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

Sommersemester 2016

Optimierung

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

Sheet 6

Deadline for hand-in: 27.06.2016 at lecture

Exercise 16 (Scaled gradient method) (4 Points)
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 the gradient-like 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 4). 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 use the gradient method you implemented for the first program sheet. Determine the number of 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 are the computed points to the exact analytical minimum? For $M = \nabla^2 f$ compute t_k exactly. What can you conclude? Explain your results.

Trust Region Method

Exercise 17 (5 Points)
We consider the function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^3 + 2x^2 - 11x - 12. \tag{1}$$

We take a look at one step in the trust region algorithm for different points $x_a \in \mathbb{R}$. For every x_a below build the quadratic model m_a of f around x_a with $H_a = \nabla^2 f(x_a)$ and plot the function f together with the quadratic model m_a (you can use Matlab).

1. $x_a = 3$. Determine the global minimum $\tilde{x}_V$ of m_a. Compute $\rho = \mathrm{ared}_a/\mathrm{pred}_a$ for $\tilde{x}_V$. Interprete the value of ρ.

2. $x_a = -0.5$. Determine the global minimum $\tilde{x}_V$ of m_a. Then, start with $\Delta = \tilde{x}_V - x_a$ and divide Δ by two until the trial point x_V,

$$x_V = \arg\min m_a(x) \quad \text{subject to} \quad |x - x_a| \leq \Delta, \tag{2}$$

satisfies $\rho > 0.2$ (you can use Matlab).

3. $x_a = -1$. What is different here compared to point 1. and 2.? Which point solves always the trust region subproblem (2)?