Norwegian University of Science and Technology
Department of Mathematical Sciences

Exercise set 4

1 Implement both the gradient descent method and Newton's method with a line search satisfying the Wolfe conditions (you may want to use a bisection algorithm for the implementation of these conditions)

Apply your method to the minimisation of the Rosenbrock function

$$f(x,y) := 100(y-x^2)^2 + (1-x)^2.$$

The Newton direction is not necessarily a descent direction for this function, as f is not convex, and thus it might be necessary to modify the search directions in the Newton method. Do this by switching to the negative gradient direction, whenever the inequality

$$-\nabla f(x_k)^T p_k^{\text{Newton}} \le \varepsilon \|\nabla f(x_k)\| \|p_k^{\text{Newton}}\|$$

holds (here, $\varepsilon > 0$ is some fixed, small parameter).

2 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 initialisation $x_0 = 0$ for solving the linear system $Ax = b$.

3 Assume that $A \in \mathbb{R}^{m\times n}$ is a matrix and that $b \in \mathbb{R}^m$.

a) Show that $x^* \in \mathbb{R}^n$ solves the *least squares problem*

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|^2, \tag{1}$$

if and only if x^* satisfies the *normal equations*

$$A^T A x^* = A^T b.$$

b) Show that the optimization problem (1) admits a solution $x^* \in \mathbb{R}^n$.

c) Show that the solution x^* of (1) is unique, if the rank of A equals n.

d) Show that, regardless of the rank of A, the optimization problem

$$\min_{x\in\mathbb{R}^n} \|x\|^2 \qquad \text{s.t. } x \text{ solves (1)} \tag{2}$$

admits a unique solution $x^\dagger \in \mathbb{R}^n$.

$\boxed{4}$ Assume that $A \in \mathbb{R}^{n \times n}$ is symmetric and positive *semi*-definite and $b \in \operatorname{Range} A$. Show that (in exact arithmetics) the CG algorithm converges for every starting point $x_0 \in \mathbb{R}^n$ in at most $m = \dim(\operatorname{Range} A)$ iterations to a solution of $Ax = b$.

(This shows that at least theoretically the assumption of positive definiteness can be slightly relaxed.)