Exercise #8

March 07, 2023

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.

Solution.

Have already discussed in the class!

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 $\nabla f(x_k) \neq 0$.

Solution.

The strict convexity of the function f implies

$$f(x_{k+1}) - f(x_k) > \langle \nabla f(x_k), x_{k+1} - x_k \rangle. \tag{1}$$

By interchanging x_{k+1} and x_k in (1), we get

$$f(x_k) - f(x_{k+1}) > \langle \nabla f(x_{k+1}), x_k - x_{k+1} \rangle. \tag{2}$$

By adding equations 1 and 2, we obtain

$$\langle x_{k+1} - x_k, \nabla f(x_{k+1}) - \nabla f(x_k) \rangle > 0,$$

which is

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

Problem 3.

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 and Nocedal & Wright, Chapter 6) to the update

$$H_{k+1} = \left(\mathrm{Id} - \frac{s_k y_k^T}{y_k^T s_k}\right)\left(\mathrm{Id} - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{s_k s_k^T}{y_k^T s_k}.$$

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

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

(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 $\nabla f_{k+1}^T p_k = 0 = \nabla f_{k+1}^T s_k$.)

Solution.

Have already discussed in the class!

Problem 4.

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.

Solution.

See the implementation on the wiki page.

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