$$x^{k+1} = x^k + t_k d^k \quad \text{with} \quad d^k = M^{-1}(-\nabla f(x^k))$$

and with the exact stepsize t^k (Exercise 5). M is one of the following matrices:

$$M = \mathrm{Id} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M = \nabla^2 f = \begin{pmatrix} 100 & -1 \\ -1 & 2 \end{pmatrix}, \quad M = \begin{pmatrix} f_{xx} & 0 \\ 0 & f_{yy} \end{pmatrix} = \begin{pmatrix} 100 & 0 \\ 0 & 2 \end{pmatrix}.$$

As basis you can use the Gradient Method you implemented for the first program sheet. Determine the number of gradient steps required for finding the minimum of f with the different matrices M and initial value `x0 = [1.5;0.6]` (use $\epsilon = 10^{-9}$). How close is the computed point to the exact analytical minimum? Explain your observations.

Exercise 18 (Cauchy-step property)
The Cauchy step is defined as $s_a^{CP} = -t_a \nabla f(x_a)$, where t_a is given by (see the lecture notes)

$$t_a = \begin{cases} \frac{\Delta_a}{\|\nabla f(x_a)\|} & \text{if } \nabla f(x_a)^\top H_a \nabla f(x_a) \leq 0, \\ \min\left(\frac{\Delta_a}{\|\nabla f(x_a)\|}, \frac{\|\nabla f(x_a)\|^2}{\nabla f(x_a)^\top H_a \nabla f(x_a)} \right) & \text{if } \nabla f(x_a)^\top H_a \nabla f(x_a) > 0. \end{cases}$$

Once the Cauchy point $x_a^{CP} = x_a + s_a^{CP}$ is computed, show that there is a sufficient decreasing in the quadratic model, i.e, the Cauchy point satisfies

$$m_a(x_a) - m_a(x_a^{CP}) \geq \frac{1}{2} \|\nabla f(x_a)\| \min\left(\Delta_a, \frac{\|\nabla f(x_a)\|}{1 + \|H_a\|} \right).$$

Exercise 19 (Dogleg strategy)
Let us consider the quadratic model of the function f in x_a

$$m_a(x) = f(x_a) + \nabla f(x_a)^\top (x - x_a) + \frac{1}{2}(x - x_a)^\top B_a (x - x_a),$$

with B_a positive definite (Hessian matrix in x_a or an approximation of it). In the trust-region subproblem we approximately solve

$$\min_{\|x-x_a\|<\Delta_a} m_a(x). \tag{1}$$

If the trust region is big enough, i.e as if there is no constraint $\|x - x_a\| < \Delta_a$, the exact (global) solution to (1) is the (quasi-)Newton point

$$x_a^{QN} = x_a - B_a^{-1}\nabla f(x_a).$$

The idea of the dogleg method is as follows: Minimize m_a along a path consisting of two straight lines: one from the current point x_a to the Cauchy point x_a^{CP} and the other one from x_a^{CP} to the (quasi-)Newton point x_a^{QN}. This path is described as

$$x(\tau) = \begin{cases} x_a + \tau(x_a^{CP} - x_a) & \tau \in [0, 1], \\ x_a^{CP} + (\tau - 1)(x_a^{QN} - x_a^{CP}) & \tau \in (1, 2]. \end{cases}$$

Show that $h(\tau) := m_a(x(\tau))$ is a decreasing function of τ.
Hint: consider m_a on the two straight lines separately.