Aufgabe 1 (5 Punkte)

Gegeben sei die Matrix A sowie der Vektor b mit

$$A = \begin{pmatrix} 12 & 644 & -96 \\ 6 & 0 & 0 \\ 0 & 322 & -47.99 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

a) Bestimmen Sie die LR–Zerlegung von A (exakt!). Geben Sie L und R explizit an.

b) Lösen Sie mittels der LR–Zerlegung aus a) das lineare Gleichungssystem $A \cdot x = b$ in **3-stelliger** Gleitpunktarithmetik.

c) Berechnen Sie das Residuum (Taschenrechnergenauigkeit). Geben Sie eine Begründung für das *schlechte* Resultat.

Lösung:

a) Wie man leicht sieht ist

$$A = \begin{pmatrix} 12 & 644 & -96 \\ 6 & 0 & 0 \\ 0 & 322 & -47.99 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 12 & 644 & -96 \\ 0 & -322 & 48 \\ 0 & 0 & 0.01 \end{pmatrix} = L \cdot R$$

b) Vorwärtseinsetzen ergibt

$$L \cdot y = b \longrightarrow y = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$$

Erst beim Rückwärtseinsetzen *wirkt* die 3-stellige Gleitpunktarithmetik

$$(R \mid b) = \left(\begin{array}{ccc|c} 12 & 644 & -96 & 2 \\ 0 & -322 & 48 & 0 \\ 0 & 0 & 0.01 & 3 \end{array}\right)$$

$$x_3 = \frac{3}{0.01} = 300$$

$$x_{2,3} = 0 - 48 \cdot 300 = -14400$$

$$x_2 = \frac{-14400}{-322} = 44.7$$

$$x_{1,2} = 2 - 644 \cdot 44.7 = 2 - 28800 = -28800$$

$$x_{1,3} = -28800 + 96 \cdot 300 = -28800 + 28800 = 0$$

oder

$$x_{1,3} = 2 + 96 \cdot 300 = 2 + 28800 = 28800$$

$$x_{1,2} = 2880 - 644 \cdot 44.7 = 28800 - 28800 = 0$$

$$x_1 = \frac{0}{12} = 0$$

aber

$$\tilde{x}_1 = 644 \cdot 44.7 - 96 \cdot 300 = 28800 - 28800 = 0$$

$$x_1 = \frac{2}{12} = 0.167$$

$$x = \begin{pmatrix} 0 \\ 44.7 \\ 300 \end{pmatrix} \quad \text{oder} \quad x = \begin{pmatrix} 0.167 \\ 44.7 \\ 300 \end{pmatrix}$$

c) Residuum:

$$A \cdot x - b = \begin{pmatrix} -15.2 \\ -1 \\ -6.6 \end{pmatrix} \quad \text{oder} \quad A \cdot x - b = \begin{pmatrix} -13.196 \\ 0.002 \\ -6.6 \end{pmatrix}$$

Das Residuum ist groß, die Rechnung ist mit Pivotisierung durchgeführt worden ($|l_{ij}| \leq 1$), also liegt es an der schlechten Kondition.

Aufgabe 2 (3 Punkte)

Gegeben sei das lineare Gleichungssystem $A \cdot x = b$ mit

$$A = \begin{pmatrix} 10000 & 5 \\ 7 & 0.1 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 0.52 \\ 0.69 \end{pmatrix}.$$

a) Bestimmen Sie die Kondition $\kappa_1(A)$ der Matrix. Ist A gut oder schlecht konditioniert?

b) Mit welchem relativen Fehler in x (in der 1-Norm) muß man rechnen, wenn statt des ursprünglichen Gleichungssystems $A \cdot x = b$ das gestörte Gleichungssystem $\tilde{A} \cdot x = \tilde{b}$ gelöst wird, wobei $\tilde{A}$ und $\tilde{b}$ Störungen von A bzw. b mit einem relativen Fehler von maximal 2% sind?

Lösung:

Sei $\|\cdot\| := \|\cdot\|_1$ (Spaltensummennorm).

a) Es ist $\|A\| = max(10000 + 7, 5 + 0.01) = 10007$. Ferner ist

$$A^{-1} = \frac{1}{10000 \cdot 0.01 - 5 \cdot 7} \begin{pmatrix} 0.1 & -5 \\ -7 & 10000 \end{pmatrix},$$

also ist $\|A^{-1}\| = \frac{1}{965} max(0.1 + 7, 5 + 10000) = \frac{10005}{965}$ und damit $\kappa_1(A) = 10007 \cdot \frac{10005}{965} = 103751.33$. A ist also sehr schlecht konditioniert.

b) Es gilt:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa_1(A)}{1 - \kappa_1(A)\frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).$$

Diese Formel darf aber NUR angewandt werden, wenn der Nenner größer Null ist. Hier ist laut Aufgabenstellung und Teil a):

$$\kappa_1(A)\frac{\|\Delta A\|}{\|A\|} = 103751 \cdot 0.02 = 2075 > 1.$$

Der Nenner ist also negativ, da die Kondition zu schlecht ist. Es kann also keine Aussage über den Ausgabefehler gemacht werden.

Aufgabe 3 (7 Punkte)

Gesucht sind die Lösungen des folgenden nichtlinearen Gleichungssystems:

$$\begin{aligned} x^2 + y^2 &= 4 \\ \tfrac{x^2}{16} + y^2 &= 1 \end{aligned}$$

a) Fertigen Sie eine Skizze an, die die Lage der Lösungen verdeutlicht. Bestimmen Sie für den 1. Quadranten einen *guten* ganzzahligen Startwert (x_0, y_0).

b) Gesucht ist nun die Lösung im ersten Quadranten. Geben Sie eine geeignete Fixpunktgleichung an, und weisen Sie hierfür die Voraussetzungen des Fixpunksatzes von Banach nach.

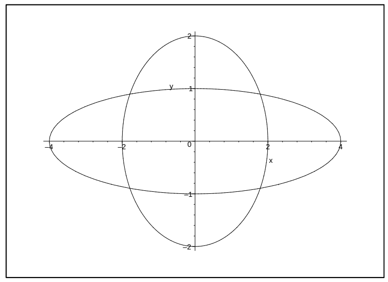
c) Wieviele Schritte sind (ausgehend von dem in a) gewählten Startwert) höchstens erforderlich, um eine Genauigkeit (welche Norm?) von $\varepsilon = 10^{-4}$ zu erzielen. Verwenden Sie die a–priori–Fehlerabschätzung.

d) Geben Sie eine a–posteriori–Fehlerabschätzung für (x_2, y_2) an.

e) Geben Sie die anderen Lösungen mit der selben Genauigkeit (ohne weitere Rechnung) an.

Lösung:

a) Die erste Gleichung beschreibt einen Kreis mit Radius 2, während die zweite einer Ellipse mit y-Achsenabschnitt 1 und x-Achsenabschnitt 4 entspricht. Die Lösungen des Gleichungssystems sind die Schnittpunkte des Kreises mit der Ellipse. Ein guter Startwert ist also $(x_0, y_0) = (2, 1)$.



b) Wir lösen die erste Gleichung nach x und die zweite nach y auf:

$$\begin{aligned} x &= \sqrt{4-y^2} &&=: f_1 \\ y &= \tfrac{1}{4}\sqrt{16-x^2} &&=: f_2 \end{aligned}$$

Gesucht ist nun die Lösung im ersten Quadranten, dazu wählen wir eine abgeschlossene Menge zur Anwendung des Fixpunksatzes von Banach. In y-Richtung kommt wegen der Form der Ellipse maximal das Intervall $[0,1]$ in Frage. Die Funktion f_1 ist monoton fallend auf dem Intervall, wir betrachten also zunächst die Randwerte:

$$f_1(0) = 2, \quad f_1(1) = \sqrt{3} = 1.732 < 2.$$

Da f_1 y auf x abbildet, bietet sich das Intervall $[1.5, 2]$ für die x-Werte an. Dort ist f_2 monoton fallend. Wir überprüfen wieder die Randwerte:

$$f_2(1.5) = 0.927 < 1, \quad f_2(2) = 0.866 > 0.5 > 0.$$

Also wird die abgeschlossene Menge $D = [1.5, 2] \times [0.5, 1]$ auf sich selbst abgebildet unter der Funktion

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f_1(y) \\ f_2(x) \end{pmatrix} = \begin{pmatrix} \sqrt{4-y^2} \\ \frac{1}{4}\sqrt{16-x^2} \end{pmatrix}.$$

Nun zur Kontraktion. Zunächst berechnen wir die Jacobi-Matrix:

$$F'\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -\frac{y}{\sqrt{4-y^2}} \\ -\frac{1}{4}\frac{x}{\sqrt{16-x^2}} & 0 \end{pmatrix}$$

Wir rechnen in der ∞-Norm, für die 1-Norm läuft der folgende Schritt analog. Zunächst berechnen wir die Maxima der Ableitungen. Da der Zähler jeweils monoton steigend und der Nenner monoton fallend ist, wird das Maximum am rechten Intervallrand angenommen. Es gilt also:

$$\max_{y\in[0.5,1]} |f_1'| = \max_{y\in[0.5,1]} \left| \frac{y}{\sqrt{4-y^2}} \right| = \frac{1}{\sqrt{3}} = 0.5774$$

$$\max_{x\in[1.5,2]} |f_2'| = \max_{x\in[1.5,2]} \frac{1}{4} \left| \frac{x}{\sqrt{16-x^2}} \right| = \frac{1}{4}\frac{2}{\sqrt{12}} = 0.1443$$

Damit ist $\|F'(x,y)\|_\infty = 0.5774$ und wir wählen $L = 0.6 < 1$. Damit ist F kontraktiv auf D und alle Voraussetzungen des Fixpunktsatzes von Banach sind erfüllt.

c) Aus der a-priori-Fehlerschätzung erhalten wir folgende Ungleichung für die Anzahl der benötigten Iterationen:

$$n \geq \frac{\ln\left(\frac{\|\mathbf{x}_1-\mathbf{x}_0\|}{\epsilon(1-L)}\right)}{\ln\frac{1}{L}} = \frac{\ln\left(\frac{\epsilon(1-L)}{\|\mathbf{x}_1-\mathbf{x}_0\|}\right)}{\ln L}.$$

Ausgehend von $(2,1)$ als Startwert erhalten wir $\mathbf{x}_1{=}(\sqrt{3}, \frac{\sqrt{3}}{2})$. Also ist

$$\|\mathbf{x}_1 - \mathbf{x}_0\|_\infty = \max(0.2679, 0.1340) = 0.2679$$

und mit $L = 0.6$ und $\epsilon = 10^{-4}$ ergibt sich:

$$n \geq \frac{\ln\left(\frac{0.2679}{10^{-4}\cdot 0.4}\right)}{\ln\frac{1}{0.6}} = 17.2...$$

Dementsprechend sind höchstens 18 Schritte erforderlich, um die Genauigkeit in der $\|\cdot\|_\infty$-Norm zu erfüllen.

d) Für (x_2, y_2) ergibt sich:

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \sqrt{4-y_1^2} \\ \frac{1}{4}\sqrt{16-x_1^2} \end{pmatrix} = \begin{pmatrix} 1.80278 \\ 0.9013878190 \end{pmatrix}.$$

Und die a-posteriori Fehlerschätzung liefert

$$\|\mathbf{x}_2 - \mathbf{x}^*\| \leq \frac{L}{1-L}\|\mathbf{x}_2 - \mathbf{x}_1\| = \frac{0.6}{0.4} \cdot 0.07072483 = 0.10608724.$$

e) Da sowohl der Kreis als auch die Ellipse symmetrisch zu beiden Achsen liegen, müssen wir ausgehend von der Lösung im ersten Quadranten nur einmal alle möglichen Kombinationen von Vorzeichen auswählen um jeweils eine Lösung zu finden. Mit (x^*, y^*) sind also auch $(-x^*, y^*)$, $(x^*, -y^*)$ und $(-x^*, -y^*)$ Lösungen.

Aufgabe 4 (5 Punkte)

Das Integral

$$I = \int_{-4}^{4} e^{-2\,x^2}\,dx$$

soll mit der summierten Trapezregel $T(h)$ und anschließender Romberg–Extrapolation zur Schrittweitenfolge $h_i := 8 \cdot 2^{-i},\ i = 0, 1, \ldots, 4$ approximiert werden.

a) Ergänzen Sie folgendes Extrapolationsschema:

$$\begin{array}{lllllllll}
T(h_0) = T_{00} = 0.00000 & & & & & & & & \\
& \searrow & & & & & & & \\
T(h_1) = T_{10} = 4.00000 & \rightarrow & T_{11} = \ldots & & & & & & \\
& \searrow & & \searrow & & & & & \\
T(h_2) = T_{20} = 2.00134 & \rightarrow & 1.33512 & \rightarrow & 1.06858 & & & & \\
& \searrow & & \searrow & & \searrow & & & \\
T(h_3) = T_{30} = \ldots & \rightarrow & 1.02801 & \rightarrow & T_{32} = \ldots & \rightarrow & 1.00656 & & \\
& \searrow & & \searrow & & \searrow & & \searrow & \\
T(h_4) = T_{40} = 1.25331 & \rightarrow & 1.24731 & \rightarrow & 1.26192 & \rightarrow & 1.26596 & \rightarrow & T_{44} = \ldots
\end{array}$$

Welcher Wert approximiert I am besten?

b) Geben Sie eine möglichst gute Fehlerabschätzung für $T(h_4)$ an.

c) Von welcher Ordnung ist (theoretisch) der Wert T_{43}?

Lösung:

Teil a) Zunächst berechnen wir den Wert $T(h_3) = T_{3,0}$. Allgemein ist die Trapezregel für $h \cdot n = b - a$ gegeben durch

$$T(h) = h\left(\frac{f(a)}{2} + \sum_{i=1}^{n-1} f(a + i\,h) + \frac{f(b)}{2}\right). \tag{1}$$

Für $h_i \cdot n_i = h_{i+1} \cdot n_{i+1} = b - a$ mit $n_{i+1} = 2\,n_i$ und $h_{i+1} = h_i/2$ können wir uns Rechenarbeit ersparen mit der Rekursionsformel

$$T(h_{i+1}) = \frac{T(h_i)}{2} + h_{i+1}\sum_{i=1}^{n_i} f(a + (2\,i-1)\,h_{i+1}).$$

Weiter gilt für den Integranden $f(x) = e^{-2\,x^2}$ Achsensymmetrie, also $f(-x) = f(x)$. Da die Integrationsgrenzen $b = 4 = -a$ symmetrisch zum Ursprung liegen, lautet somit die Rekursionsformel für $n_i \geq 2$

$$T(h_{i+1}) = \frac{T(h_i)}{2} + 2\,h_{i+1}\sum_{i=1}^{n_i/2} f(a + (2\,i-1)\,h_{i+1}) = \frac{T(h_i)}{2} + h_i\sum_{i=1}^{n_i/2} f(a + (2\,i-1)\,h_{i+1}).$$

Hier gilt

$$h_3 = 1, \quad h_2 = 2 \qquad \Longrightarrow \qquad n_3 = 8, \quad n_2 = 4 \geq 2,$$

und daher

$$T(h_3) = \frac{T(h_2)}{2} + h_2\,(f(-3) + f(-1)) = \frac{2.00134}{2} + 2.0 \cdot \big(0.152300 \cdot 10^{-8} + 0.135335\big) = 1.27134.$$

Hiermit ersparen wir uns 7 Funktionsauswertungen gegenüber der Auswertung von (1).

Allgemein lautet die Rekursionsformel des Romberg–Schemas

$$T_{i,j} = \frac{4^j\, T_{i,j-1} - T_{i-1,j-1}}{4^j - 1}, \qquad i = j, j+1, \ldots, n, \quad j = 1, 2, \ldots, n.$$

Die fehlenden Werte des Schemas ergeben sich damit zu

$$\begin{aligned}
T_{1,1} &= \frac{4^1 \cdot T_{1,0} - T_{0,0}}{4^1 - 1} = \frac{4 \cdot 4.00000 - 0.00000}{3} = 5.33333 \\
T_{3,2} &= \frac{4^2 \cdot T_{3,1} - T_{2,1}}{4^2 - 1} = \frac{16 \cdot 1.02801 - 1.33512}{15} = 1.00753 \\
T_{4,4} &= \frac{4^4 \cdot T_{4,3} - T_{3,3}}{4^4 - 1} = \frac{256 \cdot 1.26596 - 1.00656}{255} = 1.26698
\end{aligned}$$

Der Wert $T_{4,4}$ gibt (theoretisch) die beste Approximation an, da er von der Ordnung $h^{2\cdot 4+2} = h^{10}$ ist, während die anderen Werte alle niedrigerer Ordnung sind.

Teil b) Die Fehlerabschätzung für die summierte Trapezregel lautet

$$|T(h) - I| \leq \frac{h^2}{12}(b-a) \max_{z\in[a,b]} |f''(z)|.$$

Hier gilt $a = -4$, $b = 4$, $h = h_4 = 0.5$ und $f(x) = e^{-2\,x^2}$. Um die Extrema von f'' zu bestimmen, benötigen wir die Ableitungen bis zur Ordnung 3:

$$\begin{aligned}
f'(x) &= -4\,x\,e^{-2\,x^2}, \\
f''(x) &= (16\,x^2 - 4)\,e^{-2\,x^2} = 4\,(4\,x^2 - 1)\,e^{-2\,x^2}, \\
f'''(x) &= (48\,x - 64\,x^3)\,e^{-2\,x^2} = 16\,x\,(4\,x^2 - 3)\,e^{-2\,x^2}.
\end{aligned}$$

Die Nullstellen von f''' geben die Extremstellen von f'' an; diese sind $x_0 = -\sqrt{3}/2$, $x_1 = 0$ und $x_2 = \sqrt{3}/2$ mit den Werten $f''(x_0) = f''(x_2) = 8\,e^{-3/2} = 1.78504$ und $f''(x_1) = -4$, also

$$\max_{z\in[a,b]} |f''(z)| = 4$$

und somit insgesamt für den Fehler

$$|T(h_4) - I| \leq \frac{0.25}{12} \cdot 8 \cdot 4 = \frac{2}{3} = 0.66667.$$

(Tatsächlich beträgt der Fehler $-0.4137 \cdot 10^{-5}$.)

Teil c) Theoretisch nimmt die Ordnung in jeder Spalte um h^2 zu. Da die Trapezregel von der Ordnung h^2 ist, ist der Wert $T_{4,3}$ in der 3. Spalte (von 0 ab gezählt) gerade $h^{2\cdot 3+2} = h^8$.