# Exercise Sheet 9.

Due: **Monday, 24.11.2025, 12:00.**

**Exercise 1** (Gauss-Newton algorithm). Let $F: \mathbb{R} \to \mathbb{R}^2$ be the function given by

$$F(x) = \begin{bmatrix} x+1 \\ \lambda x^2 + x - 1 \end{bmatrix}, \quad \lambda \in \mathbb{R},$$

and define the non-linear least-squares problem

$$\phi(x) = \|F(x)\|_2^2 \to \min. \tag{1}$$

(a) Show that the function $\phi(x)$ for $\lambda < 1$ has a local minimum at $x^\star = 0$. Show also that when $\lambda < 7/16$ then $x^\star$ even is the only local minimum.

(b) Formulate the Gauss-Newton method for the problem (1). Prove that when $\lambda < -1$ then $x^\star = 0$ is a repulsive fixed point of the Gauss-Newton method, i.e. there exists a $\delta > 0$, such that $|x_{k+1} - 0| > |x_k - 0|$ holds for all $x_k$ with $0 < |x_k - 0| < \delta$.

(c) For $|\lambda| < 1$, the Gauss-Newton method is convergent. What is the asymptotic convergence in this case?

**Exercise 2** (Levenberg-Marquardt algorithm). Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ with $m \geq n$, $\mathbf{b} \in \mathbb{R}^m$ and $\lambda > 0$. Furthermore, let $\mathbf{d}$ be the solution of

$$\left(\mathbf{A}^\intercal \mathbf{A} + \lambda \mathbf{I}\right) \mathbf{d} = \mathbf{A}^\intercal \mathbf{b}$$

and $\mathbf{g}$ be the solution of

$$\left(\mathbf{A}^\intercal \mathbf{A} + \lambda \mathbf{I}\right) \mathbf{g} = \mathbf{d}.$$

(a) Show that $\mathbf{d}$ and $\mathbf{g}$, respectively, are the solutions of the linear least-squares problems

$$\min_{\mathbf{v}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{I} \end{bmatrix} \mathbf{v} - \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} \right\|_2 \qquad \text{and} \qquad \min_{\mathbf{w}\in\mathbb{R}^n} \left\| \begin{bmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{I} \end{bmatrix} \mathbf{w} - \begin{bmatrix} \mathbf{0} \\ \mathbf{d}/\sqrt{\lambda} \end{bmatrix} \right\|_2. \tag{2}$$

(b) Let the QR-decomposition of $\mathbf{A}$ be defined as $\mathbf{A} = \mathbf{QR}$. Describe a method that solves the linear least squares problems (2) using $n(n+1)/2$ Givens rotations.

Hint. *Use the fact that a linear least squares problem can be solved using the QR decomposition.*

**Exercise 3** (Hebden algorithm).

(a) Let $I \subset \mathbb{R}$ be an interval and $f: I \to \mathbb{R}$ be a differentiable, strictly monotonically increasing, and concave function with a unique zero $x^\star \in I$. Show that the Newton method for finding the zeros converges monotonically to $x^\star$ for all initial values $x_0 \in I$ with $x_0 \leq x^\star$.

Hint. *Use the following property on differentiable concave functions:*

$$f(y) \leq f(x) + f'(x)[y-x] \qquad \textit{for all } x, y \in I.$$

(b) Consider the equation

$$r(x) := \sum_{i=1}^{n} \frac{z_i^2}{(d_i + x)^2} = \rho \tag{3}$$

with $z_i$ and $d_i$ positive for all $i = 1, \dots, n$ and $\rho > 0$. Furthermore, assume that $d_i > d_{i+1}$ for all $i = 1, \dots, n-1$ and $r(0) > \rho$. To solve the equation (3) using the Hebden method, the Newton method for finding the roots is applied to the transformed equation.

$$g(x) = \frac{1}{\sqrt{r(x)}} - \frac{1}{\sqrt{\rho}} = 0.$$

Show that the Hebden method converges for the initial value $x_0 = 0$.

Hint. *Study the function* $g(x)$ *to use part (a).*

**Exercise 4** (Trust-Region algorithm). Let $F(\mathbf{x}) := \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}$ with $\mathbf{b} \in \mathbb{R}^n$ and a symmetric positive definite matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. For $F$ and the trust region radius $\Delta > 0$, the trust region problem is:

$$\min \{F(\mathbf{x}) \ : \ \|\mathbf{x}\|_2 \le \Delta\}.$$

Let $\mathbf{x}^\star$ be the corresponding optimal solution. Consider the following conditions:

(i) $\lambda^\star \ge 0$, $\|\mathbf{x}^\star\|_2 \le \Delta$, $\lambda^\star(\Delta - \|\mathbf{x}^\star\|_2) = 0$,

(ii) $(\mathbf{A} + \lambda^\star \mathbf{I})\mathbf{x}^\star = -\mathbf{b}$,

(iii) $\mathbf{A} + \lambda^\star \mathbf{I}$ is positive semi-definite.

Show that:

(a) If $\|\mathbf{x}^\star\|_2 < \Delta$ holds, then $\mathbf{x}^\star$ and $\lambda^\star = 0$ satisfy the conditions (i)–(iii).

(b) Let $\|\mathbf{x}^\star\|_2 = \Delta$ and $\mathbf{v} \in \mathbb{R}^n$ with $\|\mathbf{v} + \mathbf{x}^\star\|_2 \le \Delta$. Then $(\mathbf{b} + \mathbf{A}\mathbf{x}^\star)^\top \mathbf{v} \ge 0$ holds.