Interdisziplinäres Projekt

im Rahmen der Diplomhauptprüfung für Informatiker im Nebenfach Mathematik

an der

Technischen Universität München

Fakultät für Informatik

Thema:	**Musterlösungen zur Vorlesung Konkrete Mathematik**
Bearbeiter:	Nils Nitsch Fred-Hartmann-Weg 19 85435 Erding
Prüfer:	Professor Dr. Thomas Huckle Technische Universität München Lehrstuhl für Informatik mit Schwerpunkt Wissenschaftliches Rechnen
Abgabedatum:	29. September 2005

Inhaltsverzeichnis

1 Rundungsfehler

1.1 Aufgabe 1 (Buch, Übung: 24)

1.1.1 Aufgabenstellung

Geben Sie alle Funktionen $f(x)$ an, die für sämtliche x eine konstante Konditionszahl $cond_x = k$ besitzen.
Insbesondere: für welche Funktion wird der Fehler in den Eingangsdaten überhaupt nicht verstärkt ($k = 0$)?

1.1.2 Lösung

Tip

Finden Sie die Lösung der Differentialgleichung

$$y' = k \cdot \frac{y}{x} \quad \text{mit } , k \in \mathbb{R}$$

(z.B. Nachschlagen in einer Formelsammlung)

Lösungsschritt

Die Lösung der Differentialgleichung lautet:

$$f(x) = x^k$$

Lösungsschritt

Daraus folgt:

$$cond_x = \left| \frac{x \cdot f'(x)}{f(x)} \right| = \left| \frac{x \cdot k \cdot x^{k-1}}{x^k} \right| = \left| \frac{x^k \cdot k}{x^k} \right| = |k|$$

Lösungsschritt

Insbesondere gilt für $k = 0$:

$$cond_x = 0$$

D.h. nur für $f(x) = const$ wird der Fehler in den Eingangsdaten überhaupt nicht verstärkt.

1.2 Aufgabe 2 (Buch, Übung: 25)

1.2.1 Aufgabenstellung

Ersetzen Sie den Differentialoperator $d_x := d/dx$ für stetig differenzierbare Funktionen $f \in C^1[0,1]$ durch den Differenzenoperator Δ_x^h mittels

$$\Delta_x^h f(x) := \frac{f(x+h) - f(x)}{h},\ h > 0.$$

Zur Vereinfachung betrachten Sie nur den Rundungsfehler, der bei der Differenz entsteht und den Fehler, der durch die Diskretisierung, d.h. das Ersetzen des Differential- durch den Differenzoperator, gemacht wird. Der resultierende Gesamtfehler f_d ergibt sich aus der Bildung der Differenz $-_{\mathbb{M}}$ und der angenäherten Ableitung Δ_x^h in Abhängigkeit der Schrittweite h. Für welches h ist der Gesamtfehler minimal?

1.2.2 Lösung

Tip

Analysieren Sie den Diskretisierungsfehler mit Hilfe der Taylor-Reihe für $f(x+h)$. Geben Sie auch die auftretenden Konstanten an! Der bei der Differenz auftretende Fehler sei ϵ.

Lösungsschritt

Für den Diskretisierungsfehler gilt (mit Hilfe der Taylor-Reihe):

$$f(x+h) = f(x) + f'(x) \cdot h + \frac{f''(\xi)}{2!} \cdot h^2, \xi \in [x, x+h]$$

$$\Rightarrow f'(x) = \frac{f(x+h) - f(x)}{h} - \frac{f''(\xi)}{2} \cdot h$$

$$\Rightarrow \left| f'(x) - \Delta_x^h f(x) \right| \leq \left| \frac{f''(\xi)}{2} \cdot h \right|$$

Tip

Rundungsfehler bei der Differenzbildung:

$$a -_{\mathbb{M}} b = (a - b) \cdot (1 + \epsilon)$$

Lösungsschritt

Rundungsfehler bei der Differenzbildung:

$$\widetilde{\Delta}_x^h f(x) = \Delta_x^h f(x) \cdot (1 + \epsilon)$$

wobei ϵ der relative Fehler bei der Subtraktion ist mit $|\epsilon| \leq \epsilon_{masch}$.
Daraus ergibt sich für den Rundungsfehler:

$$\left|\widetilde{\Delta}_x^h f(x) - \Delta_x^h f(x)\right| \leq \left|\Delta_x^h f(x)\right| \cdot \epsilon_{masch} \leq \frac{|f(x+h)| + |f(x)|}{|h|} \cdot \epsilon_{masch} \leq c_1 \cdot \epsilon_{masch} \frac{1}{h}$$

Lösungsschritt

Der Gesamtfehler lautet damit:

$$\left|\frac{\widetilde{\Delta}_x^h f(x) - f'(x)}{f'(x)}\right| \leq \frac{|\widetilde{\Delta}_x^h f(x) - \Delta_x^h f(x)| + |\Delta_x^h f(x) - f'(x)|}{|f'(x)|} \leq c_1 \cdot \epsilon_{masch} \cdot \frac{1}{h} + \underbrace{\left|\frac{f''(\xi)}{2f'(x)}\right|}_{:=c_2} \cdot h$$

$$= c_1 \cdot \frac{1}{h} \cdot \epsilon_{masch} + c_2 \cdot h = f_d(h)$$

Der Gesamtfehler ist nun minimal für $f_d' = 0$, d.h.

$$f_d' = -\frac{1}{h^2} \cdot c_1 \cdot \epsilon_{masch} + c_2 = 0$$

$$\Leftrightarrow h^2 = \frac{c_1}{c_2} \cdot \epsilon_{masch}$$

$$\Rightarrow h = \sqrt{\frac{c_1}{c_2} \cdot \epsilon_{masch}} \Rightarrow h = \mathcal{O}(\sqrt{\epsilon_{masch}})$$

1.3 Aufgabe 3 (Buch, Übung: 26)

1.3.1 Aufgabenstellung

Analysieren Sie mit Hilfe der Epsilontik den Ausdruck $y = a^2 - b^2$.

a) Berechnen Sie die Konditionszahlen bezüglich a und b. Für welchen Fall ist das Problem schlecht konditioniert?

b) Zeigen Sie, dass für den relativen Fehler des Resultats gilt

$$\epsilon_y = -\frac{2a^2}{a^2-b^2} \cdot \epsilon_a + \frac{2b^2}{a^2-b^2} \cdot \epsilon_b - \frac{a^2}{a^2-b^2} \cdot \epsilon_1 + \frac{b^2}{a^2-b^2} \cdot \epsilon_2 - \epsilon_3.$$

Die Werte ϵ_a und ϵ_b repräsentieren die relativen Fehler in den Eingabedaten a und b und ϵ_1, ϵ_2 und ϵ_3 die relativen Fehler in den Berechnungen von $a \cdot_{\mathbb{M}} a, b \cdot_{\mathbb{M}} b$ und $a^2 -_{\mathbb{M}} b^2$.

c) Führen Sie die Analyse auch für den alternativen Berechnungsweg $(a +_{\mathbb{M}} b) \cdot_{\mathbb{M}} (a -_{\mathbb{M}} b)$ durch. Vergleichen und diskutieren Sie die Ergebnisse.

d) Welchen Berechnungsweg würden Sie empfehlen? Warum?

1.3.2 Lösung

zu a)

Lösungsschritt

Für $y = a^2 - b^2$ gilt bezüglich a und b:

$$f(a) = a^2 - b^2$$

$$f(b) = a^2 - b^2$$

d.h.

1) bzgl. a: $f(a) = a^2 - b^2$; $f'(a) = 2a$

$$\Rightarrow cond_a = \left| \frac{a \cdot (2a)}{a^2 - b^2} \right| = \left| \frac{2a^2}{a^2 - b^2} \right|$$

2) bzgl. b: $f(b) = a^2 - b^2$; $f'(b) = -2b$

$$\Rightarrow cond_b = \left| \frac{b \cdot (-2b)}{a^2 - b^2} \right| = \left| -\frac{2b^2}{a^2 - b^2} \right|$$

Lösungsschritt

D.h. das Problem ist schlecht konditioniert für den Fall, dass $a \approx b$

<u>zu b)</u>

Lösungsschritt

$f_{rel}(y) = \frac{y-\tilde{y}}{y}$

$$\begin{aligned}
\tilde{y} &= a(1+\epsilon_a) \cdot_{\mathbb{M}} a(1+\epsilon_a) -_{\mathbb{M}} b(1+\epsilon_b) \cdot_{\mathbb{M}} b(1+\epsilon_b) \\
&\doteq a^2(1+2\epsilon_a) \cdot (1+\epsilon_1) -_{\mathbb{M}} b^2(1+2\epsilon_b)(1+\epsilon_2) \\
&\doteq ((a^2+\epsilon_1 a^2 + 2\epsilon_a a^2) - (b^2+\epsilon_2 b^2 + 2\epsilon_b b^2)) \cdot (1+\epsilon_3) \\
&\doteq a^2 - b^2 + \epsilon_a 2a^2 - \epsilon_b 2b^2 + \epsilon_1 a^2 - \epsilon_2 b^2 + \epsilon_3(a^2-b^2)
\end{aligned}$$

Lösungsschritt

Nun gilt für den relativen Fehler:

$$\begin{aligned}
f_{rel}(y) &= \frac{y-\tilde{y}}{y} \\
&= \frac{a^2-b^2-\tilde{y}}{a^2-b^2} \\
&= -\frac{2a^2}{a^2-b^2} \cdot \epsilon_a + \frac{2b^2}{a^2-b^2} \cdot \epsilon_b - \frac{a^2}{a^2-b^2} \cdot \epsilon_1 + \frac{b^2}{a^2-b^2} \cdot \epsilon_2 - \epsilon_3 \qquad \square
\end{aligned}$$

<u>zu c)</u>

Lösungsschritt

$y = (a+b)(a-b)$

$$\begin{aligned}
\tilde{y} &= (a(1+\epsilon_a) +_{\mathbb{M}} b(1+\epsilon_b)) \cdot_{\mathbb{M}} (a(1+\epsilon_a) -_{\mathbb{M}} b(1+\epsilon_b)) \\
&\doteq ((a + a\epsilon_a + b + b\epsilon_b) \cdot (1+\epsilon_1)) \cdot_{\mathbb{M}} ((a + a\epsilon_a - b - b\epsilon_b)(1+\epsilon_2)) \\
&\doteq (a^2 - b^2 + (a^2-b^2)\epsilon_1 + (a^2-b^2)\epsilon_2 + 2a^2\epsilon_a - 2b^2\epsilon_b) \cdot (1+\epsilon_3) \\
&\doteq a^2 - b^2 + 2a^2\epsilon_a - 2b^2\epsilon_b + (a^2-b^2)\epsilon_1 + (a^2-b^2)\epsilon_2 + (a^2-b^2)\epsilon_3
\end{aligned}$$

Somit gilt für den relativen Fehler:

$$f_{rel} = -\frac{2a^2}{a^2-b^2}\epsilon_a + \frac{2b^2}{a^2-b^2}\epsilon_b - \epsilon_1 - \epsilon_2 - \epsilon_3$$

Lösungsschritt

In dem Verfahren aus a) liegt ein höhere relativer Fehler vor, da ϵ_1 und ϵ_2 durch den vorangestellten Faktor $\frac{a^2}{a^2-b^2}$ bzw. $\frac{b^2}{a^2-b^2}$ verstärkt werden.

<u>zu d)</u>

Lösungsschritt

Hier ist das Verfahren aus c) zu empfehlen, da es weniger Fehleranfällig ist und zudem auch nur 3 Rechenoperationen benötigt, d.h. gleich schnell ist.

1.4 Aufgabe 4 (Buch, Übung: 27)

1.4.1 Aufgabenstellung

Entscheiden Sie, welche der folgenden Ausdrücke fehleranfälliger ist. Versuchen Sie gegebenenfalls ein besseres Verfahren zur Berechnung anzugeben und untersuchen Sie auch die Kondition.

a) $2\sin^2\frac{x}{2} = 1 - \cos x$ für $x \approx 0$

b) $\log(\frac{x}{y}) = \log(x) - \log(y)$ für $x > 0$ und $y > 0$

c) $f(x) = 1 + \sin x$

d) $f(x) = (x-1)^2 - (x+1)^2$ für $x \approx 0$

e) $f(x) = \exp(x) - 1 - x$ für $x \approx 0$

1.4.2 Lösung

zu a)

Lösungsschritt

Um den Rundungsfehler im Verlauf der verschiedenen Berechnungsarten zu verfolgen, betrachten wir ungestörte Eingabedaten, also x ohne Rundungsfehler. Es ergibt sich für die jeweiligen relativen Fehler im Endergebnis bei Verwendung der rechten Formel

$$\begin{aligned} f_{rel} &= \frac{[1-(1+\epsilon_1)\cos x](1+\epsilon_2) - (1-\cos x)}{1-\cos x} \\ &\doteq \frac{\epsilon_2 - \cos x(\epsilon_1+\epsilon_2)}{1-\cos(x)} \\ &= \epsilon_2 - \epsilon_1\frac{\cos x}{1-\cos x} \end{aligned}$$

Da der Term $\frac{\cos x}{1-\cos x}$ beliebig groß werden kann, gibt es keine obere Schranke für den relativen Fehler. Diese Berechnungsvorschrift ist also nicht geeignet.

Lösungsschritt

Für die linke Formel ergibt sich ein relativer Fehler von:

$$\begin{aligned} \tilde{y} &= 2(\sin\frac{x}{2}\cdot(1+\epsilon_1))^2\cdot(1+\epsilon_2) \\ &= 2\sin^2\frac{x}{2}\cdot(1+\epsilon_1)^2\cdot(1+\epsilon_2) \end{aligned}$$

$$\Rightarrow |f_{rel}| = 2|\epsilon_1| + |\epsilon_2| \leq \epsilon_{masch}$$

Diese Art der Berechnung ist offenbar stabil, wobei der Fehler im Bereich der Maschinengenauigkeit bleibt.

Tip

Reihenentwicklung von sin und cos einsetzten.

Lösungsschritt

Die Konditionszahl der Funktion lautet:

$$cond_x = \frac{x \cdot \sin x}{1 - \cos x}$$

Für $x \approx 0$ kann man die Reihenentwicklung von cos und sin einsetzten und erhält in erster Näherung

$$cond_x = \frac{x^2 + \dots}{1 - (1 - x^2/2 + \dots)} \approx 2$$

<u>zu b)</u>

Lösungsschritt

Für den linken Term ergibt sich:

$$\begin{aligned}\tilde{y} &= \log\left(\frac{x}{x} \cdot (1 + \epsilon_1)\right) \cdot (1 + \epsilon_2) \\ &= \left(\log\left(\frac{x}{y}\right) + \log(1 + \epsilon_1)\right) \cdot (1 + \epsilon_2)\end{aligned}$$

Daraus ergibt sich für den relativen Fehler:

$$\begin{aligned}f_{rel} &= \frac{\left|\tilde{y} - \log \frac{x}{y}\right|}{\left|\log \frac{x}{y}\right|} \\ &= \left|(1 + \epsilon_2)\left(1 + \frac{\log(1 + \epsilon_1)}{\log \frac{x}{y}}\right) - 1\right| \\ &\doteq \left|\epsilon_2 + \frac{\log(1 + \epsilon_1)}{\log \frac{x}{y}}\right|\end{aligned}$$

Lösungsschritt

Für den rechten Term ergibt sich:

$$\tilde{y} = ((1 + \epsilon_1) \log x - (1 + \epsilon_2) \log y) \cdot (1 + \epsilon_3)$$

Daraus ergibt sich für den relativen Fehler:

$$\begin{aligned} f_{rel} &= \frac{|\tilde{y} - \log x + \log y|}{|\log x - \log y|} \\ &= \left|(1+\epsilon_3)\left(1 + \frac{\epsilon_1 \log x - \epsilon_2 \log y}{\log x - \log y}\right)\right| \\ &\doteq \left|\epsilon_3 + \frac{\epsilon_1 \log x - \epsilon_2 \log y}{\log x - \log y}\right| \end{aligned}$$

Da $\log(1+\epsilon_1) \approx \epsilon_1$ ist $\log\left(\frac{x}{y}\right)$ i.A. günstiger. Falls jedoch $x \approx y \approx 1$ führt der untere Ausdruck zu besseren Ergebnissen.

Lösungsschritt

Die Konditionszahlen für beide Terme lauten:

$$cond_x = \frac{1}{\log \frac{x}{y}} \text{ und } cond_y = \frac{-1}{\log \frac{x}{y}}$$

Für $x \approx y$ ist dieses Problem schlecht konditioniert, daher kann man kein stabiles Verfahren erwarten.

zu c)

Lösungsschritt

Es gilt:

$$\begin{aligned} \tilde{y} &= (1 + \sin x(1+\epsilon_1)) \cdot (1+\epsilon_2) \\ &\doteq 1 + \sin x + \sin x \epsilon_1 + (1+\sin x)\epsilon_2 \end{aligned}$$

Daraus ergibt sich für den relativen Fehler:

$$\begin{aligned} f_{rel} &= \left|\frac{\tilde{y} - 1 - \sin x}{1 + \sin x}\right| \\ &= \left|\frac{\sin x}{1+\sin x} \cdot \epsilon_1 + \epsilon_2\right| \end{aligned}$$

Da der Term $\frac{\sin x}{1+\sin x}$ für den Fall $x \approx \frac{3}{2}\pi + k \cdot 2\pi$ beliebig groß werden kann, gibt es in diesem Fall keine obere Schranke für den relativen Fehler. Diese Berechnungsvorschrift ist hier also nicht geeignet.

Tip

Einsetzen von $x = \frac{x}{2} + \frac{x}{2}$ und umformen.

Lösungsschritt

Formt man den Term $1 + \sin x$ um, so erhält man:

$$1 + \sin x = 1 + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)$$

für den gilt:

$$\begin{aligned}
\tilde{y} &= \left[1 + 2 \cdot (\sin\left(\frac{x}{2}\right)(1+\epsilon_1)) \cdot (\cos\left(\frac{x}{2}\right)(1+\epsilon_2)) \cdot (1+\epsilon_3)\right](1+\epsilon_4) \\
&= 1 + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right) + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)(\epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4) + \epsilon_4
\end{aligned}$$

Daraus ergibt sich für den relativen Fehler:

$$\begin{aligned}
f_{rel} &= \left| \frac{\tilde{y} - 1 - 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)}{1 + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)} \right| \\
&= \left| \frac{2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)(\epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4) + \epsilon_4}{1 + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)} \right|
\end{aligned}$$

Hier erhält man nun für den Fall $x \approx \frac{3}{2}\pi + k \cdot 2\pi$ ein Fehler $\approx \epsilon_{masch}$.

Lösungsschritt

Ein Vergleich der Konditionszahlen ergibt nun für $f(x) = 1 + \sin x$:

$$cond_x = \frac{x \cdot f'(x)}{f(x)} = \frac{x \cdot \cos x}{1 + \sin x}$$

und für $g(x) = 1 + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)$:

$$cond_x = \frac{x \cdot g'(x)}{g(x)} = \frac{x \cdot cos(x)}{1 + 2 \cdot \sin\left(\frac{x}{2}\right) \cos\left(\frac{x}{2}\right)}$$

Hierbei ist ersichtlich, dass Verfahren 1 für den Fall $x \approx \frac{3}{2}\pi + k \cdot 2\pi$ schlecht konditioniert ist, Verfahren 2 jedoch überall gut konditioniert ist.

<u>zu d)</u>

Lösungsschritt

Es gilt:

$$\tilde{y} = (((x-1)(1+\epsilon_1))^2(1+\epsilon_2) - ((x+1)(1+\epsilon_3))^2)(1+\epsilon_4))(1+\epsilon_5)$$

Daraus ergibt sich für den relativen Fehler:

$$f_{rel} = \left| \frac{\tilde{y} - y}{x} \right| = \left| \frac{(2\epsilon_1 + \epsilon_2)(x-1)^2}{y} - \frac{(2\epsilon_3 + \epsilon_4)(x+1)^2}{y} + \epsilon_5 \right|$$

Hier ist offensichtlich für $y \approx 0$ mit großen relativen Fehler zu rechnen.

Tip

Ausmultiplizieren

Lösungsschritt

Einfaches Ausmultiplizieren ergibt:

$$f(x) = (x-1)^2 - (x+1)^2 = -4x$$

Daraus ergibt sich für den relativen Fehler:

$$f_{rel} = \left| \frac{(-4x)(1+\epsilon) - (-4x)}{-4x} \right| = |\epsilon|$$

Lösungsschritt

$$cond_x = \frac{x \cdot f'(x)}{f(x)} = \frac{x(-4)}{-4x} = 1$$

Daher haben wir hier ein gut konditioniertes Problem, welches mit der zweiten Rechenmethode auch mit kleinem Fehler gelöst werden kann.

zu e)

Lösungsschritt

Es gilt:

$$\tilde{y} = ((e^x - 1)(1+\epsilon_1) - x)(1+\epsilon_2) = e^x - 1 - x + e^x\epsilon_2 - \epsilon_2 - x\epsilon_2 - e^x\epsilon_1 - \epsilon_1$$

Daraus ergibt sich für den relativen Fehler:

$$f_{rel} = \left| \epsilon_2 - \frac{e^x + 1}{e^x - 1 - x}\epsilon_1 \right|$$

Hier ist für $x \approx 0$ mit großen relativen Fehler zu rechnen.

Tip

Reihentwicklung von e^x einsetzen.

Lösungsschritt

Ein stabiles Verfahren kann man über die Reihenentwicklung von e^x konstruieren:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \ldots \Rightarrow e^x - 1 - x = \frac{x^2}{2!} + \frac{x^3}{3!}$$

Bricht man diese Reihe nach n Elementen ab, so erhält man einen Fehler, der kleiner ist als $\frac{x^n}{n!}$.

Lösungsschritt

Die Konditionszahl der Funktion lautet:

$$cond_x = \frac{x \cdot (e^x - 1)}{e^x - 1 - x} = \frac{xe^x - x}{e^x - 1 - x}$$

Das Problem ist also für $x \approx 0$ schlecht konditioniert.

2 Lineare Gleichungssysteme

2.1 Aufgabe 5 (Buch, Übung: 39)

2.1.1 Aufgabenstellung

Gegeben ist ein lineares Gleichungssystem mit einer Tridiagonalmatrix

$$\begin{pmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ d_2 & a_2 & c_2 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & c_{n-1} \\ 0 & \cdots & \cdots & d_n & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{pmatrix}$$

a) Formulieren Sie das Verfahren der Gauß-Elimination ohne Pivotsuche in einem Pseudocode zur Lösung des obigen Gleichungssystems.

b) Wieviele Additionen, Multiplikationen, Divisionen benötigt dieser Algorithmus?

c) Man untersuche die Fragestellung von Teilaufgabe a) und b) für den Fall von Spalten- und Total-Pivotsuche.

2.1.2 Lösung

zu a)

Lösungsschritt

```
FOR i = 1,...,n-1
     factor = A[i+1,i] / A[i,i]
     A[i+1,i+1] = A[i+1,i+1] - factor * A[i,i+1]
     b[i+1] = b[i+1] - factor * b[i]
     A[i+1,i] = 0
ENDFOR
```

Durch das Ausnutzen der Tridiagonalstruktur kommt dieser Algorithmus ohne verschachtelte Schleifen aus.

zu b)

Lösungsschritt

Pro Eliminationsschritt ergeben sich folgende Operationen:
- n-1 Divisionen
- 2(n-1) Multiplikationen
- 2(n-1) Additionen

Dieser Algorithmus braucht größenordnungsmäßig also $\mathcal{O}(n)$ Operationen.

zu c)

Lösungsschritt

mit Spalten-Pivotsuche:

```
//------------------------------------
// Spalten-Pivotsuche
//------------------------------------
FOR i = 1,...,n-1
     pivot = |A[i,i]|
     IF( |A[i+1,i]| > pivot ) THEN
           //----------------------
           // Zeilentausch
           //----------------------
           FOR colIdx = i,...,i+3
                 pivot = A[i,colIdx]
                 A[i,colIdx] = A[i+1,colIdx]
```

```
                A[i+1,colIdx] = pivot
            ENDFOR
            pivot = b[i]
            b[i] = b[i+1]
            b[i+1] = pivot
        ENDIF
        factor = A[i+1,i] / A[i,i]
        A[i+1,i+1] = A[i+1,i+1] - factor * A[i,i+1]
        //-----------------------------
        // eliminieren der i+2-ten Spalte
        // (kann durch Zeilentausch aufgefüllt worden sein)
        //-----------------------------
        IF (i+2 <= n ) THEN
            A[i+1,i+2] = A[i+1,i+2] - factor * A[i,i+2]
        ENDIF
        b[i+1] = b[i+1] - factor * b[i]
ENDFOR
```

Die Anzahl der benötigten Operationen beläuft sich hierbei auf:

- n-1 Divisionen

- 2(n-1) + n-2 = 3n-4 Additionen

- 2(n-1) + n-2 = 3n-4 Multiplikationen

Dabei könnte man den Algorithmus noch dahingehend optimieren, dass man die n-2 Additionen/Multiplikationen nur ausführt, falls das Element an der Stelle A[i,i+2] != 0 ist. Damit würde die oben genannte Anzahl an Operationen nur im schlechtesten Fall ausgeführt werden. In jedem Fall kommt bei diesem Algorithmus der Aufwand für die Spalten-Pivotsuche hinzu.

Trotz der erhöhten Kosten die durch den Zeilentausch enstehen braucht dieser Algorithmus größenordnungsmäßig immer (nur) noch $\mathcal{O}(n)$ Operationen.

Lösungsschritt

mit Zeilen-Pivotsuche:

```
//---------------------------------
// Zeilen-Pivotsuche
//---------------------------------
FOR i = 1,...,n-1
    pivot = |A[i,i]|
```

```
    IF( |A[i,i+1]| > pivot ) THEN
          //----------------------
          // Spaltentausch
          //----------------------
          FOR rowIdx = i,...,i+3
                pivot = A[rowIdx,i]
                A[rowIdx,i] = A[rowIdx,i+1]
                A[rowIdx,i+1] = pivot
          ENDFOR
          pivot = x[i]
          x[i] = x[i+1]
          x[i+1] = pivot
    ENDIF
    factor = A[i+1,i] / A[i,i]
    A[i+1,i+1] = A[i+1,i+1] - factor * A[i,i+1]
    b[i+1] = b[i+1] - factor * b[i]
    //----------------------------
    // eliminieren der i+2-ten Zeile
    // (kann durch Spaltentausch aufgefüllt worden sein)
    //----------------------------
    IF (i+2 <= n ) THEN
          factor = - A[i+1] / A[i,i]
          A[i+2,i+1] = factor * A[i,i+1]
          b[i+1] = b[i+1] - factor * b[i]
    ENDIF
ENDFOR
```

Die Anzahl der benötigten Operationen beläuft sich hierbei auf:

- n-1 + n-2 = 2n-3 Divisionen
- 2(n-1) + n-2 = 3n-4 Additionen
- 2(n-1) + 2(n-2) = 4n-6 Multiplikationen

Die Kosten erhöhen sich gegenüber dem Gauß-Algorithmus mit Spalten-Pivotsuche (5n-5) auf 9n-13 Operationen.

Trotz der erhöhten Kosten die durch den Spaltentausch enstehen braucht dieser Algorithmus größenordnungsmäßig immer (nur) noch $\mathcal{O}(n)$ Operationen.

2.2 Aufgabe 6 (Buch, Übung: 40)

2.2.1 Aufgabenstellung

Gegeben ist ein lineares Gleichungssystem der Form

$$\begin{pmatrix} a_{1,1} & a_{1,2} & 0 & \cdots & 0 & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & 0 & \cdots & 0 \\ 0 & a_{3,2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & 0 & \cdots & 0 & a_{n,n-1} & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

a) Beschreiben Sie, wie sich die Besetztheitsstruktur der Matrix ändert, wenn Sie Gauß-Elimination ohne Pivotsuche durchführen.

b) Geben Sie ein Verfahren an, die Einträge der Matrix so zu speichern, dass im Verlauf der Gauß-Elimination ohne Pivotsuche keine überflüssigen Nullen gespeichert werden.

c) Formulieren Sie das Verfahren der Gauß-Elimination ohne Pivotsuche in einem Pseudocode **zur Lösung** des obigen Gleichungssystems. Benutzen Sie dabei die in b) definierte Speicherung der Matrix. Schätzen Sie größenordnungsmäßig die Anzahl der anfallenden Rechenoperationen ab.

d) Wie kann sich die Struktur der Matrix ändern, wenn Sie Spalten-Pivotsuche durchführen?

2.2.2 Lösung

zu a)

Lösungsschritt

Beim Eliminieren von $a_{2,1}$ ergibt sich ein neuer Eintrag in der letzten Spalte bei $a_{2,n}$. Beim Eliminieren von $a_{n,1}$ ergibt sich ein neuer Eintrag bei $a_{n,2}$. Im Verlauf der Gauß-Elimination wandert also in der untersten Zeile ein Eintrag nach rechts, und die rechteste Spalte wird schrittweise aufgefüllt.

zu b)

Lösungsschritt

Wir speichern die drei Diagonalen in drei Vektoren, nämlich $d_i = a_{i,i}$ für $i = 1 \dots n, c_i = a_{i+1,i}$ für $i = 1 \dots n-1$ und $e_i = a_{i,i+1}$ für $i = 1 \dots n-1$.Da wir auf jeden Fall zur Bestimmung der Lösung am Schluss das Dreiecksgleichungssystem auflösen müssen, benötigen wir alle Einträge in der letzten Spalte, die wir daher ebenfalls in einem eigenen Vektor $f_i = a_{i,n}$ für $i = 1 \dots n-2$ speichern ($i = n-1$ und $j = n$ liegt im Vektor e und

d). Die zusätzlichen Einträge in der letzten Zeile benötigen wir später nicht mehr (es sei denn, wir wollen eine LU-Zerlegung berechnen). Daher können wir dieses jeweils einzelne Element in einer Zahl α speichern, die jeweils dieses linke untere Element im jeweiligen Eliminationsschritt enthält. (Man könnte sogar den Vektor f einsparen, und die jeweils zu Null gemachten Stellen in c durch die entstehenden Einträge der letzten Spalte nutzen).

zu c)

Lösungsschritt

```
FOR i = 1,...,n-1
     factor = c[i] / d[i]
     d[i+1] = d[i+1] - factor * e[i]
     b[i+1] = b[i+1] - factor * b[i]
     IF( i < n-1 ) THEN
           factor = α / d[i]
           d[n] = d[n] - factor * f[i]
           b[n] = b[n] - factor * b[i]
     ENDIF
     IF( i < n-2 ) THEN
           f[i+1] = -f[i] * (c[i] / d[i])
           α = -e[i] * (α / d[i])
     ENDIF
     IF( i = n-2 ) THEN
           e[n-1] = e[n-1] - f[i] * (c[i] / d[i])
           c[n-1] = c[n-1] - e[i] * (α / d[i])
     ENDIF
ENDFOR
```

Danach noch Auflösen des Dreiecksgleichungssystems (bidiagonal + letzte Spalte):

```
x[n] = b[n] / d[n]
x[n-1] = (b[n-1] - e[n-1] * x[n]) / d[n-1]
FOR i = n-2,...,1
     x[i] = (b[i] - f[i] * x[n] - e[i] * x[i+1]) / d[i]
ENDFOR
```

Die Kosten liegen im Bereich $\mathcal{O}(n)$ da nur einfache Schleifen der Länge n durchlaufen werden.

zu d)

Lösungsschritt

Vernachlässigen wir zunächst die einzelnen Einträge links unten und rechts oben, so kann es durch Vertauschen zweier Nachbarzeilen passieren, dass oben eine zusätzlich Diagonale auftritt $(a_{1,3}, \ldots, a_{n-2,n})$. Dazu kommen auf jeden Fall die letzte Spalte und letzte Zeile wie unter a). Tritt der Fall auf, dass die jeweilige Zeile mit der untersten Zeile vertauscht wird, so werden dadurch in der vorletzten Spalte zusätzliche Einträge erzeugt $(a_{1,n-1}, \ldots, a_{n-3,n-1})$.

2.3 Aufgabe 7 (Buch, Übung: 41)

2.3.1 Aufgabenstellung

Gegeben ist das lineare Gleichungssystem in Hessenberg-Form

$$\begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & \cdots & h_{1,n} \\ h_{2,1} & h_{2,2} & \cdots & \cdots & h_{2,n} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{n,n-1} & h_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

a) Formulieren Sie in einem Pseudocode das Verfahren der Gauß-Elimination ohne Pivotsuche unter Ausnutzung der Struktur der gegebenen Matrix.

b) Wie viele Operationen benötigen Sie größenordnungsmäßig?

c) Womit ist zu rechnen, wenn Sie Spalten- oder Zeilen-Pivotsuche zulassen?

2.3.2 Lösung

zu a)

Lösungsschritt

```
FOR i=1,...,n-1
     factor = A[i+1,i] / A[i,i]
     A[i+1,i] = 0
     FOR j=i+1,...,n
           A[i+1,j] = A[i+1,j] - factor * A[i,j]
     ENDFOR
     b[i+1] = b[i+1] - factor * b[i]
ENDFOR
```

Durch die gegebene Struktur der Matrix A muss nur eine Zeile pro Eliminationsschritt bearbeitet werden.

zu b)

Lösungsschritt

Es fallen folgenden Operationen an:
- n-1 Divisionen,

$$\sum_{i=1}^{n-1} i = \frac{(n-1)(n-2)}{2} = \frac{n^2}{2} - \frac{3n}{2} + 1 \quad \text{Additionen und Multiplikationen}$$

für die Elimination des Elements an der Stelle `A[i+1,i]` und noch einmal n-1 Additionen und Multiplikationen für das Element `b[i+1]`.
Da hier der Faktor $\frac{n^2}{2}$ überwiegt, liegt die Anzahl der Operationen größenordnungsmäßig bei $\mathcal{O}(n^2)$.

zu c)

Lösungsschritt

Für die Spalten-Pivotsuche ändert sich nichts, da hier die Zeile `i` nur mit der Zeile `i+1` vertauscht werden kann, da alle anderen Spalteneinträge der Spalte `i` (beginnend in Zeile `i`) gleich Null sind. Daher werden dann nur Zeilen gleicher Struktur miteinander vertauscht.

Bei der Zeilen-Pivotsuche kann es nun aber passieren, dass stark besetzte Spalten mit schwach besetzten Spalten vertauscht werden. Dadurch kann es passieren, dass die "fast-schon"-Dreickecksmatrix zerstört wird und dieses durch mehrfaches Subtrahieren einer Zeile zu anderen Zeilen in einem Eliminationsschritt wieder ausgeglichen werden muss. Dabei entsteht dann wieder der "standard" Gauß-Algorithmus mit einer Operationenanzahl in der Größenordnung von $\mathcal{O}(n^3)$.

2.4 Aufgabe 8 (Buch, Übung: 46)

2.4.1 Aufgabenstellung

Bestimmen Sie die effiziente Speicherung und effiziente Varianten der Gauß-Elimination bei linearen Gleichungssystemen mit folgender Gestalt:

$$A_X = \begin{pmatrix} * & & & & * \\ & * & & * & \\ & & * & & \\ & * & & * & \\ * & & & & * \end{pmatrix}$$

2.4.2 Lösung

Lösungsschritt

Zur Speicherung werden drei Vektoren benötigt. Die Diagonaleinträge werden in dem Vektor `Diag` gespeichert (`Diag[i] =` $a_{i,i}$ mit $i = 1, ..., n$). Die Einträge unterhalb der Diagonalen werden in dem Vektor `LowDiag` gespeichert (`LowDiag[i] =` $a_{n-i+1,i}$ mit $i = 1, ..., n/2$ (wobei $n/2$ abgerundet). Analog werden die Einträge oberhalb der Diagonalen in dem Vektor `UppDiag` gespeichert (`UppDiag[i] =` $a_{i,n-i+1}$ mit $i = 1, ..., n/2$ (wobei $n/2$ abgerundet).

Lösungsschritt

```
FOR i=1,...,n/2
     factor = LowDiag[i] / Diag[i]
     Diag[n-i+1] = Diag[n-i+1] - factor * UppDiag[i]
     b[n-i+1] = b[n-i+1] - factor * b[i]
ENDFOR
```

2.5 Aufgabe 9 (Buch, Übung: 46)

2.5.1 Aufgabenstellung

Bestimmen Sie die effiziente Speicherung und effiziente Varianten der Gauß-Elimination bei linearen Gleichungssystemen mit folgender Gestalt:

$$A_N = \begin{pmatrix} * & & & & * \\ * & * & & & * \\ * & & * & & * \\ * & & & * & * \\ * & & & & * \end{pmatrix}$$

2.5.2 Lösung

Lösungsschritt

Zur Speicherung werden drei Vektoren benötigt. Die Diagonaleinträge werden in dem Vektor `Diag` gespeichert (`Diag[i] =` $a_{i,i}$ mit $i = 1, ..., n$). Die Einträge der linken Spalte ($a_{1,1}$ ausgenommen) werden in dem Vektor `LeftCol` gespeichert (`LeftCol[i] =` $a_{i+1,1}$ mit $i = 1, ..., n-1$). Die Einträge der rechten Spalte ($a_{n,n}$ ausgenommen) werden in dem Vektor `RightCol` gespeichert (`RightCol[i] =` $a_{i+1,n}$ mit $i = 1, ..., n-1$).

Lösungsschritt

```
FOR i=1,...,n-1
     factor = LeftCol[i] / Diag[i]
     IF( i+1 < n ) THEN
           RightCol[i+1] = RightCol[i+1] - factor * RightCol[i]
     ELSE
           Diag[i+1] = Diag[i+1] - factor * RightCol[i]
     ENDIF
     b[i+1] = b[i+1] - factor * b[1]
ENDFOR
```

2.6 Aufgabe 10 (Buch, Übung: 46)

2.6.1 Aufgabenstellung

Bestimmen Sie die effiziente Speicherung und effiziente Varianten der Gauß-Elimination bei linearen Gleichungssystemen mit folgender Gestalt:

$$A_Z = \begin{pmatrix} * & * & * & * & * \\ & & & * & \\ & & * & & \\ & * & & & \\ * & * & * & * & * \end{pmatrix}$$

2.6.2 Lösung

Tip

Umwandeln der Matrix durch Zeilen und Spaltentausch.

Lösungsschritt

Zur Speicherung werden drei Vektoren benötigt. Die Diagonaleinträge werden in dem Vektor `Diag` gespeichert (`Diag[i] =` $a_{i,n-i+1}$ mit $i = 1, ..., n$). Die obere Zeile wird in dem Vektor `UppRow` gespeichert (`UppRow[i] =` $a_{1,i}$ mit $i = 1, ...n$). Die untere Zeile wird in dem Vektor `LowRow` gespeichert (`LowRow[i] =` $a_{n,i}$ mit $i = 1, ..., n$). Hierbei erleichtert es den späteren Algorithmus, dass die Einträge $a_{1,n}$ und $a_{n,1}$ doppelt abgespeichert werden.

Um nun mit möglichst wenig Operationen (Additionen, Divisionen etc.) das Gleichungssystem zu lösen, wird die Matrix wie folgt umgewandelt:
- Umkehren der Diagonalen

$$A_Z = \begin{pmatrix} * & * & * & * & * \\ & & & * & \\ & & * & & \\ & * & & & \\ * & * & * & * & * \end{pmatrix} \rightarrow \begin{pmatrix} * & * & * & * & * \\ & * & & & \\ & & * & & \\ & & & * & \\ * & * & * & * & * \end{pmatrix}$$

- Tauschen der 2. und letzten Zeile

$$\begin{pmatrix} * & * & * & * & * \\ & * & & & \\ & & * & & \\ & & & * & \\ * & * & * & * & * \end{pmatrix} \rightarrow \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ & & * & & \\ & & & * & \\ & * & & & 0 \end{pmatrix}$$

- Tauschen der 2. und letzten Spalte

$$\begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ & & * & & \\ & & & * & \\ & * & & & 0 \end{pmatrix} \rightarrow \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ & & * & & \\ & & & * & \\ & & & & * \end{pmatrix}$$

Nun muss man nur noch ein Vielfaches der ersten Zeilen von der Zweiten Zeile abziehen und es werden dadurch keine Nullen in der Matrix verstört.

Lösungsschritt

```
//--------------------------------
// Diagonale umkehren
//--------------------------------
FOR i=2,...,n/2
     val = Diag[i]
     Diag[i] = Diag[n-i+1]
     Diag[n-i+1] = val
ENDFOR
//--------------------------------
// Tausch Zeile 2 mit Zeile n
//--------------------------------
val = Diag[2]
Diag[2] = LowRow[2]
//--------------------------------
// Tausch Spalte 2 mit Spalte n
//--------------------------------
Diag[n] = val
val = Diag[2]
Diag[2] = LowRow[n]
LowRow[n] = val
val = UppRow[2]
UppRow[2] = UppRow[n]
UppRow[n] = val
//--------------------------------
// Gauß
//--------------------------------
factor = LowRow[1] / Diag[1]
FOR i=2,...,n
     LowRow[i] = LowRow[i] - factor * UppRow[i]
ENDFOR
//--------------------------------
```

```
// Letztes Setzten der Werte
//-------------------------------
Diag[2] = LowRow[2]
LowRow[1] = 0
```

3 Interpolation

3.1 Aufgabe 11 (Buch, Übung: 57)

3.1.1 Aufgabenstellung

Bestimmen Sie die Anzahl der Operationen zur Berechnung des interpolierenden Polynoms $p_0, \ldots, p_n(c)$ an der Stelle c mittels

a) Lagrange-Polynom

b) Neville-Schema

c) Newton-Form des interpolierenden Polynoms

3.1.2 Lösung

zu a)

Tip

Aufstellen eines Pseudo-Codes für die Lagrange-Polynome:

$$p(x) = \sum_{j=0}^{n} y_j L_j(x) \qquad \text{mit } L_j(x) = \prod_{i=0; i \neq j}^{n} \frac{x - x_i}{x_j - x_i}$$

Lösungsschritt

Ein Pseudo-Code zur Berechnung des Polynoms an der Stelle c mittels Lagrange-Polynomen lautet:

```
FOR j=0,...,n
     L = 1
     FOR i=0,...,n
           IF( i != j ) THEN
                 L = L * ( c - x[i] ) / ( x[j] - x[i] )
           ENDIF
     ENDFOR
     p = p + y[j] * L
ENDFOR
```

Lösungsschritt

Durch die ineinander geschachtelte FOR-Schleife benötigen wir hier $\mathcal{O}(n^2)$ Operationen.

zu b)

Lösungsschritt

Ein Pseudo-Code zur Berechnung des Polynoms an der Stelle c mittels des Neville-Schemas lautet:

```
FOR i=0,...,n
     p[i] = y[i]
     FOR k=i-1,...,0
           p[k] = p[k+1] + ( p[k+1] - p[k] ) * ( c - x[i] ) / x[i] - x[k]
     ENDFOR
ENDFOR
```

Lösungsschritt

Durch die ineinander geschachtelte FOR-Schleife benötigen wir hier $\mathcal{O}(n^2)$ Operationen.

zu c)

Lösungsschritt

Ein Pseudo-Code zur Berechnung des Polynoms an der Stelle c mittels der Newton-Form des Interpolationspolynoms lautet:

```
//------------------------------------------------------
// Berechnung der dividierten Differenzen:
//------------------------------------------------------

FOR i=0,...,n
     p[i] = y[i]
     FOR k=i-1,...,0
           p[k] = p[k+1] + ( p[k+1] - p[k] ) / x[i] - x[k]
     ENDFOR
ENDFOR

//------------------------------------------------------
```

```
// Berechnung des Polynoms:
//-----------------------------------------------------

p = f[n]
FOR j=n-1,...,0
     p = p * ( c - x[j] ) + f[j]
ENDFOR
```

Lösungsschritt

Durch die ineinander geschachtelte FOR-Schleife benötigen wir hier $\mathcal{O}(n^2)$ Operationen.

3.2 Aufgabe 12 (Buch, Übung: 58)

3.2.1 Aufgabenstellung

Wir betrachten die Funktion $f(x) = x^8$. Bestimmen Sie zu f das interpolierende Polynom vom Grad kleiner gleich 6 zu den Stützstellen -3,-2,-1,0,1,2,3.

3.2.2 Lösung

Tip

In der Lösung wir der Ansatz über die Newton-Form des Interpolierenden Polynoms verwendet.

Lösungsschritt

Die Stützstellen zu der Funktion $f(x) = x^8$ lauten:

$$\begin{aligned}
x_0 &= (-3, 6561) \\
x_1 &= (-2, 256) \\
x_2 &= (-1, 1) \\
x_3 &= (0, 0) \\
x_4 &= (1, 1) \\
x_5 &= (2, 256) \\
x_6 &= (3, 6561)
\end{aligned}$$

Unser gesuchtes Polynom (Newton-Form des Interpolierenden Polynoms) hat die Form (nach dem Horner-Schema):

$$\begin{aligned} p(x) &= (((((f[x_0,...,x_6](x-x_5)+f[x_0,...,x_5])(x-x_4)+ \\ &\quad +f[x_0,...x_4])(x-x_3)+f[x_0,...,x_3])(x-x_2)+ \\ &\quad +f[x_0,x_1,x_2])(x-x_1)+f[x_0,x_1])(x-x_0)+f[x_0] \end{aligned}$$

Wir müssen nun also die Koeffizienten (dividierte Differenzen) $f[x_0,...,x_j], j=1,...,6$ berechnen.

Lösungsschritt

Die Koeffizienten (mit $*$ gekennzeichnet) und die Zwischenergebnisse lauten:

$$\begin{array}{llllllll}
f[x_0] & = & y_0 & = & & = & 6561 & (*) \\
f[x_0,x_1] & = & \frac{y_1-y_0}{x_1-x_0} & = & \frac{256-6561}{-2+3} & = & -6305 & (*) \\
f[x_1,x_2] & = & \frac{y_2-y_1}{x_2-x_1} & = & \frac{1-256}{-1+2} & = & -255 & \\
f[x_2,x_3] & = & \dots & = & \dots & = & -1 & \\
f[x_3,x_4] & = & \dots & = & \dots & = & 1 & \\
f[x_4,x_5] & = & \dots & = & \dots & = & 255 & \\
f[x_5,x_6] & = & \dots & = & \dots & = & 6305 & \\
f[x_0,x_1,x_2] & = & \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0} & = & \frac{-255+6305}{-1+3} & = & 3025 & (*) \\
f[x_1,x_2,x_3] & = & \dots & = & \dots & = & 127 & \\
f[x_2,x_3,x_4] & = & \dots & = & \dots & = & 1 & \\
f[x_3,x_4,x_5] & = & \dots & = & \dots & = & 127 & \\
f[x_4,x_5,x_6] & = & \dots & = & \dots & = & 3025 & \\
f[x_0,x_1,x_2,x_3] & = & \dots & = & \dots & = & -966 & (*) \\
f[x_1,x_2,x_3,x_4] & = & \dots & = & \dots & = & -42 & \\
f[x_2,x_3,x_4,x_5] & = & \dots & = & \dots & = & 42 & \\
f[x_3,x_4,x_5,x_6] & = & \dots & = & \dots & = & 966 & \\
f[x_0,x_1,x_2,x_3,x_4] & = & \dots & = & \dots & = & 231 & (*) \\
f[x_1,x_2,x_3,x_4,x_5] & = & \dots & = & \dots & = & 21 & \\
f[x_2,x_3,x_4,x_5,x_6] & = & \dots & = & \dots & = & 231 & \\
f[x_0,x_1,x_2,x_3,x_4,x_5] & = & \dots & = & \dots & = & -42 & (*) \\
f[x_1,x_2,x_3,x_4,x_5,x_6] & = & \dots & = & \dots & = & 42 & \\
f[x_0,x_1,x_2,x_3,x_4,x_5,x_6] & = & \dots & = & \dots & = & 14 & (*)
\end{array}$$

Lösungsschritt

Einsetzen der Koeffizienten ergibt nun:

$$\begin{aligned} p(x) &= (((((14\cdot(x-x_5)-42)(x-x_4)+ \\ &\quad +231)(x-x_3)-966)(x-x_2)+ \\ &\quad +3025)(x-x_1)-6305)(x-x_0)+6561 \end{aligned}$$

Zuletzt Einsetzen der Stützstellen:

$$\begin{aligned} p(x) &= (((((14 \cdot (x-2) - 42)(x-1) + \\ &\quad +231)x - 966)(x+1) + \\ &\quad +3025)(x+2) - 6305)(x+3) + 6561 \end{aligned}$$

Unser gesuchtes Polynom lautet also:

$$p(x) = 14x^6 - 49x^4 + 36x^2$$

3.3 Aufgabe 13 (Buch, Übung: 59)

3.3.1 Aufgabenstellung

Approximieren Sie die Funktion $ld(x) = \log_2(x)$.

a) Berechnen Sie das Interpolationspolynom zu $ld(x)$ mit den Stützstellen 16,32 und 64.

b) Bestimmen Sie die linearen Ausgleichsgrade zu diesem Stützstellen und und vergleichen Sie sie mit dem Interpolationspolynom aus a).

3.3.2 Lösung

zu a)

Tip

In der Musterlösung wurde der Ansatz mittels Lagrange-Polynomen gewählt.

Lösungsschritt

Die Stützstellen zu der Funktion $f(x) = \log_2(x)$ lauten:

$$\begin{aligned} x_0 &= (16, 4) \\ x_1 &= (32, 5) \\ x_2 &= (64, 6) \end{aligned}$$

Unser gesuchtes Polynom (Lagrange-Polynome) hat die Form:

$$p(x) = \sum_{j=0}^{2} y_j L_j(x) \qquad \text{mit } L_j(x) = \prod_{i=0;i\neq j}^{2} \frac{x - x_i}{x_j - x_i}$$

Lösungsschritt

Die Lagrange Polynome lauten:

$$\begin{aligned}
L_0(x) &= \tfrac{(x-32)(x-64)}{(16-32)(16-64)} &= \tfrac{x^2-96x+2048}{768} \\
L_1(x) &= \tfrac{(x-16)(x-64)}{(32-16)(32-64)} &= -\tfrac{x^2-80x+1024}{512} \\
L_2(x) &= \tfrac{(x-16)(x-32)}{(64-16)(64-32)} &= \tfrac{x^2-48x+512}{1536}
\end{aligned}$$

Lösungsschritt

Einsetzen in unser Polynom ergibt:

$$\begin{aligned}
p(x) &= 4 \cdot L_0(x) + 5 \cdot L_1(x) + 6 \cdot L_2(x) \\
&= -\frac{1}{1536}x^2 + \frac{3}{32}x + 2\frac{2}{3}
\end{aligned}$$

<u>zu b)</u>

Tip

Die lineare Ausgleichsgerade $g(x) = a \cdot x + b$ setzt sich aus der linken und der rechten Stützstelle zusammen.

Lösungsschritt

Die lineare Ausgleichsgerade $g(x) = a \cdot x + b$ zu den gegeben Stützstellen errechnet sich wie folgt:

$$\begin{aligned}
g(x) &= \frac{f(x_2) - f(x_0)}{x_2 - x_0} \cdot (x - x_0) + f(x_0) \\
&= \frac{1}{24}x - \frac{2}{3} + 4 \\
&= \frac{1}{24}x + 3\frac{1}{3}
\end{aligned}$$

Lösungsschritt

Für unsere Stützstelle x_1 erhält man hier nun:

$$g(32) = \frac{1}{24} \cdot 32 + 3\frac{1}{3} = 4\frac{2}{3}$$

3.4 Aufgabe 14 (Buch, Übung: 63)

3.4.1 Aufgabenstellung

Zu den Stützstellen $x_0 < x_1 < x_2$ mit bekannten Funktionswerten $f(x_0), f(x_1)$ und $f(x_2)$ soll das Integral $\int_a^b f(x)dx$ näherungsweise bestimmt werden; dabei sei $a \leq x_0 < x_2 \leq b$. Beschreiben Sie, wie man mittels eines Polynoms zweiten Grades dieses Integral näherungsweise berechnen kann und geben Sie eine explizite Quadraturregel der Form

$$\int_a^b f(x)dx \approx \sum_{j=0}^{2} w_j f(x_j).$$

3.4.2 Lösung

Tip

Gewichtung der Stützstellen mittels Lagrange-Polynomen.

Lösungsschritt

Wir suchen hier ein Polynom $p(x)$ zweiten Grades für das gilt:

$$\int_a^b f(x)dx \approx \int_a^b p(x)dx = \sum_{i=0}^{2} w_j f(x_i)$$

Da wir die Werte für $f(x_i), i = 0, 1, 2$ bereits geben haben, brauchen wir also nur die Gewichte für unsere Stützstellen zu suchen, z.B. mittels Lagrange-Polynome.

Lösungsschritt

Durch Einsetzten der Lagrange-Polynome erhalten wir nun unsere Quadraturregel

$$\int_a^b f(x)dx \approx \sum_{j=0}^{2} w_j f(x_j) = \sum_{i=0}^{2} f(x_i) \int_a^b L_i(x)dx$$

wobei gilt:

$$\begin{aligned}
L_0(x) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \\
L_1(x) &= \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \\
L_2(x) &= \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}
\end{aligned}$$

3.5 Aufgabe 15 (Buch, Übung: 65)

3.5.1 Aufgabenstellung

Gegeben sei eine Integrationsregel mit nur einer Stützstelle x_0 und einem Gewicht α:

$$\alpha \cdot f(x_0) \approx \int_a^b f(x)\, dx.$$

Bestimmen Sie x_0 und α so, dass die Integrationsregel für Polynome möglichst hohen Grades exakt gilt.

3.5.2 Lösung

Tip

$f(x)$ als Polynom 0-ten bzw. 1-ten Grades betrachten.

Lösungsschritt

Wir betrachten:

i) $f(x) = x^0 = 1$

ii) $f(x) = x^1$

Lösungsschritt

Daraus folgt

$$\int_a^b 1\, dx = b - a \Leftrightarrow \alpha \cdot 1 = b - a \Leftrightarrow \alpha = b - a$$

Lösungsschritt

sowie

$$\int_a^b x\, dx = \frac{b^2 - a^2}{2} \Leftrightarrow \alpha \cdot x_0 = \frac{b^2 - a^2}{2} \Leftrightarrow x_0 = \frac{b + a}{2}$$

3.6 Aufgabe 16 (Buch, Übung: 66)

3.6.1 Aufgabenstellung

Bestimmen Sie zu den Stützstellen $x_0 = 0$, $x_1 = a$, $x_2 = 1$ Gewichte w_0, w_1, w_2 und a so, dass die Integrationsregel

$$\int_0^1 f(x)\, dx = w_0 f(x_0) + w_1 f(a) + w_2 f(1)$$

für Polynome möglichst hohen Grades exakt ist.

3.6.2 Lösung

Tip

Betrachte $f(x)$ als Polynom vom Grad $n = 0, 1, 2, 3$.

Lösungsschritt

Wir betrachten:

i) $f(x) = x^0 = 1$

ii) $f(x) = x^1$

iii) $f(x) = x^2$

iv) $f(x) = x^3$

Daraus folgt:

$$
\begin{aligned}
i) \int_0^1 1\ dx = 1 = w_0 + w_1 + w_2 \\
ii) \int_a^b x\ dx = \frac{1}{2} = w_1 a + w_2 \\
iii) \int_a^b x^2\ dx = \frac{1}{3} = w_1 a^2 + w_2 \\
iv) \int_a^b x^3\ dx = \frac{1}{4} = w_1 a^3 + w_2
\end{aligned}
$$

Lösungsschritt

nun $ii) - iii)$:

$$
\begin{aligned}
\frac{1}{6} &= w_1 a - w_1 a^2 = w_1(a - a^2) \\
\Leftrightarrow w_1 &= \frac{1}{6(a - a^2)}
\end{aligned}
$$

und $iii) - iv)$:

$$
\begin{aligned}
\frac{1}{12} &= w_1(a^2 - a^3) = \frac{a(a - a^2)}{6(a - a^2)} = \frac{a}{6} \\
\Leftrightarrow a &= \frac{1}{2}
\end{aligned}
$$

Lösungsschritt

Einsetzen in $i)$ und $ii)$ ergibt dann:

$$w_0 = \frac{1}{6},\ w_1 = \frac{2}{3}\ w_2 = \frac{1}{6}\ a = \frac{1}{2}$$

4 Fourier-Transformation

4.1 Aufgabe 17 (Buch, Übung: 75)

4.1.1 Aufgabenstellung

Zeigen Sie für $j, k \in \mathbb{N}$:

$$\begin{aligned} i) \int_{-\pi}^{\pi} \cos(jx)\cos(kx)\;dx &= 0 \text{ für } j \neq k \\ ii) \int_{-\pi}^{\pi} \sin(jx)\sin(kx)\;dx &= 0 \text{ für } j \neq k \\ iii) \int_{-\pi}^{\pi} \cos(jx)\sin(kx)\;dx &= 0 \forall j, k \end{aligned}$$

4.1.2 Lösung

<u>zu i)</u>

Tip

$2\cos\alpha\cos\beta = \cos(\alpha - \beta) + \cos(\alpha + \beta)$

Lösungsschritt

$$\int_{-\pi}^{\pi} \cos(jx) \cdot \cos(kx)\;dx = \frac{1}{2}\int_{-\pi}^{\pi} 2\cos(jx) \cdot \cos(kx)\;dx =$$

$$= \frac{1}{2}\int_{-\pi}^{\pi} \cos((j-k)x) + \cos((j+k)x)\;dx = \frac{1}{2}\int_{-\pi}^{\pi} \cos((j-k)x)\;dx + \frac{1}{2}\int_{-\pi}^{\pi} \cos((j+k)x)\;dx =$$

Lösungsschritt

Für $j = k$ gilt:

$$\frac{1}{2}\int_{-\pi}^{\pi} \cos(0 \cdot x)\;dx \;+\; \frac{1}{2}\int_{-\pi}^{\pi} \cos(2k \cdot x)\;dx = \pi$$

Für $j \neq k$ gilt:

$$\frac{1}{2}\left|\sin((j-k)x)\right|_{-\pi}^{\pi} + \frac{1}{2}\left|\sin((j+k)x)\right|_{-\pi}^{\pi} = 0 - 0 = 0$$

zu ii)

Tip

$2\sin\alpha\sin\beta = \cos(\alpha-\beta) - \cos(\alpha+\beta)$

Lösungsschritt

$$\int_{-\pi}^{\pi} \sin(jx)\cdot\sin(kx)\,dx = \frac{1}{2}\int_{-\pi}^{\pi} 2\sin(jx)\cdot\sin(kx)\,dx =$$

$$= \frac{1}{2}\int_{-\pi}^{\pi} \cos((j{-}k)x) - \cos((j{+}k)x)\,dx = \frac{1}{2}\int_{-\pi}^{\pi}\cos((j{-}k)x)\,dx - \frac{1}{2}\int_{-\pi}^{\pi}\cos((j{+}k)x)\,dx =$$

Lösungsschritt

Für $j = k$ gilt:

$$\frac{1}{2}\int_{-\pi}^{\pi}\cos(0\cdot x)\,dx \;-\; \frac{1}{2}\int_{-\pi}^{\pi}\cos(2k\cdot x)\,dx = \pi$$

Für $j \neq k$ gilt:

$$\frac{1}{2}\,|\sin((j-k)x)|_{-\pi}^{\pi} - \frac{1}{2}\,|\sin((j+k)x)|_{-\pi}^{\pi} = 0 - 0 = 0$$

zu iii)

Tip

$2\sin\alpha\cos\beta = \sin(\alpha-\beta) + \sin(\alpha+\beta)$

Lösungsschritt

$$\int_{-\pi}^{\pi}\cos(jx)\cdot\sin(kx)\,dx = \frac{1}{2}\int_{-\pi}^{\pi} 2\cos(jx)\cdot\sin(kx)\,dx =$$

$$= \frac{1}{2}\int_{-\pi}^{\pi}\sin((j{-}k)x) + \sin((j{+}k)x)\,dx = \frac{1}{2}\int_{-\pi}^{\pi}\sin((j{-}k)x)\,dx + \frac{1}{2}\int_{-\pi}^{\pi}\sin((j{+}k)x)\,dx =$$

$$= \frac{1}{2}\,|-\cos((j-k)x)|_{-\pi}^{\pi} + \frac{1}{2}\,|-\cos((j+k)x)|_{-\pi}^{\pi}$$

Lösungsschritt

Da $\cos(-\varphi) = \cos(\varphi)$ gilt $\forall j,k \in \mathbb{N}$:

$$\frac{1}{2}(-\cos((j-k)\pi) + \cos((j-k)\pi) - \cos((j+k)\pi) + \cos((j+k)\pi)) = 0$$

4.2 Aufgabe 18 (Buch, Übung: 76)

4.2.1 Aufgabenstellung

Sei f eine stückweise stetige periodische Funktion mit Fourier-Reihe

$$f(x) = \frac{a_0}{2} + \sum (a_k \cos(kx) \ + \ b_k \sin(kx)).$$

Zeigen Sie, dass die Fourier-Koeffizienten gegeben sind durch

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx)\ dx \qquad \text{und} \qquad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx)\ dx$$

4.2.2 Lösung

Tip

$$\int_{-\pi}^{\pi} \cos(jx) \cos(kx)\ dx = \pi \text{ für } j = \pm k \text{ und } \int_{-\pi}^{\pi} \cos(jx) \sin(kx)\ dx = 0\ \forall j, k \in \mathbb{N}$$

Lösungsschritt

Beweis von a_k:
Einsetzen ergibt:

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \left[\frac{a_0}{2} + \sum (a_k \cos(kx) \ + \ b_k \sin(kx)) \right] \cdot \cos(jx)\ dx =$$

$$= \frac{1}{\pi} \left[\underbrace{\int_{-\pi}^{\pi} \frac{a_0}{2} \cdot \cos(jx)\ dx}_{=0} + \sum a_k \underbrace{\int_{-\pi}^{\pi} \cos(kx) \cdot \cos(jx)\ dx}_{=\pi} + \sum b_k \underbrace{\int_{-\pi}^{\pi} \sin(kx) \cdot \cos(jx)\ dx}_{=0} \right] =$$

$$= \frac{1}{\pi} a_k \cdot \pi = a_k$$

Tip

$$\int_{-\pi}^{\pi} \sin(jx) \sin(kx)\ dx = \pi \text{ für } j = \pm k \text{ und } \int_{-\pi}^{\pi} \cos(jx) \sin(kx)\ dx = 0\ \forall j, k \in \mathbb{N}$$

Lösungsschritt

Beweis von b_k:
Einsetzen ergibt:

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\left[\frac{a_0}{2}+\sum(a_k\cos(kx)\ +\ b_k\sin(kx))\right]\cdot\sin(jx)\,dx=$$

$$=\frac{1}{\pi}\left[\underbrace{\int_{-\pi}^{\pi}\frac{a_0}{2}\cdot\sin(jx)\,dx}_{=0}\ +\ \sum a_k\underbrace{\int_{-\pi}^{\pi}\cos(kx)\cdot\sin(jx)\,dx}_{=0}\ +\ \sum b_k\underbrace{\int_{-\pi}^{\pi}\sin(kx)\cdot\sin(jx)\,dx}_{=\pi}\right]=$$

$$=\frac{1}{\pi}b_k\cdot\pi=b_k$$

4.3 Aufgabe 19 (Buch, Übung: 79)

4.3.1 Aufgabenstellung

Sei $n=3m$, also durch drei teilbar. Beschreiben Sie, wie

$$w_j=\sum c_k e^{2\pi ijk/n},\qquad j=0,1,...,n-1$$

in drei Summen der Länge m aufgespalten werden kann. Wie kann man mit dieser Aufspaltung die Berechnung der IDFT eines Vektors der Länge n auf die Berechnung von mehreren IDFT's kürzerer Vektoren zurückgeführt werden?

4.3.2 Lösung

Tip

Zur Lösung für $j=0,...,m-1$:

- Summe von Bereich $k=0,...,n-1$ auf Bereich $k=0,...,m-1$ anpassen

Lösungsschritt

$$w_j\sum_{k=0}^{n-1}c_k\exp(\frac{2\pi ikj}{n})=$$

$$=\sum_{k=0}^{n/3-1}\left(c_{3k}\exp(\frac{2\pi i3kj}{n})+c_{3k+1}\exp(\frac{2\pi i(3k+1)j}{n})+c_{3k+2}\exp(\frac{2\pi i(3k+2)j}{n})\right)=$$

$$=\sum_{k=0}^{m-1}\left(c_{3k}\exp(\frac{2\pi ikj}{m})\right)+\exp(\frac{2\pi ij}{3m})\sum_{k=0}^{m-1}\left(c_{3k+1}\exp(\frac{2\pi ikj}{m})\right)+\exp(\frac{4\pi ij}{3m})\sum_{k=0}^{m-1}\left(c_{3k+2}\exp(\frac{2\pi ikj}{m})\right)$$

für $j=0,...,m-1$.

Tip

Übergang von $j \to m + j$

Lösungsschritt

Durch Ersetzen von $j \to m + j$ erhält man die zweite Gruppe von Einträgen:

$$w_{m+j} = \sum \left(c_{3k} \exp(\frac{2\pi ik(m+j)}{m}) \right) + \ldots =$$

$$= \sum \left(c_{3k} \exp(\frac{2\pi ikj}{m}) \right) +$$

$$+ \exp(2\pi i/3) \exp(\frac{2\pi ij}{3m}) \sum \left(c_{3k+1} \exp(\frac{2\pi ikj}{m}) \right) +$$

$$+ \exp(4\pi i/3) \exp(\frac{4\pi ij}{3m}) \sum \left(c_{3k+2} \exp(\frac{2\pi ikj}{m}) \right)$$

Tip

Übergang von $j \to 2m + j$

Lösungsschritt

Durch Ersetzen von $j \to 2m + j$ erhält man die dritte Gruppe von Einträgen:

$$w_{2m+j} = \sum \left(c_{3k} \exp(\frac{2\pi ik(2m+j)}{m}) \right) + \ldots =$$

$$= \sum \left(c_{3k} \exp(\frac{2\pi ikj}{m}) \right) +$$

$$+ \exp(4\pi i/3) \exp(\frac{2\pi ij}{3m}) \sum \left(c_{3k+1} \exp(\frac{2\pi ikj}{m}) \right) +$$

$$+ \exp(8\pi i/3) \exp(\frac{4\pi ij}{3m}) \sum \left(c_{3k+2} \exp(\frac{2\pi ikj}{m}) \right)$$

Berechnet man also die IDFT der drei kurzen Vektoren mit Komponenten c_{3k}, c_{3k+1} und c_{3k+1}, so kann man diese drei sich ergebenden Vektoren, mit passenden Faktoren versehen, zu der gesuchten Lösung kombinieren.

5 Iterative Verfahren

5.1 Aufgabe 20 (Buch, Übung: 82)

5.1.1 Aufgabenstellung

Zeigen Sie, dass die Iteration

$$x_{n+1} = \cos(x_n)$$

für alle $x_0 \in \mathbb{R}$ gegen den einzigen Fixpunkt χ, $\chi = \cos(\chi)$ konvergiert.

a) Formulieren Sie ferner ein Newton-Verfahren zur Berechnung von χ. Konvergiert dieses Verfahren ebenfalls für jeden beliebigen Startwert?

b) Vergleichen Sie die Konvergenzgeschwindigkeit der beiden Interationsverfahren.

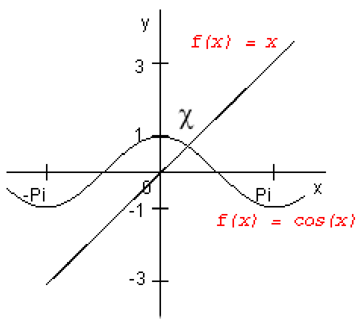
5.1.2 Lösung

Tip

Betrachte $x_{n+2} = \cos(\cos(x_n))$, also nur jeden zweiten Iterationsschritt.

Lösungsschritt

Die folgende Grafik zeigt, dass es nur einen Fixpunkt für $x = \cos(x)$ gibt:



Wir betrachten $h(x) := \cos(\cos(x))$, also nur jeden zweiten Iterationsschritt. Konvergiert dies, so konvergiert auch die ursprüngliche Folge. Dann gilt:

$$|h'(x)| = |\sin(\cos(x))| \cdot |\sin(x)| \leq \sin(1) < 1$$

$$\Rightarrow |h(x) - h(y)| = |h'(z)| \cdot |x - y| \leq \sin(1) \cdot |x - y|$$

mit z als Zwischenstelle. Daraus folgt h kontrahierend und stets konvergent gegen den eindeutigen Fixpunkt. Hieraus folgt wiederum $f(x)$ konvergent.

zu a)

Lösungsschritt

Newton liefert:

$$\Phi(x) = x - \frac{f(x)}{f'(x)} = x + \frac{x - \cos(x)}{1 + \sin(x)} = \frac{x\sin(x) + \cos(x)}{1 + \sin(x)}$$

Lösungsschritt

Schlechte Konvergenz bzw. keine Konvergenz falls $f'(x_0) \approx 0$ bzw. $f'(x_0) = 0$.

$$f'(x) = 1 + \sin(x) = 0 \Leftrightarrow x = \frac{3}{2}\pi + 2k\pi$$

D.h. für einen Startwert $x_0 = \frac{3}{2}\pi + 2k\pi$ konvergiert unser Verfahren nicht.

zu b)

Lösungsschritt

Da $\chi \neq \frac{3}{2}\pi + 2k\pi$ gilt:

$$f'(\chi) = 1 + \sin(\chi) \neq 0$$

Somit ist χ eine einfache Nullstelle und das Newton-Verfahren konvergiert quadratisch.

Lösungsschritt

$f(x) = \cos(x)$ konvergiert offensichtlich linear.

5.2 Aufgabe 21 (Buch, Übung: 83)

5.2.1 Aufgabenstellung

Wir suchen eine einfache Nullstelle χ einer Funktion $g(x)$. Natürlich ist dann χ auch eine Nullstelle von $g^2(x)$.

a) Formulieren Sie das Newton-Verfahren zur Bestimmung von χ einmal mittels der Funktion $g(x)$ und einmal mittels der Funktion $g^2(x)$ (unter Verwendung von $g'(x)$).

b) Da χ eine einfache Nullstelle ist, können wir $g(x) = (x - \chi)h(x)$ schreiben mit $h(\chi) \neq 0$. Welches der beiden Verfahren konvergiert schneller? Wie ist der Unterschied zu erklären?

c) Sei nun χ eine zweifache Nullstelle einer Funktion $f(x)$. Wie kann daher das Newton-Verfahren modifiziert werden, so dass schnelle Konvergenz erreicht wird? Begründung!

5.2.2 Lösung

zu a)

Lösungsschritt

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}$$

Lösungsschritt

$$x_{k+1} = x_k - \frac{g^2(x_k)}{(g^2(x_k))'} = x_k - \frac{g^2(x_k)}{2g(x_k)g'(x_k)} = x_k - \frac{1}{2}\frac{g(x_k)}{g'(x_k)}$$

zu b)

Tip

Konvergenz des Newton-Verfahren bei einfacher bzw. doppelter Nullstelle.

Lösungsschritt

Bei einer einfachen Nullstelle hat $g^2(x)$ natürlich eine doppelte Nullstelle. Das Newton-Verfahren konvergiert bei einer einfachen Nullstelle quadratisch, bei einer doppelten Nullstelle nur linear.

zu c)

Tip

Betrachtung von $\sqrt{|g(x)|}$.

Lösungsschritt

Hat $g(x)$ eine doppelte Nullstelle, so hat $\sqrt{|g(x)|}$ dort eine einfache Nullstelle. Verwende also das Newton-Verfahren z.B. mit der Wurzel. Dies entspricht der Iteration

$$x_{k+1} = x_k - \frac{\sqrt{g(x_k)}}{\frac{1}{2\sqrt{g(x_k)}} g'(x_k)} = x_k - 2\frac{g(x_k)}{g'(x_k)}$$

Tip

Betrachtung von $g'(x)$.

Lösungsschritt

Hat $g(x)$ eine doppelte Nullstelle, so hat $g'(x)$ dort eine einfache Nullstelle. Verwende also das Newton-Verfahren z.B. mit der Ableitung. Dies entspricht der Iteration

$$x_{k+1} = x_k - \frac{g'(x)}{g''(x)}$$

5.3 Aufgabe 22 (Buch, Übung: 84)

5.3.1 Aufgabenstellung

Zur Berechnung von $1/b$ suchen wir eine Nullstelle χ der Funktion $f(x) = 1/x - b, b \neq 0$.

a) Formulieren Sie das Newton-Verfahren zur Bestimmung von χ. Für welche Startwerte konvergiert das Verfahren in einem Schritt?

b) Für welche b ist das Newton-Verfahren lokal quadratisch Konvergent?

c) Diskutierten Sie die Iteration, die durch die Festlegung $\Phi(x) = bx^2, x_{k+1} = \Phi(x)$ definiert ist. Wo liegen evtl. Fixpunkte? Ist dieses Verfahren zur Berechnung von $1/b$ zu empfehlen?

5.3.2 Lösung

zu a)

Lösungsschritt

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{1/x_k - b}{-1/x_k^2} = 2x_k - bx_k^2$$

Tip

Finde Lösung zu $x_1 = 1/b$.

Lösungsschritt

Konvergenz in einem Schritt, falls gilt:

$$x_1 = 1/b = 2x_0 - bx_0^2 \Leftrightarrow$$
$$(bx_0 - 1)^2 = 0 \Leftrightarrow$$
$$x_0 = 1/b$$

Konvergenz also in einem Schritt wenn der Startwert gleich der gesuchten Lösung ist.

<u>zu b)</u>

Tip

Wann ist das Newton-Verfahren quadratisch konvergent?

Lösungsschritt

Das Newton-Verfahren ist quadratisch konvergent, falls nur einfache Nullstellen vorliegen. Nun gilt:

$$f'(x) = -\frac{1}{x^2}$$

also

$$f'(1/b) = -b^2 \neq 0$$

Somit ist $1/b$ eine einfache Nullstelle und daher liegt lokal quadratische Konvergenz vor.

<u>zu c)</u>

Tip

Fixpunkt liegt vor, wenn $\Phi(x) = x$.

Lösungsschritt

Fixpunkt bei:

$$\Phi(x) = x \Leftrightarrow$$
$$x = bx^2 \Leftrightarrow$$
$$x = 0 \text{ oder } x = 1/b$$

Tip

Anstossender/anziehender Fixpunkt?

Lösungsschritt

Es gilt:

$$\Phi'(x) = 2bx \text{ und somit } \Phi'(1/b) = 2b/b = 2$$

Daher ist 1/b ein abstossender Fixpunkt. Die Folge x_k wird also gegen 0 oder $\pm\infty$ konvergieren, aber nur für den exakten Startwert $x_0 = 1/b$ auch gegen $1/b$. Zudem wäre maximal mit linearer Konvergenz zu rechen, das Verfahren ist also nicht geeignet.

5.4 Aufgabe 23 (Buch, Übung: 85)

5.4.1 Aufgabenstellung

Wir betrachten die Iterationsfunktion $\Phi(x) = (x^2 + 1)/(2x)$. Welche Fixpunkte hat Φ? Sind diese Fixpunkte abstossend oder anziehend?
Zeigen Sie dass für einen Startwert x_0 gilt:

$$|x_{k+1} - 1| \leq \frac{1}{2x_k}(x_k - 1)^2$$

Was folgt daraus für die Konvergenzordnung dieses Verfahrens?
Man bestimme ein Intervall I so, dass für $x \in I$ die Funktion Φ kontrahierend ist und eine Abbildung $\Phi : I \to I$ darstellt.

5.4.2 Lösung

Tip

Fixpunkt liegt vor, wenn $\Phi(x) = x$.

Lösungsschritt

$$\begin{array}{lcll} \Phi(x) & = & x & \Leftrightarrow \\ x & = & (x^2+1)/(2x) & \Leftrightarrow \\ x^2+1 & = & 2x^2 & \Leftrightarrow \\ x & = & \pm 1 & \end{array}$$

Somit sind die Fixpunkte $x_{1,2} = \pm 1$.
Weiter gilt:

$$\begin{array}{rcl} \Phi'(x) & = & \frac{1}{2} - \frac{1}{x^2} \\ \Phi(x_{1,2}) & = & 0 \end{array}$$

Daher sind $x_{1,2}$ anziehende Fixpunkte.

Lösungsschritt

$$\begin{aligned} x_{k+1} - 1 &= \tfrac{1}{2}(x_k + \tfrac{1}{x_k} - 1 - 1) &= \\ &= \tfrac{1}{2}(x_k - 1 - \tfrac{1}{x_k}(x_k - 1)) &= \\ &= \tfrac{1}{2x_k}(x_k - 1)(x_k - 1) &= \\ &= \tfrac{(x_k-1)^2}{2x_k} \end{aligned}$$

Somit quadratische Konvergenz, da $\frac{1}{2x_k}$ für den Fall der Konvergenz beschränkt bleibt.

Lösungsschritt

z.B. $I = [1,2]$; dann ist $0 \leq \Phi'(x) \leq 1/2$ in I (nach a)), also kontrahierend. Ausserdem gilt für $x \in I$: $\Phi(x) \geq 1$ und $\Phi(x)$ ist näher an 1 als x, insgesamt also $\Phi(x) \in I$.

5.5 Aufgabe 24 (Buch, Übung: 86)

5.5.1 Aufgabenstellung

Das Heron-Verfahren ist ein Verfahren, um die Quadratwurzel einer beliebigen positiven Zahl mit festem Aufwand auf Maschinengenauigkeit zu berechnen. Es basiert auf der Bestimmung der Nullstelle der Funktion $f(x) = x^2 - a$ mittels des Newton-Verfahrens.

a) Bestimmen Sie zunächst mittels linearer Interpolation für $a \in [1,4)$ einen guten Näherungswert für $\sqrt{a}$. Verwenden Sie die Funktion

$$s(x) = \frac{1}{24}(8x + 17)$$

als Näherung an die Wurzelfunktion. Schätzen Sie den relativen Fehler des gefundenen Näherungswertes ab (also die Größe von $(s(x) - \sqrt{x})/\sqrt{x}$).

b) Formulieren Sie das Newton-Verfahren zur Bestimmung der Nullstelle $\sqrt{a}$ von $f(x)$. Wie entwickelt sich der Fehler während der Iteration? Wie viele Iterationen sind nötig, um $\sqrt{a}$ bis auf Maschinengenauigkeit zu berechnen, falls $a \in [1,4)$ ist?

c) Wie kann das obige Verfahren angewandt werden, falls $a \notin [1,4)$ gilt? Dazu schreiben wir a in der Form $a = \alpha 2^{2p}$ mit passendem p und $\alpha \in [1,4)$.

5.5.2 Lösung

zu a)

Tip

Für die lineare Interpolation $L(x)$ gilt:

$$x_0 < x < x_1, \qquad y_0 = f(x_0), \qquad y_1 = f(x_1), \qquad h = x_1 - x_0$$

$$\Rightarrow L(x) = y_0 + \frac{y_1 - y_0}{h}(x - x_0)$$

Lösungsschritt

Hier gilt:

$$a_0 < a < a_1 \quad \text{mit } a_0 = 1, \quad a_1 = 4 \quad \Rightarrow h = 3$$

$$\begin{aligned} f(a_0) &= \frac{25}{24} = y_0 \\ f(a_1) &= \frac{49}{24} = y_1 \end{aligned}$$

Somit ist $L(x)$:

$$\begin{aligned} L(a) &= \frac{25}{24} + \frac{\frac{49}{25} - \frac{25}{24}}{3}(a-1) \\ &= \frac{25}{24} + \frac{1}{3}(a-1) \\ &= \frac{1}{3}a + \frac{17}{24} \end{aligned}$$

Lösungsschritt

Für den relativen Fehler des Näherungswertes gilt:

$$f_{rel} = \frac{\tilde{L}(a) - \sqrt{a}}{\sqrt{a}}$$

wobei

$$\begin{aligned} \tilde{L}(a) &= ((\frac{1}{3}a)(1+\epsilon_1) + \frac{17}{24})(1+\epsilon_2) \\ &\doteq \frac{1}{3}a\epsilon_1 + (\frac{1}{3}a + \frac{17}{24})\epsilon_2 + \frac{1}{3}a + \frac{17}{24} \end{aligned}$$

Somit gilt für den relativen Fehler:

$$f_{rel} = \frac{\frac{1}{3}a\epsilon_1 + (\frac{1}{3}a + \frac{17}{24})\epsilon_2 + \frac{1}{3}a + \frac{17}{24} - \frac{1}{3}a - \frac{17}{24}}{\frac{1}{3}a + \frac{17}{24}} = \frac{a}{a + \frac{17}{8}}\epsilon_1 + \epsilon_2$$

<u>zu b)</u>

Lösungsschritt

Es gilt:

$$f(x) = x^2 - a \Rightarrow \Phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x^2 - a}{x} = \frac{x^2 + a}{2x}$$

Tip

Für den relativen Fehler im k-ten Schritt gilt:

$$f_{rel_k} = \frac{x_k - \sqrt{a}}{\sqrt{a}}$$

Einsetzten der Heron-Rekursion(Newton-Verfahren im (k-1)-ten Schritt) für x_k.

Lösungsschritt

Für den relativen Fehler im k-ten Schritt gilt:

$$f_{rel_k} = \frac{x_k - \sqrt{a}}{\sqrt{a}}$$

Setzen wir die vorher ermittelte Newton-Form für x_k ein, so erhalten wir:

$$\begin{aligned}
f_{rel_k} &= \frac{\frac{x_{k-1}^2 + a}{x_{k-1}} - \sqrt{a}}{\sqrt{a}} = \frac{1}{\sqrt{a}}\left(\frac{x_{k-1}^2 + a}{2x_{k-1}} - \sqrt{a}\right) \\
&= \frac{1}{2\sqrt{a}x_{k-1}}\left(x_{k-1}^2 - 2\sqrt{a}x_{k-1} + a\right) = \frac{\sqrt{a}}{2x_{k-1}}\left(\frac{x_{k-1} - \sqrt{a}}{\sqrt{a}}\right)^2 \\
&= \frac{1}{2\left(\frac{x_{k-1} - \sqrt{a}}{\sqrt{a}} + 1\right)}(f_{rel_{k-1}})^2 = \frac{1}{2(1 + f_{rel_{k-1}})}(f_{rel_{k-1}})^2
\end{aligned}$$

Tip

Abschätzen einer oberen Schranke für f_{rel_k}, z.B. $f_{rel_k} \leq \frac{1}{2}(f_{rel_{k-1}})^2$ und anschließendes Auflösen nach k.

Lösungsschritt

Wir schätzen unseren relativen Fehler im k-ten Schritt ab:

$$\begin{aligned}
f_{rel_k} &\leq \frac{1}{2}(f_{rel_{k-1}})^2 \leq \frac{1}{2}\left(\frac{1}{2}(f_{rel_{k-2}})^2\right)^2 \leq \ldots \leq 2^{-1}2^{-2}\cdots 2^{-2^{k-2}}(f_{rel_1})^{2^{k-1}} \\
&= 2^{-1}2^{-2}\cdots 2^{-2^{k-2}}2^{-2^{k-1}}(1+f_{rel_0})^{-2^{k-1}}(f_{rel_0})^{2^k} \\
&= 2^{-(2^k-1)}(1+f_{rel_0})^{-2^{k-1}}(f_{rel_0})^{2^k}, \quad k \geq 1
\end{aligned}$$

Nun müssen wir noch nach k auflösen:

$$\begin{aligned}
\left(\frac{1}{2}\right)^{2^k-1}\left(\frac{1}{1+f_{rel_0}}\right)^{2^{k-1}}(f_{rel_0})^{2^k} &\leq \epsilon_{masch} \\
2\cdot\left(\underbrace{\frac{1}{2}\cdot\sqrt{\frac{1}{1+f_{rel_0}}}\cdot f_{rel_0}}_{=:\alpha}\right)^{2^k} &\leq \epsilon_{masch} \\
(\alpha)^{2^k} &\leq \frac{\epsilon_{masch}}{2} \\
2^k\cdot\log(\alpha) &\leq \log\left(\frac{\epsilon_{masch}}{2}\right) \\
2^k &\geq \frac{\log(\frac{\epsilon_{masch}}{2})}{\log(\alpha)} \\
k &\geq \log\left(\frac{\log(\frac{\epsilon_{masch}}{2})-\log(\alpha)}{\log(2)}\right)
\end{aligned}$$

<u>zu c)</u>

Lösungsschritt

Hierzu passen wir einfach $s(x)$ und unser gefundenes Newton-Verfahren an und benutzen $s_0(\hat{=}s(\alpha))$ als Startwert.

$$\begin{aligned}
s(x) = \frac{1}{24}(8x+17) \quad &\rightarrow \quad s(\alpha) = \frac{1}{24}(8\alpha+17) \\
x_{k+1} = \frac{x_k^2+a}{2x_k} \quad &\rightarrow \quad s_{k+1} = \frac{s_k^2+\alpha}{2s_k}, \quad k = 0,...,N-1
\end{aligned}$$

wobei N die Anzahl der Iterationsschritte ist.

Lösungsschritt

Die gesuchte Lösung für $\sqrt{a}$ ist dann: $s_N\cdot 2^p$

5.6 Aufgabe 25 (Buch, Übung: 88)

5.6.1 Aufgabenstellung

Sei f eine zweimal stetig stetig differenzierbare Funktion. Wir wollen mittels Newton-Verfahren eine Nullstelle unbekannter Ordnung bestimmen, unter Verwendung der Funktion $h(x) := f(x)/f'(x)$. Was lässt sich über die Konvergenz des Newton-Verfahrens, angewandt auf die Funktion h, sagen?

5.6.2 Lösung

Tip

Schreiben von $f(x)$ als $f(x) = (x - x_0)^k \cdot g(x)$

Lösungsschritt

Wir schreiben $f(x)$ als $f(x) = (x - x_0)^k \cdot g(x)$ mit x_0 als Nullstelle der Ordnung k.

$$\Rightarrow f'(x) = k(x - x_0)^{k-1} \cdot g(x) + (x - x_0)^k \cdot g'(x)$$

Lösungsschritt

Somit folgt für $h(x)$:

$$\begin{aligned} h(x) &= \frac{f(x)}{f'(x)} = \frac{(x - x_0)^k \cdot g(x)}{k(x - x_0)^{k-1} \cdot g(x) + k(x - x_0)^k \cdot g(x)} \\ &= \frac{g(x)}{\frac{k \cdot g(x)}{x - x_0} + g'(x)} \\ &= (x - x_0) \cdot \frac{g(x)}{k \cdot g(x) + (x - x_0) \cdot g'(x)} \end{aligned}$$

Lösungsschritt

Somit ist x_0 also eine einfache Nullstelle und das Newton-Verfahren konvergiert quadratisch.

6 Differentialgleichungen

6.1 Aufgabe 26 (Buch, Übung: 114)

6.1.1 Aufgabenstellung

Gegeben ist ein Anfangswertproblem der Form $y'(x) = \Phi(y, x), y(a) = y_0$. Bei äquidistanter Schrittweite h betrachten wir die Iterationsvorschrift

$$y_{k+1} = y_k + \frac{h}{2} \left(\Phi(y_k, x_k) + \Phi(y_{k+1}, x_{k+1}) \right).$$

a) Aus welcher Quadraturregel ergibt sich dieses Verfahren? Man beschreibe den Zusammenhang.

b) Zeigen Sie, wie man aus der obigen Beziehung bei bekannten Werten von x_k, y_k und h den neuen Wert y_{k+1} berechnen kann. Um welche Art von Problem handelt es sich dabei? Welche Art von Verfahren kann man benutzen? Beschreiben Sie die Startbedingungen eines solchen Verfahrens.
Geben Sie einen Algorithmus an zu Lösung des AWPs mit näherungsweiser Berechnung von $y(b)$ für ein $b > a$.

6.1.2 Lösung

zu a)

Lösungsschritt

Dieses Verfahren stammt aus der Trapezregel, und zwar

$$\int_{x_k}^{x_{k+1}} y' dx = y_{k+1} - y_k \approx \frac{x_{k+1} - x_k}{2} \cdot (y'_{k+1} + y'_k) = \frac{h}{2}(\Phi(y_k, x_k) + \Phi(y_{k+1}, x_{k+1}))$$

zu b)

Lösungsschritt

Das Problem besteht hierin, dass der neue Wert y_{k+1} nicht explizit gegeben ist, sondern nur implizit, also erst noch durch ein zusätzliches Verfahren aus der definierten Gleichung berechnet werden muss. Hier bieten sich zum einen eine einfache Fixpunktiteration oder

zum anderen das Newton-Verfahren an.
Fixpunktiteration: Betrachte die Funktion

$$f(y) = y_k + \frac{h}{2}(\Phi(y_k, x_k) + \Phi(y, x_{k+1}))$$

und bestimme den Fixpunkt $y = f(y)$. Da $|df/dy|$ für kleine Werte von h auch klein wird, ist f kontrahierend.
Newton-Verfahren: Betrachte die Funktion

$$f(y) = y_k + \frac{h}{2}(\Phi(y_k, x_k) + \Phi(y, x_{k+1})) - y$$

und suche die Nullstelle.

Lösungsschritt

Als Startwert kann man jeweils das alte y_k benutzen.
Konvergiert die erzeugte Folge y gegen eine Zahl $\tilde{y}$, so ist $\tilde{y}$ ein Fixpunkt bzw. eine Nullstelle, und wir können $y_{k+1} = \tilde{y}$ setzen.

6.2 Aufgabe 27 (Buch, Übung: 115)

6.2.1 Aufgabenstellung

Gegeben ist das Anfangswertproblem

$$y'(x) = \varphi(y, x) = x, \qquad x_0 = 0, \quad y(x_0) = 0.$$

Wir benutzen die äquidistante Einteilung $x_0 = 0, x_1 = h, \ldots, x_n = nh = b$.

a) Bestimmen Sie die exakte Lösung $y(x)$ und die Näherrungswerte $y_j \approx y(x_j), j = 1, \ldots, n$, die durch das (Vorwärts-)Euler-Verfahren gegeben werden.

b) Wie groß ist der lokale Diskretisierungsfehler $|y_1 - y(x_1)|$? Wie groß ist der globale Fehler $|y_n - y(b)|$?

6.2.2 Lösung

zu a)

Tip

Trennung der Variablen.

Lösungsschritt

Hier liegt eine Gleichung vom Typ

$$y'(x) = f(x)g(y)$$

mit $g(y) = 1$ und $f(x) = x$ vor. Mittels Trennung der Variablen erhalten wir:

$$\begin{array}{lcl} \frac{1}{g(y)}dy & = & f(x)dx \quad \Leftrightarrow \\ \\ dy & = & xdx \end{array}$$

Durch Integration jeder Seite (unbestimmt) erhalten wir:

$$G(y) = \int dy = y, \qquad F(x) = \int xdx = \frac{x^2}{2}$$

Mit der Allgemeinen Lösung

$$G(y) - F(x) = c, \qquad c \in \mathbb{R}$$

können wir die exakte Lösung für $c_0 = G(y_0) - F(x_0)$ berechnen. Hier also $c_0 = 0, \Rightarrow$ unsere Lösung lautet:

$$y = \frac{x^2}{2}$$

Lösungsschritt

Das Euler-Verfahren liefert

$$\begin{array}{lclclcl} y_1 & = & y_0 + h(\phi(y_0, x_0) & = & 0 + hx_0 & = & 0 \\ y_2 & = & y_1 + h(\phi(y_1, x_1) & = & 0 + hx_1 & = & h^2 \\ y_3 & = & x_2 + h(\phi(y_2, x_2) & = & h^2 + 2h^2 & = & 3h^2 \\ \ldots \end{array}$$

oder allgemein

$$y_k = h^2 \cdot \sum_{j=1}^{k-1} j = h^2 \frac{(k-1)k}{2}$$

zu b)

Lösungsschritt

lokaler Fehler:

$$y(x_1) - y_1 = \frac{x_1^2}{2} - y_1 = \frac{h^2}{2} - 0 = \frac{h^2}{2} = \frac{b^2}{2n^2}$$

globaler Fehler:

$$y(b) - y_n = y(x_n) - y_n = \frac{b^2}{2} - \frac{h^2(n-1)n}{2} = \frac{b^2}{n} - \frac{(b^2/n^2)(n^2-n)}{2} = \frac{(bh)}{2} = \frac{b^2}{2n}.$$

Der lokale Fehler ist also $\mathcal{O}(h^2)$ und der globale Fehler $\mathcal{O}(h)$.

6.3 Aufgabe 28 (Buch, Übung: 116)

6.3.1 Aufgabenstellung

Gegeben ist das Anfangswertproblem

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} -1000 & \alpha \\ \alpha & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, x_0 =, u(x_0) = 1, v(x_0) = 1.$$

Wir benutzen die äquidistante Einteilung $x_0 = 0, x_1 = x_0 + h, \ldots, x_n = x_0 + nh = b$, also $h = b/n$.

a) Formulieren Sie das (Vorwärts-)Euler-Verfahren zur Bestimmung von Näherungslösungen $u_j \approx u(x_j)$ und $v_j \approx v(x_j)$.

Im Folgenden betrachten wir den Spezialfall $\alpha = 0$.

b) Bestimmen Sie die exakten Lösungen $u(x)$ und $v(x)$. Man gebe eine einfache Formel für die Näherungswerte $u_j \approx u(x_j)$ und $v_j \approx v(x_j)$ an, die durch das (Vorwärts-)Euler-Verfahren gegeben werden, in Abhängigkeit von h.

b) Wie groß darf h gewählt werden, so dass u_j und v_j beide für $j \to \infty$ das richtige Verhalten zeigen? Wie groß darf h gewählt werden, so dass v_j für $j \to \infty$ das richtige Verhalten zeigt?

6.3.2 Lösung

zu a)

Lösungsschritt

$$\begin{pmatrix} u_{k+1} \\ v_{k+1} \end{pmatrix} = \begin{pmatrix} u_k \\ v_k \end{pmatrix} + h \begin{pmatrix} -1000 & \alpha \\ \alpha & -1 \end{pmatrix} \begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} (1-1000h)u_j + \alpha v_j \\ \alpha u_j + (1-h)v_j \end{pmatrix}$$

für $j = 1, \ldots, n$ und $u_0 = v_0 = 1$.

zu b)

Tip

Homogene lineare Differentialgleichung 1. Ordnung.

Lösungsschritt

Uns liegt eine lineare Differentialgleichung 1. Ordnung, also $y'(x) = a(x)y$, vor. Die Lösung hierzu lautet:

$$y(x) = ce^{A(x)}, \qquad c \in \mathbb{R} \qquad \text{mit } A(x) = \int a(x)dx$$

In unseren Fall lauten die Lösungen, mit $u(x_0) = v(x_0) = 1$:

$$u(x) = e^{-1000x}, \qquad v(x) = e^{-x}$$

Lösungsschritt

Das Euler-Verfahren liefert für $u(x)$:

$$\begin{array}{lllll} u_1 & = & (1-100hu_0) & = & (1-1000)h \\ u_2 & = & (1-1000h)(1-1000h) & = & (1-1000)^2 h \\ \dots & & & & \end{array}$$

Insgesamt also:

$$u_j = (1-1000h)^j \qquad \text{und analog} \qquad v_j = (1-h)^j$$

zu c)

Lösungsschritt

Es gilt: $u(x) = v(x) = 0$ für $x \to \infty$. Daher sollten die Näherungslösungen dasselbe Verhalten zeigen. Das stimmt aber für $u_j(x)$ und $v_j(x)$ nur gemeinsam, falls $0 < h < 1/1000$ (dann auch monoton fallendes Verhalten), oder $0 < h < 2/1000$ (dann alternierend gegen 0).
Betrachten wir dagegen nur die Folge v_j, so ist diese Bedingung schon für $h < 1$, bzw. $h < 2$, erfüllt.