Exercise 8

March 05, 2024

Problem 1.

The secant method for the solution of one-dimensional optimisation problems is given by the iteration

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f'(x_k) - f'(x_{k-1})} f'(x_k).$$

Show that this method coincides with both the BFGS and the DFP Quasi-Newton methods without line search.

Problem 2. (See N&W, Problem 6.1.a)

Consider a line search method $x_{k+1} = x_k + \alpha_k p_k$ with search direction

$$p_k = -B_k^{-1} \nabla f(x_k),$$

where $B_k \in \mathbb{R}^{d \times d}$ is a positive definite, symmetric matrix. Denote $s_k = x_{k+1} - x_k$ and $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$.

Show that the curvature condition (see N&W, eq. (6.7))

$$\langle s_k, y_k \rangle > 0$$

holds for all step lengths $\alpha_k > 0$, if f is a strictly convex function and $\nabla f(x_k) \neq 0$.

Problem 3.

Implement the BFGS method for the minimisation of the Rosenbrock function.

Note that you will require a Wolfe line search in order to ensure that the matrices stay positive definite and the search directions are actually descent directions.

Problem 4.

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 j-th step for some fixed number j.[1] For $j = 1$ this leads (with the notation of the lecture) to the update

$$H_{k+1} = \left(\mathrm{Id} - \frac{s_k \otimes y_k}{\langle y_k, s_k\rangle}\right)\left(\mathrm{Id} - \frac{y_k \otimes s_k}{\langle y_k, s_k\rangle}\right) + \frac{s_k \otimes s_k}{\langle y_k, s_k\rangle}.$$

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

$$\beta_{k+1} = \frac{\langle \nabla f_{k+1}, \nabla f_{k+1} - \nabla f_k\rangle}{\langle \nabla f_{k+1} - \nabla f_k, p_k\rangle}$$

(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 $\langle \nabla f_{k+1}, p_k\rangle = 0 = \langle \nabla f_{k+1}, s_k\rangle$*.)*

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