Diplom – VP HM III/IV / **Numerik** 21. März 2002

Aufgabe N1

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

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 9 \\ 6 \\ 10 \end{pmatrix}.$$

a) Berechnen Sie die LR–Zerlegung von A mit Spaltenpivotisierung, d.h. $PA = LR$, wobei P eine geeignete Permutationsmatrix ist. Geben sie L und R explizit an.

b) Lösen Sie das Gleichungssystem $A \cdot x = b$ mit Hilfe der unter a) berechneten LR–Zerlegung.

Lösung:

a) Wir speichern die Permutation in einem Vektor, den wir an A anhängen.

$$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 1 \\ 0 & 1 & 2 & 2 \\ 2 & 1 & 3 & 3 \end{array}\right) \to \left(\begin{array}{c|cc|c} \multicolumn{3}{c|}{2 \quad\quad 1 \quad\quad 3} & 3 \\ \hline 0 & 1 & 2 & 2 \\ 1/2 & 3/2 & 1/2 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 2 & 1 & 3 & 3 \\ 1/2 & 3/2 & 1/2 & 1 \\ 0 & 2/3 & 5/3 & 2 \end{array}\right)$$

Und somit

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 2/3 & 1 \end{pmatrix} \quad \text{und} \quad R = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 3/2 & 1/2 \\ 0 & 0 & 5/3 \end{pmatrix} \quad \text{sowie} \quad p = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}.$$

b) Wir benutzen p für die Permutation von b und erhalten:

(i) Vorwärtseinsetzen:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 10 \\ 1/2 & 1 & 0 & 9 \\ 0 & 2/3 & 1 & 6 \end{array}\right) \Big\downarrow \to y = \begin{pmatrix} 10 \\ 4 \\ 10/3 \end{pmatrix}$$

(ii) Rückwärtseinsetzen:

$$\left(\begin{array}{ccc|c} 2 & 1 & 3 & 10 \\ 0 & 3/2 & 1/2 & 4 \\ 0 & 0 & 5/3 & 10/3 \end{array}\right) \Big\uparrow \to x = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Aufgabe N2

Gegeben sei die folgende Meßreihe

x_i	1	2	4
f_i	2.5	1.5	2

Aufgrund theoretischer Überlegungen weiß man, daß diese Meßdaten einer Funktion

$$f(x) = a \cdot (x - 1) + \frac{b}{x}$$

genügen. Bestimmen Sie die Parameter a und b optimal im Sinne der kleinsten Fehlerquadrate. Formulieren Sie dazu zunächst das entsprechende lineare Ausgleichsproblem. Lösen sie dieses anschließend mittels Normalgleichungen.

Lösung:

Das lineare Ausgleichsproblem lautet $\|A \cdot x - b\|_2 \to \min$. Dabei ist $x = (a, b)^T$ und

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1/2 \\ 3 & 1/4 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 5/2 \\ 3/2 \\ 2 \end{pmatrix}.$$

Die Dazugehörigen Normalgleichungen ($A^T A \cdot x = A^T b$) ergeben

$$\begin{pmatrix} 10 & 5/4 \\ 5/4 & 21/16 \end{pmatrix} x = \begin{pmatrix} 15/2 \\ 15/4 \end{pmatrix}.$$

Wir lösen das Gleichungssystem ausnahmsweise mit Gaußelimination (eigentlich Cholesky):

$$\begin{pmatrix} 10 & 5/4 & | & 15/2 \\ 5/4 & 21/16 & | & 15/4 \end{pmatrix} \to \begin{pmatrix} 10 & 5/4 & | & 15/2 \\ 0 & 37/32 & | & 45/16 \end{pmatrix} \longrightarrow x = \begin{pmatrix} 33/74 \\ 90/37 \end{pmatrix} = \begin{pmatrix} 0.445946 \\ 2.43243 \end{pmatrix}.$$

Wir haben somit als Lösung:

$$f(x) = \frac{33}{74}(x-1) + \frac{90}{37 \cdot x}$$

Aufgabe N3
Gesucht ist eine Nullstelle der Funktion

$$f(x) = e^{-x}\,\cos(x) - x$$

im Intervall $I = [0, 1]$.

a) Zeigen Sie mit Hilfe des Banach'schen Fixpunktsatzes, daß f in I genau eine Nullstelle besitzt.

b) Führen Sie, ausgehend vom Startwert $x_0 = 0.5$, zwei Iterationen des Fixpunktverfahrens durch und geben Sie für die letzte Approximation eine Fehlerabschätzung an.

c) Nennen Sie ein Verfahren, das ein besseres Konvergenzverhalten hat (Begründung!). Geben Sie die Iterationsvorschrift für dieses Problem explizit an.

Lösung:

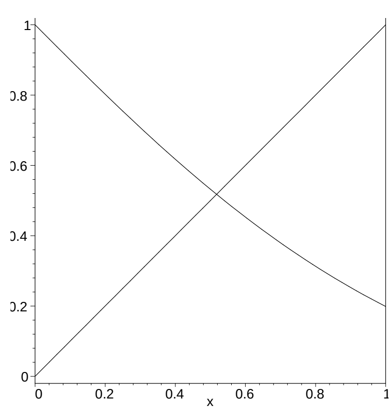
Teil a) Zunächst muß das Nullstellenproblem in ein Fixpunktproblem transformiert werden:

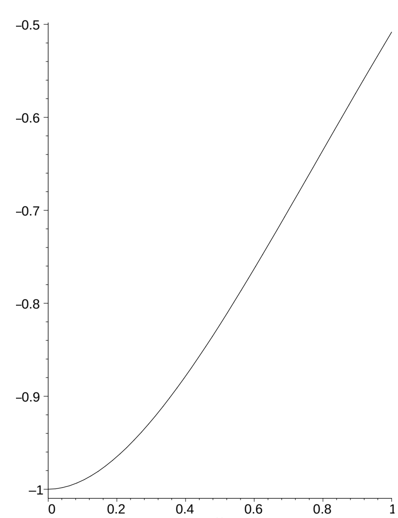
$$f(x) \stackrel{!}{=} 0 \qquad \Longleftrightarrow \qquad F(x) := e^{-x}\,\cos(x) \stackrel{!}{=} x.$$

Die Ableitung ist gegeben durch

$$F'(x) := -e^{-x}\,(\cos(x) + \sin(x)) < 0.$$

Eine Skizze der Funktion (links zusammen mit der Geraden $y = x$) und ihrer Ableitung (rechts) liefert:





Für den Banach'schen Fixpunktsatz müssen drei Punkte nachgewiesen werden: 1. Abgeschlossenheit des Intervalls (offensichtlich), 2. Selbstabbildung und 3. Kontraktivität von F. Wegen der Negativität der Ableitung ist F streng monoton fallend, und es gilt für $x \in I$

$$F(x) \in [F(1), F(0)] = [e^{-1}\,\cos(1), 1] = [0.19877, 1] \subset I,$$

und damit die Selbstabbildung. Wegen $F'(0) = -1$ ist F auf I aber nicht kontraktiv, daher arbeiten wir weiter mit dem (abgeschlossenen) Intervall

$$\tilde{I} := [0.19, 1],$$

für das wir die Selbstabbildung nicht mehr nachzuweisen brauchen. Weiterhin gilt

$$F''(x) = 2\,e^{-x}\,\sin(x) > 0, \qquad x \in \tilde{I},$$

also ist F' monoton steigend in $\tilde{I}$. Wegen der Negativität von F' auf $\tilde{I}$ gilt daher

$$\max_{x\in\tilde{I}} |F'(x)| = |F'(0.19)| = 0.96826 < 1,$$

damit ist F kontraktiv mit der Lipschitzkonstanten $L := 0.97$. Mit dem Fixpunktsatz von Banach hat F also genau einen Fixpunkt x^* in $\tilde{I}$ und somit f genau eine Nullstelle $x^* \in \tilde{I}$.

Betrachte nun noch das Restintervall $\hat{I} := I \setminus \tilde{I}$. Für $x \in \hat{I}$ gilt also $x < 0.19$ und somit wegen der Monotonie von F

$$F(x) > F(0.19) = 0.81208 > 0.19 > x,$$

also kann kein weiterer Fixpunkt von F in $\hat{I}$ liegen und damit auch keine weitere Nullstelle von f.

Teil b) Zwei Iterationen zum Startwert $x_0 = 0.5$ liefern die Werte

$$\begin{aligned} x_1 &:= F(x_0) = e^{-x_0}\cos(x_0) = 0.53228, \\ x_2 &:= F(x_1) = e^{-x_1}\cos(x_1) = 0.50602. \end{aligned}$$

Mit dem Banach'schen Fixpunktsatz erhalten wir mit der Lipschitzkonstanten L aus Teil a) die a-posteriori-Fehlerabschätzung

$$|x_2 - x^*| \le \frac{L}{1-L}|x_2 - x_1| = \frac{0.97}{0.03}|0.50602 - 0.53228| = 32.333 \cdot 0.026264 = 0.84919.$$

(Diese ist wegen des hohen Wertes von L eigentlich unbrauchbar.)

Teil c) Das Fixpunktverfahren ist nur von der Ordnung $p = 1$, während das Newton–Verfahren von der Ordnung $p = 2$ ist (bei einfachen Nullstellen, was hier der Fall ist wegen $f'(x) < 0$ in I). Dieses lautet

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)},$$

wobei jetzt natürlich die ursprüngliche Funktion $f(x) = e^{-x}\cos(x) - x$ mit der Ableitung

$$f'(x) = -e^{-x}\,(\cos(x) + \sin(x)) - 1$$

zu nehmen ist.

Aufgabe N4

Das Integral

$$I := \int_0^4 f(x)\,dx, \qquad f(x) := \arctan(x),$$

soll mit summierten Newton–Cotes–Formeln näherungsweise bestimmt werden.

a) Approximieren Sie I mit der summierten Simpsonregel für $n = 2$ und geben Sie eine möglichst gute Fehlerabschätzung an.

b) Bestimmen Sie die Anzahl n der Teilintervalle so, dass die summierte Trapezregel den Fehler $\epsilon = 10^{-4}$ nicht überschreitet.

c) Skizzieren Sie die summierte Trapezregel mit $n = 2$ für dieses Integral. Warum ist der approximative Wert für diese Funktion kleiner als der tatsächliche Integralwert? Gilt das für jede Anzahl n von Teilintervallen?

Hinweis: Es gilt

$$f''(x) = \frac{-2\,x}{(1+x^2)^2},\; f'''(x) = \frac{2\,(3\,x^2-1)}{(1+x^2)^3},\; f^{(4)}(x) = \frac{24\,x\,(1-x^2)}{(1+x^2)^4},\; f^{(5)}(x) = \frac{24\,(5\,x^4-10\,x^2+1)}{(1+x^2)^5}.$$

Lösung:

Teil a) Für $n=2$ und die Intervallgrenzen $a=0$ sowie $b=4$ gilt für die Schrittweite und die Stützstellen

$$h = \frac{b-a}{n} = \frac{4-0}{2} = 2, \qquad x_k = a + k\,h = 2\,k, \quad k = 0,1,2.$$

Damit ergibt sich für die Simpsonregel

$$\begin{aligned} S(h) &= \sum_{k=0}^{n-1} \frac{h}{6}\left(f(x_k) + 4\,f(x_k+h/2) + f(x_{k+1})\right) \\ &= \frac{h}{6}\left(f(a) + 2\sum_{k=1}^{n-1} f(x_k) + 4\sum_{k=0}^{n-1} f(x_k+h/2) + f(b)\right) \\ &= \frac{h}{6}\left(f(x_0) + 4\,f(x_0+h/2) + 2\,f(x_1) + 4\,f(x_1+h/2) + f(x_2)\right) \\ &= \frac{1}{3}\left(f(0) + 4\,f(1) + 2\,f(2) + 4\,f(3) + f(4)\right) \\ &= \frac{1}{3}\left(0.0 + 4\cdot 0.785398 + 2\cdot 1.10715 + 4\cdot 1.24905 + 1.32582\right) \\ &= 3.89264. \end{aligned}$$

Für den Fehler der Simpsonregel gilt

$$|I - S(h)| \le \sum_{k=0}^{n-1} \frac{1}{90}\left(\frac{h}{2}\right)^5 |f^{(4)}(\xi_k)| \le n\cdot\frac{1}{90}\left(\frac{h}{2}\right)^5 \max_{z\in[a,b]} |f^{(4)}(z)| = \frac{1}{45}\max_{z\in[a,b]} |f^{(4)}(z)|,$$

wobei $x_k \le \xi_k \le x_{k+1}$, $k = 0,\dots,n-1$ gilt. Um die Extrema der vierten Ableitung zu bestimmen, berechnen wir die Nullstellen der fünften Ableitung (Hinweis!), wobei eine biquadratische Gleichung zu lösen ist:

$$f^{(5)}(x) = 0 \quad\Longleftrightarrow\quad 5\,x^4 - 10\,x^2 + 1 = 0 \quad\Longleftrightarrow\quad x = \pm\sqrt{1 \pm \frac{2}{5}\sqrt{5}} = \begin{cases} \pm 0.32492 \\ \pm 1.3764 \end{cases}.$$

Damit gilt

$$\begin{aligned} \max_{z\in[a,b]} |f^{(4)}(z)| &= \max\{|f^{(4)}(0)|,\ |f^{(4)}(0.32492)|,\ |f^{(4)}(1.3764)|, |f^{(4)}(4)|\} \\ &= \max\{0,\ 4.6686,\ 0.42096,\ 0.017241\} \;=\; 4.6686, \end{aligned}$$

und somit

$$|I - S(h)| \le \frac{4.6686}{45} = 0.10375.$$

Teil b) Für den Fehler der Trapezregel gilt

$$\begin{aligned} |I - T(h)| &\le \sum_{k=0}^{n-1} \frac{h^3}{12}|f''(\xi_k)| \;\le\; n\cdot\frac{h^3}{12}\max_{z\in[a,b]} |f''(z)| \\ &= \frac{n}{12}\cdot\left(\frac{b-a}{n}\right)^3 \max_{z\in[a,b]} |f''(z)| \;=\; \frac{(b-a)^3}{12} n^{-2} \max_{z\in[a,b]} |f''(z)| \\ &= \frac{16}{3} n^{-2} \max_{z\in[a,b]} |f''(z)| \;\overset{!}{\le}\; \epsilon, \end{aligned}$$

wobei nun h und n natürlich variabel sind und $x_k \le \xi_k \le x_{k+1}$, $k = 0,\dots,n-1$ gilt. Aufgelöst nach n ergibt sich

$$n^2 \ge \frac{16}{3}\cdot\frac{\max_{z\in[a,b]}|f''(z)|}{\epsilon} \qquad\Longrightarrow\qquad n \ge \sqrt{\frac{16}{3}\cdot\frac{\max_{z\in[a,b]}|f''(z)|}{\epsilon}}.$$

Um die Extrema der zweiten Ableitung zu bestimmen, setzen wir die dritte Ableitung zu Null und erhalten

$$f'''(x) = 0 \quad \Longleftrightarrow \quad 3\,x^2 - 1 = 0 \quad \Longleftrightarrow \quad x = \pm\frac{\sqrt{3}}{3} = \pm 0.57735.$$

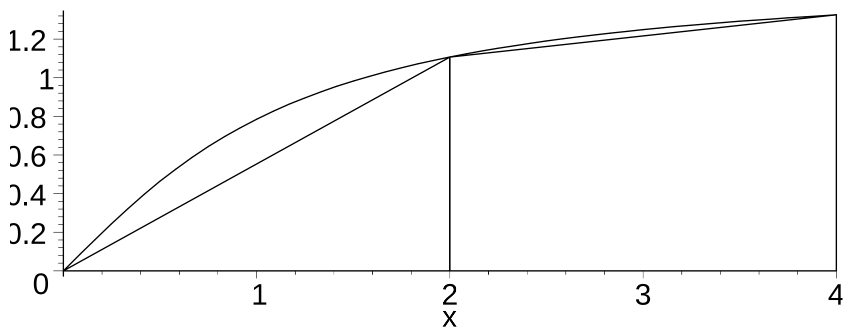
Also gilt

$$\begin{aligned} \max_{z\in[a,b]} |f''(z)| &= \max\{|f''(0)|,\ |f''(0.57735)|,\ |f''(4)|\} \\ &= \max\{0,\ 0.64952,\ 0.027682\} \ = \ 0.64952, \end{aligned}$$

und mit $\epsilon = 10^{-4}$ erhalten wir

$$n \geq 186.12\ldots \qquad \Longrightarrow \qquad n \geq 187.$$

Teil c) Graphisch ergibt die Trapezregel für $n = 2$ folgendes Resultat:



Wie auch die Skizze zeigt, ist f wegen $f''(x) < 0$ oder auch wegen

$$f(x) \geq \frac{x_{k+1} - x}{h}\, f(x_k) + \frac{x - x_k}{h}\, f(x_{k+1}), \quad x_k \leq x \leq x_{k+1} = x_k + h,$$

eine konkave Funktion. Damit liegt die Verbindungsstrecke zwischen $(x_k, f(x_k)$ und $(x_{k+1}, f(x_{k+1})$ immer unterhalb vom Graphen dieser Funktion. Da die summierte Trapezregel geometrisch genau die Fläche zwischen x–Achse und diesen Verbindungsstrecken bedeutet, während das Integral die Fläche zwischen x–Achse und dem Funktionsgraphen angibt, muß der Wert der summierten Trapezregel immer kleiner sein als der tatsächliche Integralwert. Dieses ist natürlich unabhängig von der Anzahl n der Teilintervalle und gilt für jeden Bereich $[a, b]$, in dem $f''(x) < 0$ ist, also für $0 \leq a < b$.