problem sheet 10

discussion: week of Monday, 13.12.2021

10.1. The Frobenius norm of a matrix $\mathbf{A}$ is given by $\|\mathbf{A}\|_F^2 = \sum_{i,j} |\mathbf{A}_{ij}|^2$.

a) Show for an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ and an arbitrary matrix $\mathbf{X} \in \mathbb{R}^{n\times k}$ that $\|\mathbf{Q}\mathbf{X}\|_F^2 = \|\mathbf{X}\|_F^2$.

b) Let the SVD of $\mathbf{A} \in \mathbb{R}^{m\times n}$ be $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$, where the diagonal entries of Σ are the singular values $\sigma_1, \dots, \sigma_{\min\{m,n\}}$. Show: $\|\mathbf{A}\|_F^2 = \sum_i \sigma_i^2$.

10.2. Program Newton's method in 1D. To that end, realize a `Matlab`/`python` function `newton(x, f, df)` that realizes one step of the method. f and df are *function handles* for the function f and its derivative f'. Plot (use `semilogy`) the error versus the number of Newton steps for the following 3 functions:

$$f_1(x) = x^2, \qquad f_2(x) = e^x - 2, \qquad f_3(x) = |x|^{3/2}.$$

Use $x_0 = 0.5$ as the starting value. What do you observe? Which assumptions that underlie the proof of quadratric convergence are not satisfied? Consider Newton's method for

$$f_4(x) = \frac{1}{x} - 1$$

and initial value $x_0 = 2.1$. What do you observe? *Remark:* This reduces the problem of a division to one of multiplication only.

10.3. (**to be uploaded on TUWEL**) Consider the nonlinear system of equations $\mathbf{f}(\mathbf{x}) = 0$ given by

$$\begin{aligned} 3x_1 - \cos(x_2x_3) - 3/2 &= 0 \\ 4x_1^2 - 625x_2^2 + 2x_3 - 1 &= 0 \\ 20x_3 + e^{-x_1x_2} + 9 &= 0 \end{aligned}$$

Compute the derivative $\mathbf{f}'(\mathbf{x})$ and formulate Newton's method. Program Newton's method in `matlab`/`python`. The program should additionally estimate the error (e.g., in the $\|\cdot\|_2$-norm) and also return it. Use the initial vector $(1, 1, 1)^\top$. Plot (using `semilogy`) the estimated error versus the iteration number.

10.4. Consider the system of equations

$$\mathbf{A}x = \mathbf{b} + \varepsilon\mathbf{f}(\mathbf{x}), \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$

with

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \qquad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} (x_1 - x_2)^2 \\ 0 \end{pmatrix}, \qquad \varepsilon = 0.01.$$

a) Formulate the Newton method and write a program to compute the iterates $\mathbf{x}_n$, $n = 1, 2, \dots,$. Initial value: $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$.

b) Consider the following method with initial value $\mathbf{x}_0$ given by $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$: For $n = 0, 1, \dots$ one determines $\mathbf{x}_{n+1} \in \mathbb{R}^2$ such that $\mathbf{A}\mathbf{x}_{n+1} = \mathbf{b} + \varepsilon\mathbf{f}(\mathbf{x}_n)$. Write a program to compute the iterates. Taking the last value of the Newton method as the "exact solution" you can compute the errors. Plot the error for both methods in `loglog`-plot (error versus iteration index n). Can you relate this "simple" method to Newton's method for small ε?

10.5. Show that $\|\mathbf{x}_{n+1} - \mathbf{x}_n\|$ of Problem 10.4 is a good estimate for the error $\|\mathbf{x}_* - \mathbf{x}_n\|$ for the method of Problem 10.4b. To see this, you may use the following facts and hints:

- One step of Newton's method with starting value $\mathbf{x}_n$ has the form $\mathbf{x}_{n+1}^{Newton} = \mathbf{x}_n - (\mathbf{F}'(\mathbf{x}_n))^{-1}\mathbf{F}(\mathbf{x}_n)$ with $\mathbf{F}(x) = \mathbf{A}\mathbf{x} - \mathbf{b} - \varepsilon\mathbf{f}(\mathbf{x})$ and $\mathbf{F}'(\mathbf{x}) = \mathbf{A} - \varepsilon\mathbf{f}'(\mathbf{x})$.

- one has $\|\mathbf{F}(\mathbf{x}_n)\| \leq C\|\mathbf{e}_n\|$ for the error $\mathbf{e}_n = \mathbf{x}_\star - \mathbf{x}_n$.
- Geometric series yield, for small ε and matrices $\mathbf{B}$, $\mathbf{C}$ with $\mathbf{B}$ invertible: $(\mathbf{B} + \varepsilon\mathbf{C})^{-1} = \mathbf{B}^{-1} + O(\varepsilon)$ (here, $O(\varepsilon)$ is a matrix whose entries are of size $O(\varepsilon)$; the exact matrix entries depend in a complicated way on $\mathbf{B}$ and $\mathbf{C}$, but are not so important for our purposes)
- Try to estimate $\|\mathbf{x}_{n+1}^{Newton} - \mathbf{x}_{n+1}\| \leq O(\varepsilon)\|\mathbf{e}_n\|$
- recall that $\|\mathbf{x}_{n+1}^{Newton} - \mathbf{x}_n\|$ is an excellent estimator for $\|\mathbf{e}_n\|$