Exercises for Lecture Course on Numerical Optimization (NUMOPT)
Albert-Ludwigs-Universität Freiburg – Winter Term 2015-2016

Exercise 5: Unconstrained Newton-type Optimization

(to be completed during exercise session on Nov 25, 2015 or sent by email to dimitris.kouzoupis@imtek.uni-freiburg.de before Nov 27, 2015)

Prof. Dr. Moritz Diehl, Dimitris Kouzoupis and Andrea Zanelli

Aim of this exercise is to become familiar with different Newton-type methods and learn their characteristics in practice.

Exercise Tasks

1. **Regularization:** Prove that a regularized Newton-type step $x_{k+1} = x_k - (B_k + \alpha I)^{-1}\nabla f(x_k)$ with B_k a Hessian approximation, α a positive scalar and I the identity matrix of suitable dimensions, converges to a small gradient step $x_{k+1} = x_k - \frac{1}{\alpha}\nabla f(x_k)$ as $\alpha \to \infty$.

   (2 points)

2. **Unconstrained minimization:** In this task we will implement different Newton-type methods that minimize the nonlinear function

$$f(x,y) = \frac{1}{2}(x-1)^2 + \frac{1}{2}(10(y-x^2))^2 + \frac{1}{2}y^2. \tag{1}$$

   (a) Derive, first on paper, the gradient and Hessian matrix of the function in (1). Then, re-write it in the form $f(x,y) = \frac{1}{2}||R(x,y)||_2^2$ where $R : \mathbb{R}^2 \to \mathbb{R}^3$ is the residual function. Derive the Gauss-Newton Hessian approximation and compare it with the exact one. When do the two matrices coincide?

   (2 points)

   (b) Implement your own Newton method with exact Hessian information and full steps. Start from the initial point $(x_0, y_0) = (-1, -1)$ and use as termination condition $||\nabla f(x_k, y_k)||_2 \le 10^{-3}$. Keep track of the iterates (x_k, y_k) and use the provided function to plot the results.

   (2 points)

   (c) Update your code to use the Gauss-Newton Hessian approximation instead. Compare the performance of the two algorithms and plot the difference between exact and approximate Hessian as a function of the iterations (use the MATLAB function `norm` to measure this difference).

   (2 points)

   (d) Check how the steepest descent method performs on this example. Your Hessian now becomes simply αI where α is a positive scalar and I the identity matrix. Try $\alpha = 100, 200$ and 500. For which values does your algorithm converge? How does its performance compare with the previous methods?

   (1 point)

   (e) Imagine you remove the term $\frac{1}{2}y^2$ from $f(x,y)$ and compare the exact Newton's method with the Gauss-Newton. What do you expect?

   (1 point)

3. **Lifted Newton method:** Consider the scalar nonlinear function $F(w) = w^{16} - 2$.

   (a) Implement in MATLAB the Newton method in order to numerically find a root of $F(w)$. Plot how the residuals evolve. Test the algorithm for different initial guesses and analyze the behaviour of the algorithm.

   (1 point)

   (b) Implement now a Newton-type algorithm that exploits a fixed approximation of the gradient

   $$w^{k+1} = w^k - M^{-1}F(w^k),$$

   where $M = \nabla F(w_0)$ is the gradient of F at the initial guess w_0. For which range of values $a \leq w_0 \leq b$ does the algorithm converge?

   (1 point)

   (c) An equivalent problem to (a) can be obtained by *lifting* the original one to a higher dimensional space

   $$\tilde{F}(w) = \begin{bmatrix} w_2 & - & w_1^2 \\ w_3 & - & w_2^2 \\ w_4 & - & w_3^2 \\ -2 & - & w_4^2 \end{bmatrix}.$$

   Implement the Newton method for this lifted problem and compare the convergence of the two algorithms.

   (1 point)

   (d) Show that the Newton method is guaranteed to converge to a root of any monotonocally increasing convex differentiable function $F : \mathbb{R} \to \mathbb{R}$.

   (1 point)

This sheet gives in total 14 points.