Otto-von-Guericke-University Magdeburg
Max Planck Institute for Dynamics of Complex Technical Systems
Computational Methods in Systems and Control Theory

Dr. Jens Saak, Dipl.-Math. Martin Köhler
Website: http://www.mpi-magdeburg.mpg.de/2802004/sc1/

Winterterm 2014/2015

Scientific Computing 1
Solution to the tutorial part of Homework 3
11/21/2014

Solution

Exercise 1:

Determine the absolute and the relative error of 0.5403023059 and π in

a.) $\mathbb{M}(10,3,-2,2)$

b.) $\mathbb{M}(2,3,-2,3)$

c.) $\mathbb{M}(2,5,-2,2)$

Solution:

a) Für 0.5403023059:
$\gamma(x) = 0.540$ in $\mathbb{M}(10,3,-2,2)$
Darstellung der Zahl im Binärsystem: $(0.5403023059)_{10} = (0.100010100\ldots)_2$
Somit gilt:

$$\gamma(x) = \begin{cases} 0.100 & \text{in } \mathbb{M}(2,3,-2,3) \\ 0.10001 & \text{in } \mathbb{M}(2,5,-2,2) \end{cases}$$

Menge	$\lvert\gamma(x)-x\rvert$	$\frac{\lvert\gamma(x)-x\rvert}{\lvert x\rvert}$
$\mathbb{M}(10,3,-2,2)$	≈ 0.0003023059	≈ 0.0005595125
$\mathbb{M}(2,3,-2,3)$	≈ 0.0403023059	≈ 0.0745921412
$\mathbb{M}(2,5,-2,2)$	≈ 0.0091023059	≈ 0.0168466908

b) Für $\pi = 3.1415926535897931\ldots$:
$\gamma(x) = 0.314 \cdot 10^1$ in $\mathbb{M}(10,3,-2,2)$
Darstellung der Zahl im Binärsystem: $(3.1415926535897931)_{10} = (11.001001000011111\ldots)_2$
Somit gilt:

$$\gamma(x) = \begin{cases} 0.110 \cdot 2^2 = 3 & \text{in } \mathbb{M}(2,3,-2,3) \\ 0.11001 \cdot 2^2 = 3.125 & \text{in } \mathbb{M}(2,5,-2,2) \end{cases}$$

Menge	$\lvert\gamma(x)-x\rvert$	$\frac{\lvert\gamma(x)-x\rvert}{\lvert x\rvert}$
$\mathbb{M}(10,3,-2,2)$	≈ 0.0015926536	≈ 0.0005069574
$\mathbb{M}(2,3,-2,3)$	≈ 0.1415926536	≈ 0.0450703414
$\mathbb{M}(2,5,-2,2)$	≈ 0.0165926536	≈ 0.0052816057

Exercise 2:

Compute the smaller one of the two solutions to the quadratic equation

$$x^2 - 1.8x + 0.0001, \quad (a = 0.9,\, b = 0.0001)$$

using

a.) $x_1 = a - \sqrt{a^2 - b}$,

b.) $x_2 = a + \sqrt{a^2 - b}, \quad x_1 = \frac{b}{x_2}$.

Solve the equation in $\mathbb{M}(10, 3, -\infty, \infty)$ and $\mathbb{M}(10, 4, -\infty, \infty)$ and determine the relative error.

Solution:

exakte Berechnung	num. Erg. $(t = 4)$	num. Erg. $(t = 3)$
$a = 0.9$	.9000e0	.900e0
$b = 0.0001$	.1000e-3	.100e-3
$c = a \cdot a = 0.81$	.8100e0	.810e0
$d = c - b = 0.8099$	.8099e0	.810e0
$e = \sqrt{d} \approx 0.89994444272966$	.8999e0	.900e0
$x_1 = a - e \approx 0.55557270339 \cdot 10^{-4}$	.1000e-3	.000e0

Es ergeben sich die relativen Fehler

$$E_{rel_4} = \frac{|\Delta x_{1_4}|}{|x_1|} = \frac{0.4444272966 \cdot 10^{-4}}{0.55557270339 \cdot 10^{-4}} \approx 0.8,$$
$$E_{rel_3} = \frac{|\Delta x_{1_3}|}{|x_1|} = \frac{0.55557270339 \cdot 10^{-4}}{0.55557270339 \cdot 10^{-4}} = 1,$$

die auf Grund der Auslöschung bei $a - e$ ($|b| \ll a^2$) sehr groß werden.
Nach Methode (2) erfolgt die Berechnung von c, d, e wie oben, anschließend

exakte Berechnung	num. Erg. $(t = 4)$	num. Erg. $(t = 3)$
$x_2 = a + e \approx 1.79994444272966$	.1800e1	.180e1
$x_1 = b/x_2 = 0.55557270339 \cdot 10^{-4}$	.5556e-4	.556e-4

mit den relativen Fehlern

$$E_{rel_4} = \frac{0.272966 \cdot 10^{-8}}{0.55557270339 \cdot 10^{-4}} \approx 0.49 \cdot 10^{-4},$$
$$E_{rel_3} = \frac{0.4272966 \cdot 10^{-7}}{0.55557270339 \cdot 10^{-4}} \approx 0.77 \cdot 10^{-3}.$$

Hier gilt $E_{rel_i} < \mathbf{u}$, sodass man dem Ergebnis vertrauen darf.