Prof. Dr. M. Ehrhardt
D. Heubes, M.Sc.
Sommersemester 2011

Bergische Universität Wuppertal
Fachbereich C – Mathematik und Naturwissenschaften
Angewandte Mathematik / Numerische Analysis

Nachklausur zu

Einführung in die Numerische Mathematik

11.10.2011

Lösungen

Die besten 4 Aufgaben werden gewertet

Schreibzeit beträgt 90 Minuten

Mit 40 Punkten ist die Klausur bestanden

Viel Erfolg!

Name: Musterstudentin

Vorname: Susanne

Matrikel-Nr.: 123456

Studiengang: Ba. Applied Science

Mit der Veröffentlichung meiner Note unter Angabe der Matrikelnummer im Internet bin ich einverstanden:
☐ ja ☐ nein

Unterschrift: ______________________________

Raum G.13.18
Sitzplatz 1

Aufgabe	1			2			3				4			5		$\sum$
	a	b	c	a	b	c	a	b	c	d	a	b	c	a	b	
max. Punkte	5	4	11	8	8	4	6	3	7	4	4	12	4	8	12	
err. Punkte																

Punkte in der Klausur	
Note	

Klausuraufgabe 1: (5+4+11 Punkte)

Wir betrachten das überbestimmte Gleichungssystem $Ax = b$, mit $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, mit $m > n$, $m, n \in \mathbb{N}$.

a) Unter welchen zusätzlichen Bedingungen existiert eine eindeutige Lösung für dieses Gleichungssystem?
Hinweis: Man kann die erweiterte Matrix $(A|b)$ betrachten, welche als zusätzliche Spalte den Vektor b enthält.

b) Es sei eine Zerlegung $A = Q \cdot R$ mit einer orthogonalen Matrix $Q \in \mathbb{R}^{m\times m}$ und einer Matrix $R \in \mathbb{R}^{m\times n}$

$$R = \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}$$

wobei $\tilde{R} \in \mathbb{R}^{n\times n}$ eine obere reguläre Dreiecksmatrix ist, gegeben. Zeigen Sie, dass $\tilde{R}^\top \tilde{R}$ die Cholesky-Zerlegung von $A^\top A$ ist.

c) Für die Wahl

$$A = \begin{pmatrix} -4 & 3 \\ 2 & 9 \\ 1 & -5 \\ -1 & 1 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} 3 \\ -10 \\ 7 \\ 0 \end{pmatrix}$$

besitzt das Gleichungssystem $Ax = b$ keine eindeutige Lösung. Formulieren Sie eine Aufgabe, die anstelle des Gleichungssystems, die Lösung im Sinne des kleinsten Fehlerquadrates beschreibt und bestimmen Sie damit eine Lösung $x^\star$ des so formulierten Problems. Ist die berechnete Lösung eindeutig?

Lösungsvorschlag:

a) Eine Lösung existiert nur dann, wenn das Gleichungssystem linear abhängig ist und das System äquivalent, durch höchstens n linear unabhängige Gleichungen, geschrieben werden kann. Eindeutigkeit liegt vor, wenn genau n linear unabhängige Gleichungen notwendig sind.
Betrachtet man die erweiterte Matrix $(A|b)$ muss diese Matrix Spaltenrang n haben, damit eine eindeutige Lösung existiert.

b) Nutzt man die gegebene Zerlegung $A = Q \cdot R = Q \cdot \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}$ aus, erhält man für die symmetrische Matrix $A^\top A$ die Zerlegung

$$A^\top A = R^\top \cdot \underbrace{Q^\top \cdot Q}_{=I} \cdot R = R^\top \cdot R = (\tilde{R}^\top \; 0^\top) \cdot \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix} = \tilde{R}^\top \cdot \tilde{R} + 0^\top \cdot 0 = \tilde{R}^\top \cdot \tilde{R}.$$

Nach Voraussetzung ist $\tilde{R}$ eine reguläre obere Dreiecksmatrix, daher ist $\tilde{R}^\top$ eine untere Dreiecksmatrix, es handelt sich also um die Cholesky-Zerlegung. Es folgt aus der Regularität von $\tilde{R}$ die positive Definitheit von $A^\top A$, was hier jedoch nicht untersucht werden sollte.

c) Da keine Lösung für $Ax = b$ existiert, fordern wir für $x^\star$

$$\min_{x\in\mathbb{R}^2} \|A\cdot x - b\|_2^2 = \|A\cdot x^\star - b\|_2^2.$$

Dieses Minimierungsproblem lässt sich über die Normalengleichungen $A^\top Ax = A^\top b$ lösen, wir berechnen also zunächst $A^\top A$ und $A^\top b$:

$$A^\top A = \begin{pmatrix} -4 & 2 & 1 & -1 \\ 3 & 9 & -5 & 1 \end{pmatrix} \begin{pmatrix} -4 & 3 \\ 2 & 9 \\ 1 & -5 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 22 & 0 \\ 0 & 116 \end{pmatrix}$$

$$A^\top b = \begin{pmatrix} -4 & 2 & 1 & -1 \\ 3 & 9 & -5 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -10 \\ 7 \\ 0 \end{pmatrix} = \begin{pmatrix} -25 \\ -116 \end{pmatrix}$$

Die Normalengleichungen lauten also

$$\begin{pmatrix} 22 & 0 \\ 0 & 116 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -25 \\ -116 \end{pmatrix}$$

wobei das Gleichungssystem in Form von zwei unabhängigen Gleichungen vorliegt. Es folgt unmittelbar $x_1 = -\frac{25}{22}$ und $x_2 = -1$, also

$$x^\star = \begin{pmatrix} -\frac{25}{22} \\ -1 \end{pmatrix},$$

es liegt eine eindeutige Lösung vor, da A vollen Spaltenrang hat.

Klausuraufgabe 2: (8+8+4 Punkte)

Die Lösung des linearen Gleichungssystems

$$\begin{pmatrix} 5 & 2 & -1 \\ 1 & 10 & 1 \\ 0 & -2 & -5 \end{pmatrix} \cdot x = \begin{pmatrix} 6 \\ 0 \\ 5 \end{pmatrix}$$

ist gesucht, die Matrix ist regulär eine eindeutige Lösung existiert demnach.

a) Wie lautet die Iterationsmatrix im Jacobi-Verfahren allgemein und im vorliegenden Fall. Nennen Sie eine hinreichende und notwendige Bedingung für die Konvergenz des Verfahrens. Untersuchen sie zudem, ob das Jacobi-Verfahren für den vorliegenden Fall konvergiert, hierzu reicht es mit hinreichenden Kriterien zu argumentieren.

b) Berechnen Sie, ausgehend vom Startwert $x^{(0)} = (0,0,0)^\top$, die ersten beiden Schritte des Jacobi-Verfahren, es sei dann $x^{(2)}$ die Näherungslösung. Runden Sie anschliessend die Einträge in $x^{(2)}$ auf ganze Zahlen. Verifizieren Sie auf diese Weise die exakte Lösung $x^\star$ erhalten zu haben. Wie groß ist die absolute Abweichung Δx der Näherungslösung $x^{(2)}$ von der exakten Lösung unter Verwendung der $\|\cdot\|_\infty$-Norm?

c) Geben Sie eine obere Schranke für eine Störung der rechten Seite in der $\|\cdot\|_\infty$-Norm an, die, unter Verwendung der Abschätzung aus der Vorlesung und $\|A^{-1}\|_\infty \leq 0.32$, einen absoluten Fehler von höchstens $\|\Delta x\|_\infty$ aus Aufgabenteil b) garantiert.

Lösungsvorschlag:

a) Allgemein lautet die Iterationsmatrix $I - B^{-1}A$, wobei beim Jacobi-Verfahren die Matrix B eine Diagonalmatrix mit den Diagonaleinträgen von A ist. Hat man eine Zerlegung von $A = L+D+R$ in strikte untere bzw. obere Dreiecksmatrix L bzw. R und Diagonalmatrix D gegeben, so lautet im Jacobi-Verfahren die Iterationsmatrix allgemein $J = I - D^{-1}A$. Im vorliegenden Fall gilt speziell

$$\begin{aligned} J &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1/5 & 0 & 0 \\ 0 & 1/10 & 0 \\ 0 & 0 & -1/5 \end{pmatrix} \begin{pmatrix} 5 & 2 & -1 \\ 1 & 10 & 1 \\ 0 & -2 & -5 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2/5 & -1/5 \\ 1/10 & 1 & 1/10 \\ 0 & 2/5 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2/5 & 1/5 \\ -1/10 & 0 & -1/10 \\ 0 & -2/5 & 0 \end{pmatrix} \end{aligned}$$

Äquivalent zur Konvergenz des Jacobi-Verfahrens (gilt auch allgemein für Iterationsverfahren) ist $\rho(J) < 1$ (allgemein $\rho(I - B^{-1}A) < 1$), wobei $\rho(J)$ den betragsmäßig größten Eigenwert von J bezeichnet. Hinreichend für die Konvergenz ist die Bedingung $\|J\| < 1$, wobei $\|\cdot\|$ eine beliebige Matrixnorm ist, die konsistent zu einer Vektornorm ist. Konkret für den vorliegenden Fall und die $\|\cdot\|_\infty$-Norm gilt

$$\|J\|_\infty = \max\left\{\frac{2}{5} + \frac{1}{5},\ \frac{1}{10} + \frac{1}{10},\ \frac{2}{5}\right\} = \max\left\{\frac{3}{5},\ \frac{1}{5},\ \frac{2}{5}\right\} = \frac{3}{5} < 1,$$

woraus folgt, dass das Jacobi-Verfahren im vorliegenden Fall konvergiert.

b) Die Iterationsvorschrift des Jacobi-Verfahrens lautet $x^{(k+1)} = Jx^{(k)} + D^{-1}b$. Es ist

$$D^{-1}b = \begin{pmatrix} 1/5 & 0 & 0 \\ 0 & 1/10 & 0 \\ 0 & 0 & -1/5 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 6/5 \\ 0 \\ -1 \end{pmatrix}$$

Damit gilt ausgehend vom Startvektor $x^{(0)} = (0, 0, 0)^\top$:

$$\begin{aligned} x^{(1)} &= Jx^{(0)} + D^{-1}b = D^{-1}b = \begin{pmatrix} 6/5 \\ 0 \\ -1 \end{pmatrix} \\ x^{(2)} &= Jx^{(1)} + D^{-1}b \\ &= \begin{pmatrix} 0 & -2/5 & 1/5 \\ -1/10 & 0 & -1/10 \\ 0 & 2/5 & 0 \end{pmatrix} \begin{pmatrix} 6/5 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 6/5 \\ 0 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} -1/5 \\ -6/50 + 5/50 \\ 0 \end{pmatrix} + \begin{pmatrix} 6/5 \\ 0 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ -1/50 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -0.02 \\ -1 \end{pmatrix} \end{aligned}$$

Der auf ganze Zahlen gerundete Vektor ist $\tilde{x}^{(2)} = (1, 0, -1)^\top$, wir berechnen

$$\begin{pmatrix} 5 & 2 & -1 \\ 1 & 10 & 1 \\ 0 & -2 & -5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 5 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 6 \\ 0 \\ 5 \end{pmatrix} \quad \checkmark$$

was zeigt, dass $\tilde{x}^{(2)} = (1, 0, -1)^\top$ die exakte Lösung des Gleichungssystems ist. Ferner gilt

$$\|\Delta x\|_\infty = \|\tilde{x}^{(2)} - x^{(2)}\|_\infty = \left\| \begin{pmatrix} 0 \\ 0.02 \\ 0 \end{pmatrix} \right\|_\infty = 0.02.$$

c) Gesucht ist eine obere Schranke für $\|\Delta b\|_\infty$, so dass der absolute Fehler höchstens 0.02 groß ist. Allgemein berechnet sich der absolute Fehler unter Vernachlässigung der $\|\Delta A\|_\infty$-Terme über

$$\begin{aligned} & \|\Delta x\|_\infty \leq \|A^{-1}\|_\infty \cdot \|\Delta b\|_\infty \leq 0.32 \cdot \|\Delta b\|_\infty \stackrel{!}{\leq} 0.02 \\ \Rightarrow \quad & \|\Delta b\|_\infty \leq \frac{1}{16}. \end{aligned}$$

Stört man die Komponenten in b maximal um $\frac{1}{16}$ resultiert ein absoluter Fehler im Lösungsvektor der kleiner als 0.02 ist.

Klausuraufgabe 3: (6+3+7+4 Punkte)

Wir betrachten ein lineares Gleichungssystem

$$A \cdot x = b$$

mit quadratischer Matrix $A \in \mathbb{R}^{n \times n}$ und rechter Seite $b \in \mathbb{R}^n$. Es seien

$$A = \begin{pmatrix} 5 & -7 & -4 \\ -10 & 18 & 2 \\ 5 & 7 & -3 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} 6 \\ -5 \\ 1 \end{pmatrix}$$

gegeben, hier gilt demnach explizit $n = 3$.

a) Bestimmen Sie die Determinante der Matrix A, verwenden Sie hierfür die LR-Zerlegung der Matrix A, welche vorher zu berechnen ist. Alle Zwischenschritte und die Matrizen L und R sind stets anzugeben.

b) Begründen Sie kurz, dass für die Matrix A mit Hilfe der Gauß-Elimination keine Zerlegung

$$A = L \cdot R$$

mit normierter unterer Dreiecksmatrix $L \in \mathbb{R}^{n \times n}$ und oberer Dreiecksmatrix $R \in \mathbb{R}^{n \times n}$ erzeugt werden kann, derart dass alle Einträge in $L = (\ell_{ij})$ vom Betrag kleiner oder gleich 1 sind, d.h. $|\ell_{ij}| \leq 1$.

c) Bestimmen Sie nun eine Zerlegung

$$A = C \cdot D \cdot E$$

mit normierter unterer Dreiecksmatrix $D \in \mathbb{R}^{n \times n}$, oberer Dreiecksmatrix $E \in \mathbb{R}^{n \times n}$ und einer regulären Matrix $C \in \mathbb{R}^{n \times n}$, so dass alle Einträge von $D = (d_{ij})$ vom Betrag kleiner oder gleich 1 sind, d.h. $|d_{ij}| \leq 1$. Die Zwischenschritte und die Matrizen C, D und E sind explizit anzugeben.

d) Beschreiben Sie in etwa 5 Sätzen wie die Zerlegung aus Aufgabenteil c) benutzt werden kann um das Ausgangsgleichungssystem zu lösen.

Lösungsvorschlag:

a) Die LR-Zerlegung erhält man durch

$$\begin{pmatrix} 5 & -7 & -4 \\ -10 & 18 & 2 \\ 5 & 7 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 5 & -7 & -4 \\ -2 & 4 & -6 \\ 1 & 14 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 5 & -7 & -4 \\ -2 & 4 & -6 \\ 1 & \frac{7}{2} & 22 \end{pmatrix}$$

Man liest ab:

$$L := \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & \frac{7}{2} & 1 \end{pmatrix} \qquad R := \begin{pmatrix} 5 & -7 & -4 \\ 0 & 4 & -6 \\ 0 & 0 & 22 \end{pmatrix}$$

Es gilt also

$$A = \begin{pmatrix} 5 & -7 & -4 \\ -10 & 18 & 2 \\ 5 & 7 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 14 & 1 \end{pmatrix} \begin{pmatrix} 5 & -7 & -4 \\ 0 & 4 & -6 \\ 0 & 0 & 22 \end{pmatrix} = L \cdot R$$

Weiterhin gilt mit Ausnutzung der Dreiecksgestalt von L und R (d.h. Determinante ist Produkt der Diagonaleinträge)

$$\det(A) = \det(L \cdot R) = \det(L) \cdot \det(R) = (1 \cdot 1 \cdot 1) \cdot (5 \cdot 4 \cdot 22) = 440.$$

b) Die LR-Zerlegung, das heißt eine Zerlegung in ein Produkt einer normierten unteren Dreiecksmatrix L und oberen Dreiecksmatrix R, ist eindeutig. Die in Aufgabenteil a) berechnete Zerlegung zeigt, dass die Bedingung $|\ell_{ij}| \leq 1$ nicht erfüllt ist. Aufgrund der Eindeutigkeit existiert also keine Zerlegung mit der geforderten Bedingung.

c) Die gesuchen Matrizen lassen sich über eine LR-Zerlegung mit Spaltenpivotsuche bestimmen, wobei die Permutationsmatrix mit C korrespondiert und die Matrizen D und E den Matrizen L und R aus der Zerlegung entsprechen.

$$\begin{aligned}
&\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}5&-7&-4\\-10&18&2\\5&7&-3\end{pmatrix}\\
\rightsquigarrow&\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}-10&18&2\\5&-7&-4\\5&7&-3\end{pmatrix}\\
\rightsquigarrow&\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}-10&18&2\\-1/2&2&-3\\-1/2&16&-2\end{pmatrix}\\
\rightsquigarrow&\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}-10&18&2\\-1/2&16&-2\\-1/2&2&-3\end{pmatrix}\\
\rightsquigarrow&\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}-10&18&2\\-1/2&16&-2\\-1/2&1/8&-2,75\end{pmatrix}
\end{aligned}$$

Man liest folgende Zerlegung ab

$$\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}5&-7&-4\\-10&18&2\\5&7&-3\end{pmatrix} = \begin{pmatrix}1&0&0\\-1/2&1&0\\-1/2&1/8&1\end{pmatrix}\begin{pmatrix}-10&18&2\\0&16&-2\\0&0&2,75\end{pmatrix}$$

woraus durch Umstellen folgt (Inverse von Permutationsmatrizen erhält man durch Transponieren)

$$\begin{aligned}
\begin{pmatrix}5&-7&-4\\-10&18&2\\5&7&-3\end{pmatrix} &= \begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}^{-1}\begin{pmatrix}1&0&0\\-1/2&1&0\\-1/2&1/8&1\end{pmatrix}\begin{pmatrix}-10&18&2\\0&16&-2\\0&0&2,75\end{pmatrix}\\
&= \underbrace{\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}}_{=:C}\underbrace{\begin{pmatrix}1&0&0\\-1/2&1&0\\-1/2&1/8&1\end{pmatrix}}_{=:D}\underbrace{\begin{pmatrix}-10&18&2\\0&16&-2\\0&0&2,75\end{pmatrix}}_{=:E}.
\end{aligned}$$

d) Das Gleichungssystem $Ax = b$ ist äquivalent zu $CDEx = b$, welches man ohne großen Aufwand durch Linksmultiplikation auf $DEx = C^\top b$ umformen kann. Dieses Gleichungssystem kann man in zwei Schritten lösen; zuerst bestimmt man einen Vektor y der $Dy = C^\top b$ löst, dies kann man durch Vorwärtssubstitution machen. Im nächsten Schritt bestimmt man x durch Lösen des Gleichungssystems $Ex = y$, was durch Rückwärtssubstitution effizient gelöst werden kann.

Klausuraufgabe 4: (4+12+4 Punkte)

Berechnen Sie näherungsweise die folgenden Integrale.

a) Verwenden Sie die zweifach ($n = 2$) zusammengesetzte Fassregel zur Approximation des Integrals

$$\int_0^{2\pi} \sqrt{t}\cos(t)\mathrm{d}t.$$

b) Verwenden Sie eine Gauß-Quadraturformel

$$Q[f] = \alpha f(0) + \beta f(x_1)$$

mit zwei Stützstellen zur Approximation von

$$I[f] = \int_0^1 (-x^2 + 1) f(x)\mathrm{d}x.$$

Bestimmen Sie zunächst die Unbekannten $\alpha, \beta \in \mathbb{R}$ und $0 \neq x_1 \in \mathbb{R}$ in $Q[f]$, so dass die Quadraturformel größtmöglichen Exaktheitsgrad hat. Weisen Sie nach, dass $Q[x^3] \neq I[x^3]$ ist.

c) Begründen Sie in jeweils ca. 2-3 Sätzen,

- dass man zwar eine Quadraturformel

$$J[f] = w_1 f(x_0) + w_2 f(x_1)$$

konstruieren kann, welche die Integrale $I[x^k]$, $k = 3, 4, 5$ (mit dem Integral $I[f]$ aus Aufgabenteil b)) exakt berechnen kann,

- und wieso es dennoch im Allgemeinen nicht besser ist $J[f]$ statt $Q[f]$ zur Approximation von $I[f]$ zu benutzen.

Lösungsvorschlag:

a) Die Fassregel lautet

$$J(f) = \frac{b-a}{6}\left(f(a) + 4f\left(\frac{b-a}{2}\right) + f(b)\right).$$

Bei dem vorliegenden Integral mit $a = 0$ und $b = 2\pi$ müssen bei der zweifach zusammengesetzten Fassregel Funktionsauswertungen an 5 Stellen vorgenommen werden ($f(x) = \sqrt{x}\cos(x)$):

$$\begin{aligned} f(0) &= 0 \\ f(\pi/2) &= \sqrt{\pi/2}\cos(\pi/2) = 0 \\ f(\pi) &= -\sqrt{\pi} \\ f\left(\frac{3}{2}\pi\right) &= 0 \\ f(2\pi) &= \sqrt{2\pi} \end{aligned}$$

Es gilt mit der zweifach zusammengesetzten Fassregel:

$$\begin{aligned} \int_0^{2\pi} \sqrt{t}\cos(t)\mathrm{d}t &\approx \frac{\pi}{6}\left(f(0) + 4f(\pi/2) + 2f(\pi) + 4f\left(\frac{3}{2}\pi\right) + f(2\pi)\right) \\ &= \frac{\pi}{6}\left(-2\sqrt{\pi} + \sqrt{2\pi}\right) \end{aligned}$$

b) Es gilt $\alpha, \beta \in \mathbb{R}$ und $x_1 \in \mathbb{R}$ zu finden, so dass $Q[x^k] = I[x^k]$ für $k = 0, 1, 2$. Es ist

$$Q[1] = \alpha + \beta \stackrel{!}{=} I[1] = \int_0^1 (-x^2+1)\mathrm{d}x = \left[-\frac{1}{3}x^3 + x\right]_0^1 = \frac{2}{3}$$
$$Q[x] = \beta x_1 \stackrel{!}{=} I[x] = \int_0^1 (-x^2+1)x\mathrm{d}x = \left[-\frac{1}{4}x^4 + \frac{1}{2}x^2\right]_0^1 = \frac{1}{4}$$
$$Q[x^2] = \beta x_1^2 \stackrel{!}{=} I[x^2] = \int_0^1 (-x^2+1)x^2\mathrm{d}x = \left[-\frac{1}{5}x^5 + \frac{1}{3}x^3\right]_0^1 = \frac{2}{15}$$

Dieses System von Gleichungen kann man lösen indem man die zweite Gleichung mit x_1 multipliziert (ergibt $\frac{1}{4}x_1$) und diese dann mit der dritten Gleichung gleichsetzt ($= \frac{2}{15}$):

$$\beta x_1^2 = \frac{1}{4}x_1 = \frac{2}{15} = \beta x_1^2 \quad \Rightarrow \quad x_1 = \frac{8}{15}$$

Damit erhält man aus der zweiten Gleichung $\beta \frac{8}{15} = \frac{1}{4} \Rightarrow \beta = \frac{15}{32}$. Schlußendlich mit der ersten Gleichung folgt dann $\alpha + \frac{15}{32} = \frac{2}{3} \Rightarrow \alpha = \frac{19}{96}$. Es folgt die Quadraturformel

$$Q[f] = \frac{19}{96} f(0) + \frac{15}{32} f\left(\frac{8}{15}\right).$$

Man weist nun leicht nach, dass

$$Q[x^3] = \frac{15}{32}\left(\frac{8}{15}\right)^3 = \frac{16}{225}$$

und

$$I[x^3] = \int_0^1 (-x^2+1)x^3\mathrm{d}x = \left[-\frac{1}{6}x^6 + \frac{1}{4}x^4\right]_0^1 = \frac{1}{12}$$

nicht gleich sind.

c) Die Unbekannten der Quadraturformel $J[f]$ kann man durch Gleichsetzen von $J[x^k] = I[x^k], k = 3, 4, 5$ bestimmen. Nach Konstruktion folgt dann, dass die Quadraturformel exakt ist für $x^k, k = 3, 4, 5$. Nach Konstruktion ist die Exaktheit für $x^k, k = 0, 1, 2$ aber nicht garantiert, beliebige Polynome aus $\mathbb{P}_5$ werden also nicht exakt integriert.
Vorausgesetzt, eine betrachtete Funktion f ist glatt genug, dann kann sie approximativ durch eine Taylorentwicklung als Polynom geschrieben werden. Die dominierenden Terme niedriger Ordnung werden durch $J[f]$ jedoch im Allgemeinen nicht exakt integriert, die Quadraturformel $Q[f]$ hingegen leistet dies. Es ist daher ratsamer die Quadraturformel $Q[f]$ zu benutzen.

Klausuraufgabe 5: (8+12 Punkte)

a) Bestimmen Sie in der Newtonschen Darstellung das Interpolationspolynom $p(x)$ zu den folgenden Stützpunkten:

i	0	1	2	3	4
x_i	-5	-2	-1	0	1
y_i	17	8	21	42	35

b) Zeigen Sie, dass es zu jeder stetigen Funktion $f \in C[a,b]$ und paarweise verschiedenen Stützstellen $x_0, \dots, x_n \in [a,b]$ für jedes $\varepsilon > 0$ ein interpolierendes Polynom $P(x)$ gibt mit

$$\|P(x) - f(x)\|_\infty \le \varepsilon, \qquad P(x_i) = f(x_i) \quad \text{für } i = 0, \dots, n. \qquad (\star)$$

Hinweis: Ohne Nachweis darf man verwenden, dass es für jedes $\varepsilon > 0$ ein Polynom $Q(x)$ (nicht notwendigerweise interpolierend) gibt, so dass

$$\|Q(x) - f(x)\|_\infty \le \frac{\varepsilon}{M} \qquad \text{mit} \qquad M = 1 + K \cdot (n+1),$$

wobei

$$\max_{k=0,\dots,n} \|L_k\|_\infty =: K < \infty$$

und $L_0, \dots, L_n$ die Lagrangeschen Basispolynome zu den gegebenen Stützstellen $x_0, \dots, x_n$ sind. Es bleibt nun zu zeigen, dass

$$P(x) := Q(x) + \sum_{i=0}^{n} \left(f(x_i) - Q(x_i)\right) L_i(x)$$

die Interpolations- und Genauigkeitseigenschaft $(\star)$ erfüllt.

Lösungsvorschlag:

a) Die Koeffizienten der Newtonpolynome berechnen sich über das Schema der dividierten Differenzen:

$$\begin{array}{r|ccccccccc}
-5 & \boxed{17} & & & & & & & & \\
 & & \searrow & & & & & & & \\
-2 & 8 & \to & \frac{8-17}{-2-(-5)} = \boxed{-3} & & & & & & \\
 & & \searrow & & \searrow & & & & & \\
-1 & 21 & \to & \frac{21-8}{-1-(-2)} = 13 & \to & \frac{13-(-3)}{-1-(-5)} = \boxed{4} & & & & \\
 & & \searrow & & \searrow & & \searrow & & & \\
0 & 42 & \to & \frac{42-21}{0-(-1)} = 21 & \to & \frac{21-13}{0-(-2)} = 4 & \to & \frac{4-4}{0-(-5)} = \boxed{0} & & \\
 & & \searrow & & \searrow & & \searrow & & \searrow & \\
1 & 35 & \to & \frac{35-42)}{1-0} = -7 & \to & \frac{-7-21}{1-(-1)} = -14 & \to & \frac{-14-4}{1-(-2)} = -6 & \to & \frac{-6-0}{1-(-5)} = \boxed{-1}
\end{array}$$

Die Newtonpolynome lauten:

$$\begin{aligned}
w_0(x) &= 1 \\
w_1(x) &= x+5 \\
w_2(x) &= (x+5)(x+2) = x^2+7x+10 \\
w_3(x) &= (x+5)(x+2)(x+1) = x^3+8x^2+17x+10 \\
w_4(x) &= (x+5)(x+2)(x+1)(x-0) = x^4+8x^3+17x^2+10x,
\end{aligned}$$

womit das Interpolationspolynom berechnet werden kann zu:

$$\begin{aligned}
p(x) &= 17w_0(x) - 3w_1(x) + 4w_2(x) + 0w_3(x) - w_4(x) \\
&= 17 - 3(x+5) + 4(x^2+7x+10) - (x^4+8x^3+17x^2+10x).
\end{aligned}$$

b) Für das Polynom $P(x) := Q(x) + \sum_{k=0}^{n} \left(f(x_k) - Q(x_k)\right) L_k(x)$ müssen zwei Eigenschaften nachgewiesen werden. Zuerst zeigen wir die Interpolationseigenschaft, anschliessend die Genauigkeitseigenschaft:

(i) zu zeigen: $P(x_j) = f(x_j)$ für $j = 0, \dots, n$
Es sei x_j ein beliebiger Punkt der Stützstellen. Dann gilt

$$\begin{aligned}
P(x_j) &= Q(x_j) + \sum_{k=0}^{n} \left(f(x_k) - Q(x_k)\right) \underbrace{L_k(x_j)}_{=\delta_{jk}} \\
&= Q(x_j) + f(x_j) - Q(x_j) \\
&= f(x_j)
\end{aligned}$$

(ii) zu zeigen: $\|P(x) - f(x)\|_\infty \le \varepsilon$
Wir benutzen lediglich die Dreiecksungleichung und die Voraussetzung.

$$\begin{aligned}
\|P(x) - f(x)\|_\infty &= \left\| Q(x) + \sum_{k=0}^{n} \left(f(x_k) - Q(x_k)\right) L_k(x) - f(x) \right\|_\infty \\
&\le \|Q(x) - f(x)\|_\infty + \sum_{k=0}^{n} \| \left[f(x_k) - Q(x_k)\right] L_k(x) \|_\infty \\
&\le \frac{\varepsilon}{1+K(n+1)} + \sum_{k=0}^{n} K \, \|f(x_k) - Q(x_k)\|_\infty \\
&\le \frac{\varepsilon}{1+K(n+1)} + (n+1)K \frac{\varepsilon}{1+K(n+1)} \\
&= \frac{\varepsilon(1+(n+1)K)}{1+K(n+1)} \\
&= \varepsilon
\end{aligned}$$