Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization I Spring 2020

Exercise set 4

The first and third exercise below are concerned with the usage of the CG-method for the solution of linear least squares problems (which is one approach to the solution of overdetermined linear systems).

1 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*

$$A^TAx^* = A^Tb.$$

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 $x^\dagger \in \mathbb{R}^n$.

2 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.)

3 Assume that $m > n$, that $A \in \mathbb{R}^{m\times n}$, and that $b \in \mathbb{R}^m$. Consider the following algorithm:

- Choose $x_0 \in \mathbb{R}^n$ arbitrary, set $r_0 \leftarrow Ax_0 - b$, $s_0 \leftarrow A^T r_0$, $p_0 \leftarrow -s_0$, and $k \leftarrow 0$.
- While $s_k \neq 0$:

$$\begin{aligned}
\alpha_k &\leftarrow \frac{\|s_k\|^2}{\|Ap_k\|^2},\\
x_{k+1} &\leftarrow x_k + \alpha_k p_k,\\
r_{k+1} &\leftarrow r_k + \alpha_k A p_k,\\
s_{k+1} &\leftarrow A^T r_{k+1},\\
\beta_{k+1} &\leftarrow \frac{\|s_{k+1}\|^2}{\|s_k\|^2},\\
p_{k+1} &\leftarrow -s_{k+1} + \beta_{k+1} p_k,\\
k &\leftarrow k+1.
\end{aligned}$$

Assume that the matrix A has full rank. Show that the algorithm above is actually identical with the CG-algorithm for the solution of $A^T Ax = A^T b$ (in the sense that the iterates x_k of both methods coincide).

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.

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

b) Show that the matrix $J^T J$ 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 $(x_0, y_0) = (0, 0)$.