Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimisation I Spring 2017

Exercise set 5

1 One possibility for lowering the memory requirements of the BFGS-method is to reset the matrix B_k (or its inverse H_k) to the identity matrix after each jth step for some fixed number j.[1] For $j = 1$ this leads—with the notation of the lecture and Nocedal & Wright, Chapter 6—to the update

$$H_{k+1} = \Big(\mathrm{Id} - \frac{s_k y_k^\top}{y_k^\top s_k}\Big)\Big(\mathrm{Id} - \frac{y_k s_k^\top}{y_k^\top s_k}\Big) + \frac{s_k s_k^\top}{y_k^\top s_k}.$$

Assume now that this method is implemented with an exact line search. Show that this yields a non-linear CG-method, where the search directions are defined by

$$p_{k+1} = -\nabla f_{k+1} + \beta_{k+1} p_k$$

with

$$\beta_{k+1} = \frac{\nabla f_{k+1}^\top (\nabla f_{k+1} - \nabla f_k)}{(\nabla f_{k+1} - \nabla f_k)^\top p_k}$$

(this is the *Hestenes–Stiefel method*, cf. Nocedal & Wright, p. 123).

(Hint: You may need to show in a first step that an exact line search implies that $\nabla f_{k+1}^\top p_k = 0 = \nabla f_{k+1}^\top s_k$*.)*

2 Let

$$f(x) = x_1^4 + 2x_2^4 + x_1 x_2 + x_1 - x_2 + 2.$$

Starting at the point $x_0 = (0, 0)$ compute explicitly one step for the trust region method with the model function $m(p) = f(x_0) + g^\top p + \frac{1}{2} p^\top B p$, where $g = \nabla f(x_0)$, $B = \nabla^2 f(x_0)$, and the trust region radius $\Delta = 1$.

3 Let

$$f(x) = \tfrac{1}{2} x_1^2 + x_2^2,$$

put $x_0 = (1, 1)$, and define the model function $m(p) = f(x_0) + g^\top p + \frac{1}{2} p^\top B p$ with $g = \nabla f(x_0)$ and $B = \nabla^2 f(x_0)$.

a) Compute explicitly the next step p in the trust region method using values of $\Delta = 2$ and $\Delta = 5/6$.

b) Compute for all $\Delta > 0$ the next step in the dogleg method.

[1] More sophisticated methods are described in Nocedal & Wright, Chapter 7.2.

4 In this exercise, we study the Gauß–Newton method for solving the least-squares problem corresponding to the (overdetermined and inconsistent) system of equations

$$\begin{aligned} x + y &= 1, \\ x - y &= 0, \\ xy &= 2. \end{aligned}$$

To that end, we define

$$\begin{aligned} r_1(x, y) &:= x + y - 1, \\ r_2(x, y) &:= x - y, \\ r_3(x, y) &:= xy - 2, \end{aligned}$$

and

$$f(x, y) := \frac{1}{2} \sum_{j=1}^{3} r_j(x, y)^2.$$

We denote moreover by $J = J(x, y)$ the Jacobian of $r = (r_1, r_2, r_3) \colon \mathbb{R}^2 \to \mathbb{R}^3$.

a) Show that the function f is non-convex, but that it has a unique minimiser (x^*, y^*).

b) Show that the matrix $J^\top J$ required in the Gauß–Newton method is positive definite for all x, y.

c) Show that the Gauß–Newton method with Wolfe line search for the minimisation of f converges for all initial values (x_0, y_0) to the unique solution of the non-linear least squares problems.

d) Perform one step of the Gauß–Newton method (without line search) for the solution of this least-squares problem. Use the initial value $(x_0, y_0) = (0, 0)$.