

# Modulprüfung MA2302 „Numerik“

Musterlösung

10. August 2009

Prüfer: Prof. Dr. Bernd Simeon



## Aufgabe 1 Eine 1-Punkt Quadraturformel (ca. 9 P.)

Es seien für  $f \in C^\infty$

$$I(f) = \int_0^1 f(x) \cdot \sqrt{x} \, dx \quad \text{und} \quad J(f) = g \cdot f(s).$$

Das gewichtete Integral  $I(f)$  soll durch die Quadraturformel  $J(f)$ , welche mit nur einer Stützstelle  $s$  auskommt, möglichst gut approximiert werden.

- Für Polynome welchen Grades kann diese Quadraturformel maximal exakt sein? Wie müssen dazu das Gewicht  $g$  und die Stützstelle  $s$  gewählt werden?
- Leiten Sie eine Fehlerformel der Form

$$I(f) - J(f) = C \cdot f^{(p)}(\xi)$$

her. *Hinweis:* Taylorentwicklung von  $f$  um geeignete Stelle.

Im Folgenden wird speziell das Integral

$$\int_0^1 \cos\left(\frac{\pi}{2}x\right) \cdot \sqrt{x} \, dx \approx 0.35762 \dots \quad (1)$$

betrachtet, d.h.  $f(x) = \cos\left(\frac{\pi}{2}x\right)$ .

- Approximieren Sie das Integral (1) mit Hilfe der Trapezregel und geben Sie den Fehler an.
- Welchen Fehler können Sie nach Teilaufgabe b) für das Integral (1) erwarten, wenn man es mit der Quadraturformel  $J$  approximiert?
- Ist im vorliegenden Fall die Approximation durch  $J$  im Vergleich zur Trapezregel tatsächlich genauer? Begründen Sie Ihre Aussage durch Zahlwerte.

## Lösung Aufgabe 1

- Nachdem  $J$  nur zwei Koeffizienten  $g$  und  $s$  hat, kann die QF höchstens Polynome ersten Grades exakt integrieren. Damit die Monome  $m_0(x) = 1$  und  $m_1(x) = x$  exakt integriert werden, lauten die Bedingungen

$$I(m_0) = J(m_0) \quad \text{und} \quad I(m_1) = J(m_1).$$

Wir haben

$$I(m_0) = \int_0^1 1 \cdot \sqrt{x} \, dx = \left[ \frac{2}{3} x^{\frac{3}{2}} \right]_0^1 = \frac{2}{3}$$

$$J(m_0) = g \cdot m_0(s) = g$$

$$I(m_1) = \int_0^1 x \cdot \sqrt{x} \, dx = \left[ \frac{2}{5} x^{\frac{5}{2}} \right]_0^1 = \frac{2}{5}$$

$$J(m_1) = g \cdot m_1(s) = g \cdot s .$$

Die Bedingungen sind durch

$$g = \frac{2}{3} \quad \text{und} \quad s = \frac{3}{5}$$

erfüllt. Bemerkung: die Quadraturformel  $J$  hat Ordnung 2.

b) Taylorentwicklung von  $f$  um  $s$  liefert:

$$f(x) = f(s + (x - s)) = \underbrace{f(s) + f'(s)(x - s)}_{=:p_1(x)} + \underbrace{\frac{f''(\hat{\xi})}{2}(x - s)^2}_{=:r(x)} .$$

Weil erstens  $J$  linear ist und zweitens  $J$  exakt für Polynome vom Grad 1 ist und drittens der Rest  $r$  die Nullstelle  $s$  hat, folgt

$$J(f) = J(p_1) + J(r) = I(p_1) + 0 = I(p_1) .$$

Weil  $I$  auch linear ist gilt

$$I(f) - J(f) = I(p_1) + I(r) - J(f) = I(r) .$$

Also ist der Fehler gegeben durch  $I(r)$ , was unter Verwendung des Mittelwertsatzes der Integralrechnung ausformuliert zu

$$\begin{aligned} I(r) &= \int_0^1 \frac{f''(\hat{\xi})}{2} (x - s)^2 \cdot \sqrt{x} \, dx = \frac{f''(\xi)}{2} \int_0^1 x^2 \cdot \sqrt{x} - 2sx \cdot \sqrt{x} + s^2 \cdot \sqrt{x} \, dx \\ &= \frac{f''(\xi)}{2} \left( \left[ \frac{2}{7} x^{\frac{7}{2}} \right]_0^1 - \frac{6}{5} \cdot \frac{2}{5} + \frac{9}{25} \cdot \frac{2}{3} \right) = \frac{4}{175} f''(\xi) . \end{aligned}$$

c) Bezeichne  $T$  die Trapezregel. Sie ist hier anzuwenden auf  $g(x) = f(x) \cdot \sqrt{x}$  mit

$$T(g) = \frac{1}{2}(f(0)\sqrt{0} + f(1)\sqrt{1}) = \frac{1}{2} \cos\left(\frac{\pi}{2}\right) = 0 .$$

Der absolute Fehler ist damit

$$|I(f) - T(g)| = 0.35762 \dots .$$

d) Nach Teilaufgabe b) können wir für den absoluten Fehler erwarten:

$$|I(f) - J(f)| \leq \frac{4}{175} \|f''\|_{\infty} = \frac{4}{175} \frac{\pi^2}{4} < \frac{(3.5)^2}{7 \cdot 5^2} = 0.07$$

- e) Nachdem die obere Fehlerschranke in Teilaufgabe d) echt kleiner als der Fehler in Teilaufgabe c) ist, ist  $J$  genauer.

**Aufgabe 2** *Eigenwerte (ca. 12 P.)*

Es sei

$$A = \begin{pmatrix} 5 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

- a) Begründen Sie kurz, dass  $A$  nur reelle Eigenwerte hat.  
 b) Geben Sie ein kompaktes Intervall an, welches alle Eigenwerte von  $A$  enthält. Begründen Sie Ihr Resultat.  
 c) Sei  $\mu \in \mathbb{R}$  beliebig. Welchen Rang hat die Matrix  $A - \mu I$  mindestens? Was folgt daraus für die geometrische Vielfachheit der Eigenwerte?

**Bitte wenden** →

- d) Die Matrix  $A - 3I$  hat eine Zerlegung der Form

$$A - 3I = L \cdot D \cdot L^T.$$

Dabei ist  $L$  eine untere Dreiecksmatrix mit positiven Diagonalelementen und  $D$  eine Diagonalmatrix mit Diagonalelementen aus  $\{-1, 0, 1\}$ .

Berechnen Sie die Matrix  $D$ .

- e) Wieviele Eigenwerte von  $A$  sind kleiner oder gleich 3? Beantworten Sie die Frage ohne die Eigenwerte zu berechnen, aber begründen Sie Ihre Antwort.

**Lösung Aufgabe 2** *Eigenwerte*

- a) Weil  $A$  symmetrisch sind die Eigenwerte alle reell.  
 b) Die Eigenwerte liegen bekanntlich in der Vereinigung der Gerschgorinkreise  $K_1, \dots, K_4 \subset \mathbb{C}$  mit

$$\begin{array}{ll} K_1 = \{z \in \mathbb{C} : |5 - z| \leq 1\} & K_1 \cap \mathbb{R} = [4, 6] \\ K_2 = \{z \in \mathbb{C} : |-1 - z| \leq 2\} & K_2 \cap \mathbb{R} = [-3, 1] \\ K_3 = \{z \in \mathbb{C} : |2 - z| \leq 2\} & K_3 \cap \mathbb{R} = [0, 4] \\ K_4 = \{z \in \mathbb{C} : |3 - z| \leq 1\} & K_4 \cap \mathbb{R} = [2, 4] \end{array}$$

Nach a) sind alle Eigenwerte reell und man bekommt ein kompaktes Intervall, das alle Eigenwerte enthält, z.B. so

$$I = \bigcup_{i=1}^4 (K_i \cap \mathbb{R}) = [-3, 6]$$

- c) Sei  $\mu \in \mathbb{R}$  beliebig, dann hat die Matrix  $A - \mu I$  im Vergleich zu  $A$  nur andere Diagonaleinträge. Bei  $A - \mu I$ , wie bei  $A$ , sind die ersten 3 Spalten linear unabhängig und somit ist der Rang  $A - \mu I$  mindestens 3. Aus der Dimensionsformel folgt weiterhin, dass der Kern von  $A - \mu I$  höchstens die Dimension 1 hat.

Wenn nun  $\mu = \lambda$  ein Eigenwert ist, muss die Dimension des Kerns von  $A - \lambda I$  mindestens 1 sein. Beide Aussagen zusammen ergeben, dass Kern von  $A - \lambda I$  genau die Dimension 1 hat, also die geometrische Vielfachheit des Eigenwertes  $\lambda$  gleich 1 ist.

- d) Der Gauß-Algorithmus ohne Pivoting liefert (hier mit in situ Speicherung) für  $A - 3I$ :

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & -4 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{9}{2} & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{9}{2} & 1 & 0 \\ 0 & -\frac{2}{9} & -\frac{7}{9} & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{9}{2} & 1 & 0 \\ 0 & -\frac{2}{9} & -\frac{7}{9} & 1 \\ 0 & 0 & -\frac{9}{7} & \frac{9}{7} \end{pmatrix}$$

Damit ist

$$A - 3I = L_G \cdot R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{9} & 1 & \\ 0 & 0 & -\frac{9}{7} & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & -\frac{9}{2} & 1 & 0 \\ 0 & 0 & -\frac{7}{9} & 1 \\ 0 & 0 & 0 & \frac{9}{7} \end{pmatrix},$$

aus Symmetriegründen erhält man sofort  $R = \hat{D} \cdot L_G^T$

$$A - 3I = L_G \cdot \hat{D} \cdot L_G^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{9} & 1 & \\ 0 & 0 & -\frac{9}{7} & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -\frac{9}{2} & 0 & 0 \\ 0 & 0 & -\frac{7}{9} & 0 \\ 0 & 0 & 0 & \frac{9}{7} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{9} & 1 & \\ 0 & 0 & -\frac{9}{7} & 1 \end{pmatrix}^T,$$

Die gesuchte Matrix hat nun die Form

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Man erhält sie, indem man  $D = \text{sign}(\hat{D})$  und  $\sqrt{|\hat{D}|}$  komponentenweise bildet. Der Zusammenhang ist

$$\hat{D} = \sqrt{|\hat{D}|} \cdot D \cdot \sqrt{|\hat{D}|} \quad , \quad L = L_G \cdot \sqrt{|\hat{D}|} \quad \Rightarrow \quad A = L \cdot D \cdot L^T,$$

dann ist  $L$  sicher eine untere Dreiecksmatrix mit positiven Diagonalelementen.

- e) Wenn die Matrix  $A$  das Spektrum  $\sigma(A)$  hat, so gilt

$$\sigma(A - 3I) = \sigma(A) - 3.$$

Die Frage ist also gleichbedeutend mit "Wieviele nicht-positive Eigenwerte hat  $A - 3I$  ?". Mit  $A - 3I = LDL^T$  aus Teilaufgabe d) haben wir eine Kongruenz zwischen  $A - 3I$  und  $D$ . Kongruenz bedeutet zwar im Allgemeinen nicht, dass man gleiche Eigenwerte hat, aber es bleiben die Vorzeichen mit entsprechender Vielfachheit erhalten. Man liest daher aus  $D$  ab, dass  $A - 3I$  genau zwei nicht-positive Eigenwerte hat und damit hat  $A$  genau zwei Eigenwerte kleiner oder gleich 3.

### Aufgabe 3 *Williamsons Runge–Kutta Verfahren (ca. 11 P.)*

Betrachten Sie das Runge–Kutta Verfahren zu dem folgenden Butcher-Tableau:

$c_1$			
$c_2$	$\frac{1}{3}$		
$c_3$	$-\frac{3}{16}$	$\frac{15}{16}$	
	$\frac{1}{6}$	$\frac{3}{10}$	$\frac{8}{15}$

a) Das gegebene Verfahren ist anwendbar auf autonome Anfangswertprobleme. Erweitern Sie es nun *ohne Ordnungsverlust* so, dass es auch für nicht autonome Differentialgleichungen anwendbar ist, d.h. berechnen Sie die Koeffizienten  $c_i$ . Geben Sie nicht nur das Ergebnis sondern auch die bestimmenden Gleichungen an. Begründen sie außerdem genau, warum dabei die Ordnung erhalten bleibt.

b) Zeigen Sie, dass das Verfahren mindestens die Ordnung 2 besitzt<sup>1</sup>.

Williamson<sup>2</sup> stellte das folgende Verfahren zur Lösung einer autonomen gewöhnlichen Differentialgleichung  $y' = f(y)$  vor:

$$\begin{aligned}q_1 &= hf(y_0) \\z_1 &= y_0 + \alpha_1 q_1 \\q_2 &= hf(z_1) + \beta_1 q_1 \\z_2 &= z_1 + \alpha_2 q_2 \\q_3 &= hf(z_2) + \beta_2 q_2 \\y_1 &= z_2 + \alpha_3 q_3.\end{aligned}$$

Wie üblich bezeichnet  $h$  die Schrittweite,  $y_0$  die (approximierte) Lösung zu einem Punkt  $x_0$  und  $y_1$  die (approximierte) Lösung zum Punkt  $x_0 + h$ .

c) Zeigen Sie, dass es sich auch hier um ein explizites Runge–Kutta Verfahren handelt.

d) Geben Sie das Butcher–Tableau zu Williamsons Verfahren für den autonomen Fall an.

e) Mit

$$\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{15}{16}, \alpha_3 = \frac{8}{15}, \quad \beta_1 = -\frac{5}{9}, \beta_2 = -\frac{153}{128}$$

ist Williamsons Verfahren äquivalent zu dem Verfahren aus den Teilaufgaben a) und b) (nicht zu zeigen). Williamson entwickelte sein Verfahren für hochdimensionale Probleme, d.h.  $y \in \mathbb{R}^n$  mit sehr großem  $n$ . Wieso ist für solche Systeme das Williamson-Schema im Unterschied zum üblichen Runge–Kutta Schema besonders vorteilhaft? (Kurze Begründung reicht).

### Lösung Aufgabe 3 *Williamsons Runge–Kutta Verfahren*

a) Wenn man die  $c_i$  gemäß der Autonomisierungsinvarianz setzt, d.h.

$$c_i = \sum_{j=1}^n a_{ij} \quad \Rightarrow \quad c_1 = 0, \quad c_2 = \frac{1}{3}, \quad c_3 = \frac{3}{4},$$

---

<sup>1</sup>tatsächlich ist es ein Verfahren dritter Ordnung, das ist aber nicht zu zeigen

<sup>2</sup>J. H. Williamson, J. Comput. Phys., 35:48–56, 1980

dann ist das Verfahren ohne Ordnungsverlust auf nicht autonome Systeme anwendbar.  
Begründung : Sei das nicht autonome AWP

$$y'(x) = f(x, y(x)) \qquad y(x_0) = y_0$$

gegeben. Es kann autonomisiert werden durch

$$\begin{aligned} s'(x) &= 1 & s(x_0) &= x_0 \\ y'(x) &= f(s(x), y(x)) & y(x_0) &= y_0 . \end{aligned}$$

Für das autonomisierte System arbeitet das Verfahren mit einer gewissen Ordnung  $p$ . Die Stufen  $K_i$  sehen wie nachstehend aus, wobei  $k_i$  die Stufen für den nicht autonomen Fall bezeichnen

$$K_i = \left( f \left( \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + h \sum_{j=1}^{i-1} a_{ij} K_j \right) \right) = \begin{pmatrix} 1 \\ k_i \end{pmatrix} = \left( f \begin{pmatrix} x_0 + h \sum_{j=1}^{i-1} a_{ij} 1 \\ y_0 + h \sum_{j=1}^{i-1} a_{ij} k_j \end{pmatrix} \right) .$$

Wenn man nun

$$c_i = \sum_{j=0}^n a_{ij}$$

setzt, kann man sich den Umweg über die Autonomisierung sparen. Das Verfahren leistet dann exakt das gleiche und daher auch mit der selben Ordnung.

- b) Weil das Verfahren mit Teilaufgabe c) autonomisierungsinvariant ist braucht man, um Ordnung 2 nachzuweisen, nur noch die zwei Bedingungen

$$\begin{aligned} b_1 + b_2 + b_3 &= 1 \\ b_1 c_1 + b_2 c_2 + b_3 c_3 &= \frac{1}{2} . \end{aligned}$$

zu prüfen. Sie sind erfüllt !

Bemerkung: Wegen Autonomisierungsinvarianz, hätte man, um Ordnung 3 nachzuweisen, außerdem noch die Bedingungen

$$\begin{aligned} b_1 c_1^2 + b_2 c_2^2 + b_3 c_3^2 &= \frac{1}{3} \\ b_3 a_{32} c_2 &= \frac{1}{6} , \end{aligned}$$

die auch noch erfüllt sind. Das war aber nicht gefragt !

- c) Man setzt von unten nach oben ein

$$\begin{aligned} y_1 &= z_2 + \alpha_3 q_3 \\ &= z_1 + \alpha_2 q_2 + \alpha_3 (f(z_2)h + \beta_2 q_2) \\ &= z_1 + (\alpha_2 + \beta_2 \alpha_3) q_2 + \alpha_3 f(z_1 + \alpha_2 q_2) h \\ &= y_0 + \alpha_1 q_1 + (\alpha_2 + \beta_2 \alpha_3) q_2 + \alpha_3 f(y_0 + \alpha_1 q_1 + \alpha_2 q_2) h \\ &= y_0 + \alpha_1 q_1 + (\alpha_2 + \beta_2 \alpha_3) (f(z_1)h + \beta_1 q_1) + \alpha_3 f(y_0 + \alpha_1 q_1 + \alpha_2 (f(z_1)h + \beta_1 q_1)) h \\ &= y_0 + (\alpha_1 + \beta_1 \alpha_2 + \beta_1 \beta_2 \alpha_3) q_1 + (\alpha_2 + \beta_2 \alpha_3) f(z_1) h + \alpha_3 f(y_0 + (\alpha_1 + \beta_1 \alpha_2) q_1 + \alpha_2 f(z_1) h) h \\ &= y_0 + (\alpha_1 + \beta_1 \alpha_2 + \beta_1 \beta_2 \alpha_3) q_1 + (\alpha_2 + \beta_2 \alpha_3) f(y_0 + \alpha_1 q_1) h \\ &\quad + \alpha_3 f(y_0 + (\alpha_1 + \beta_1 \alpha_2) q_1 + \alpha_2 f(y_0 + \alpha_1 q_1) h) h \\ &= y_0 + (\alpha_1 + \beta_1 \alpha_2 + \beta_1 \beta_2 \alpha_3) f(y_0) h + (\alpha_2 + \beta_2 \alpha_3) f(y_0 + \alpha_1 f(y_0) h) h \\ &\quad + \alpha_3 f(y_0 + (\alpha_1 + \beta_1 \alpha_2) f(y_0) h + \alpha_2 f(y_0 + \alpha_1 f(y_0) h) h) h . \end{aligned}$$

Setzt man nun

$$\begin{aligned} K_1 &= f(y_0) \\ K_2 &= f(y_0 + \alpha_1 K_1 h) \\ K_3 &= f(y_0 + (\alpha_1 + \beta_1 \alpha_2) K_1 h + \alpha_2 K_2 h) , \end{aligned}$$

so schreibt sich  $y_1$  als:

$$y_1 = y_0 + (\alpha_1 + \beta_1 \alpha_2 + \beta_1 \beta_2 \alpha_3) K_1 h + (\alpha_2 + \beta_2 \alpha_3) K_2 h + \alpha_3 K_3 h .$$

d) Gemäß obiger Form überträgt man die Koeffizienten in ein Butcher-Tableau, allgemein

$$\begin{array}{c|ccc} c_1 & 0 & 0 & 0 \\ c_2 & \alpha_1 & 0 & 0 \\ c_3 & \alpha_1 + \beta_1 \alpha_2 & \alpha_2 & 0 \\ \hline & \alpha_1 + \beta_1 \alpha_2 + \beta_1 \beta_2 \alpha_3 & \alpha_2 + \beta_2 \alpha_3 & \alpha_3 \end{array}$$

e) Wenn man sich das Williamson-Schema des Verfahrens scharf ansieht, fällt einem auf, dass man für alle drei Stufen der  $q_i$  und alle zwei Stufen der  $z_i$  jeweils den gleichen Speicherplatz verwenden kann. Es reicht also z.B. in Matlab zwei Hilfsvektoren  $q$  und  $z$  anzulegen, welche in einem Verfahrensschritt sukzessiv gemäß des Williamson-Schema überschrieben werden. Für das gleiche Verfahren im RK-Schema braucht man drei Hilfsvektoren  $K_1, K_2, K_3$  für die Stufen. Man spart sich einen Vektor der Länge  $n =$  Systemdimension, was bei sehr großem  $n$  relevant ist.

#### Aufgabe 4 (8 P.)

*Bepunktung dieser Aufgabe:*

*Für jede der folgenden Multiple-Choice Fragen wird die richtige Lösung mit +2 Punkten bewertet, die falsche Lösung mit -2 Punkten und keine Antwort mit 0 Punkten. Die vergebenen positiven oder negativen Punkte werden zu einer Gesamtpunktzahl der Aufgabe summiert. Falls die Summe negativ ist wird die Aufgabe mit 0 Punkten bewertet. Die Gesamtpunktzahl der Aufgabe kann also nicht negativ werden.*

Beantworten Sie die folgenden Fragen:

Hessenberg-Matrizen sind günstig für numerische Eigenwertlöser  Ja  Nein

Jedes konsistente und stabile Mehrschrittverfahren ist konvergent.  Ja  Nein

Es gibt eine Folge  $(p_n)$  von Polynomen  $p_n$ , wobei  $p_n$  Grad  $n$  besitzt, so dass

$$\int_0^\pi p_n(x) p_{2n}(x) dx = 0$$

für alle  $n > 1$  gilt.  Ja  Nein

Betrachten Sie das Runge-Kutta Verfahren zu dem Butcher-Tableau

$$\begin{array}{c|cc} 0 & & \\ \frac{1}{2} & \frac{1}{2} & \\ \hline & 0 & 1 \end{array}$$

Es sei  $\alpha, \beta, x_0, y_0 \in \mathbb{R}$  beliebig. Wenn das obige Verfahren auf

$$y' = \alpha x + \beta, \quad y(x_0) = y_0$$

angewendet wird, dann gilt

$$y_i = y(x_i)$$

für jede Schrittweite.  Ja  Nein