Exercise 4

February 04, 2025

Problem 1.

Let

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 2 \end{pmatrix} \qquad \text{and} \qquad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Use the CG-method with initialization $x_0 = 0$ for solving the linear system $Ax = b$.

Problem 2.

Implement the linear CG method (N&W, p. 112, Algorithm 5.2) and use it to solve linear systems in which A is the Hilbert matrix, whose elements are $A_{i,j} = \frac{1}{i+j-1}$. Set the right-hand-side to $b = (1, 1, \ldots, 1)^T$ and the initial point to $x_0 = 0$. Try dimensions $n = 5, 8, 12, 20$, and report the number of iterations required to reduce the residual below 10^{-6}.

Problem 3.

Implement a non-linear CG-method (see N&W, Algorithm 5.4). Test it on the *Rosenbrock function* $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$. Use the strong Wolfe conditions with $0 < c_1 < c_2 < 1/2$ for step length selection.

Hint: You can use a bracketing method as discussed in the lecture for finding a suitable step length. In this case, you have to recall, though, that the two conditions

$$\begin{aligned} f(x_k + \alpha p_k) &\le f(x_k) + c_1 \alpha \langle \nabla f(x_k), p_k \rangle, \\ \langle \nabla f(x_k + \alpha p_k), p_k \rangle &\le -c_2 \langle \nabla f(x_k), p_k \rangle, \end{aligned}$$

indicate that a step length α is not too large (that is, if one of those is violated, then the current step length *is* too large), while only the weak curvature condition

$$\langle \nabla f(x_k + \alpha p_k), p_k \rangle \ge c_2 \langle \nabla f(x_k), p_k \rangle$$

prevents too small step lengths α.

Alternatively, you can use the algorithm that is described in N&W, Section 3.5.

Problem 4.

Let $Q \in \mathbb{R}^{d\times d}$ by symmetric and positive definite and let $b \in \mathbb{R}^d$. Define $\Phi\colon \mathbb{R}^d \to \mathbb{R}$,

$$\Phi(x) = \frac{1}{2}\langle x, Qx\rangle - \langle x, b\rangle.$$

Consider the minimisation of Φ using a non-linear CG method *with exact line search*, where the search directions are chosen either with the Polak–Ribière formula (N&W, eq. (5.44)), the Hestenes–Stiefel formula (N&W, eq. (5.46)), or the Fletcher–Reeves formula (N&W, eq. (5.41a)). Show that all of these methods reduce to the CG-method for linear systems.

Hint: Use induction to show that $\langle r_{k+1}, p_k\rangle = 0$, $\langle r_k, p_k\rangle = -\|r_k\|^2$ and $\langle r_{k+1}, r_k\rangle = 0$ for all k.

Problem 5.

Consider a non-linear CG-method defined by

$$p_k = -\nabla f(x_k) + \beta_{k-1} p_{k-1},$$

where $\beta_{k-1} \in \mathbb{R}$ satisfies the estimate

$$|\beta_{k-1}| \le \beta_{k-1}^{\mathrm{FR}} := \frac{\|\nabla f(x_k)\|^2}{\|\nabla f(x_{k-1})\|^2}$$

and the step length $\alpha_k > 0$ satisfies the strong Wolfe conditions with $0 < c_1 < c_2 < 1/2$.

Show that

$$-\frac{1}{1-c_2} \le \frac{\langle \nabla f(x_k), p_k\rangle}{\|\nabla f(x_k)\|^2} \le \frac{2c_2 - 1}{1-c_2}$$

for all k.

Hint: Cf. the proof of Lemma 5.6.