Univ.-Prof. Dr. Matthias Ehrhardt
Dipl.-Math. Alexander Schranner
WiSe 2009/10

Bergische Universität
Wuppertal
Angewandte Mathematik
Numerische Analysis

Introduction to Numerical Methods for Computer Simulation

Exercises No. 1

E 1: *Floating Point Arithmetic*

Let $B = 2$, $t = 4$, and M, E be integers.

$$\begin{aligned} \mathbb{G} &:= \{g = M \cdot B^E : M = 0 \text{ or } B^{t-1} \leq |M| < B^t\} \\ &\quad \text{(t-digit floating point numbers with basis } B) \\ \mathbb{M} &:= \{g = M \cdot B^E \in \mathbb{G} : -6 \leq E \leq -2\} \end{aligned}$$

a) Determine the largest and the smallest positive number in $\mathbb{M}$.

b) Sketch $\mathbb{M}$ on the real line.

c) Determine the maximal and the minimal relative and absolute distance of two successive positive numbers in $\mathbb{M}$.

d) Try to express the numbers 1.625, 3.7, 0.02 and 4.2 as machine numbers in $\mathbb{M}$.

E 2: *Backward Analysis of Scalar Product*

We consider the scalar product $\langle x, y\rangle$ for vectors $x = (x_1, \dots, x_n)^\top$, $y = (y_1, \dots, y_n)^\top \in \mathbb{R}^n$. It is calculated with the recursive formula

$$\langle x, y\rangle := x_n y_n + \langle x^{n-1}, y^{n-1}\rangle, \qquad (\star)$$

where $x^{n-1} := (x_1, \dots, x_{n-1})^\top$ and $y^{n-1} := (y_1, \dots, y_{n-1})^\top$.

The floating point realisation of the scalar product following $(\star)$ calculates a solution $\langle x, y\rangle_{fl} \in \mathbb{G}$.

Apply backward analysis for $\langle x, y\rangle_{fl}$ to show the following proposition:

There exists $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n) \in \mathbb{R}^n$, such that

$$\langle x, y\rangle_{fl} = \langle \hat{x}, y\rangle$$

with (using a linearisation)

$$|x_i - \hat{x}_i| \dot{\leq} n\,\varepsilon_0\,|x_i|, \qquad i = 1,\ldots,n.$$

Is the scalar product stable in terms of backward analysis?

E 3: *Condition numbers*

The eigenvalues of a symmetric matrix

$$A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

are given by the two zeros of the characteristic polynomial $P(\lambda) = \lambda^2 - (a+b)\lambda + ab - c^2$, i.e.

$$\lambda_{1/2} := \frac{a+b}{2} \pm \sqrt{\frac{(a+b)^2}{4} + c^2 - ab}\,.$$

It holds $\lambda_1 = \lambda_2$ only if $a = b$ and $c = 0$. Thus we assume $a \neq b$ or $c \neq 0$. In the following, we observe just λ_1 (λ_2 behaves the same).

a) Show that the problem $\lambda_1 = \varphi(a, b, c)$ is well-conditioned with respect to perturbations in a, b and c.

b) The computation of the zero λ_1 is done in two steps now:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \xrightarrow{\tau} \begin{pmatrix} p \\ q \end{pmatrix} \xrightarrow{\sigma} \lambda_1$$

$$\text{with} \quad \tau(a,b,c) := \begin{pmatrix} -\frac{1}{2}(a+b) \\ ab - c^2 \end{pmatrix} \quad \text{and} \quad \sigma(p,q) := -p + \sqrt{p^2 - q}.$$

Compute the absolute condition numbers for each step.
Is the algorithm $\sigma \circ \tau$ for computing λ_1 stable?

E 4: *Roundoff Errors and Cancellation*

Given is the problem

$$p^2 - 2q^2 \quad \text{with} \quad p = 665857 \quad \text{and} \quad q = 470832\,.$$

In exact arithmetic, we have

$$p^2 = 443365544449, \quad q^2 = 221682772224, \quad 2q^2 = 443365544448$$

and therefore $p^2 - 2q^2 = 1$.
Now, use your pocket calculator. Compute the results using

i) t= 8 ii) t=11 iii) t=12

significant digits. Compare these results with the exact one.
How can you compute the exact solution using 8 digits?