

Klausur zur Vordiplom-Prüfung

Numerische Verfahren

17. Juli 2007

Sie haben **90** Minuten Zeit zum Bearbeiten der Klausur.

Bitte kennzeichnen Sie jedes Blatt mit Ihrem Namen und Ihrer Matrikelnummer in DRUCKSCHRIFT.

Tragen Sie bitte zunächst Ihren Namen, Ihren Vornamen und Ihre Matrikelnummer in **Druckschrift** in die folgenden jeweils dafür vorgesehenen Felder ein.

Diese Eintragungen werden auf Datenträger gespeichert.

Name:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Vorname:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Matr.-Nr.	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Ich bin darüber belehrt worden, dass meine Ausarbeitung nur dann als Prüfungsleistung bewertet wird, wenn die Nachprüfung durch das Zentrale Prüfungsamt der TUHH meine offizielle Zulassung vor Beginn der Prüfung ergibt.

(Unterschrift)

Lösen Sie die folgenden 12 Aufgaben!

Aufgabe	1	2	3	4	5	6	7	8	9	10	11	12
Punkte												

Σ	
----------	--

Aufgabe 1: (4+4 Punkte)

Was versteht man unter dem Begriff Auslöschung? Geben Sie zwei Zahlen mit einem relativen Fehler von ungefähr 10^{-3} an, deren Differenz einen relativen Fehler von ungefähr 2 hat.

Lösung zu Aufgabe 1: Wenn zwei annähernd gleich große (fehlerbehaftete) Zahlen mit gleichem Vorzeichen subtrahiert oder mit entgegengesetztem Vorzeichen addiert werden, spricht man von Auslöschung. Bei dieser Auslöschung werden die gleichen führenden Stellen zu Null und das Ergebnis der Operation ist viel kleiner als die Eingangsdaten. Das führt dazu, dass die fehlerhaften Stellen viel näher an die führende erste Stelle ungleich Null rutschen, also der relative Fehler des Ergebnisses der Operation ungleich viel größer ist als die relativen Fehler der Ausgangsdaten. (4 Punkte)

Als Beispiel nehmen wir die exakten Zahlen $x = 1.001$ und $y = 1$ und die gestörten Zahlen $\tilde{x} = 1.002$ und $\tilde{y} = 0.999$. Die relativen Fehler sind damit gegeben durch

$$\frac{|\tilde{x} - x|}{|x|} = \frac{|1.002 - 1.001|}{1.001} = \frac{0.001}{1.001} = \overline{0.000999} \approx 10^{-3},$$

$$\frac{|\tilde{y} - y|}{|y|} = \frac{|0.999 - 1|}{1} = \frac{0.001}{1} = 10^{-3},$$

also beide kleiner oder gleich 10^{-3} . Das tatsächliche Ergebnis d wäre

$$d = x - y = 1.001 - 1 = 0.001$$

gewesen, die exakte Differenz der gestörten Daten liefert das gestörte Ergebnis

$$\tilde{d} = \tilde{x} - \tilde{y} = 1.002 - 0.999 = 0.003.$$

Der relative Fehler ist also durch

$$\frac{|\tilde{d} - d|}{|d|} = \frac{|0.003 - 0.001|}{0.001} = \frac{3 - 1}{1} = 2$$

gegeben und damit exakt zwei. (4 Punkte)

Aufgabe 2: (5+5 Punkte)

Geben Sie die Lagrangsche Form der Polynominterpolation wieder. Berechnen Sie das Interpolationspolynom der Funktion

$$f(x) = \cos(\pi x)$$

zu den Punkten

$$\begin{array}{c|cccccc} i & 0 & 1 & 2 & 3 & 4 \\ \hline x_i & -2 & -1 & 0 & 1 & 2 \end{array}.$$

Lösung zu Aufgabe 2: Die Lagrangsche Interpolation zu den Datenpaaren $\{(x_j, y_j)\}_{j=0}^n$ mit $x_i \neq x_j$ für alle $i \neq j$ ist eine Vorschrift um das eindeutige Polynom p vom Höchstgrad n zu berechnen, dass die Interpolationsbedingungen

$$p(x_j) = y_j, \quad j = 0, 1, \dots, n$$

erfüllt. Es läßt sich leicht explizit angeben als

$$p(x) = \sum_{j=0}^n y_j \ell_j(x),$$

wobei die Lagrangeschen Basispolynome gegeben sind durch die Interpolationspolynome zu den speziellen Aufgaben

$$\ell_j(x_i) = \delta_{ij}, \quad i = 0, 1, \dots, n, \quad j = 0, 1, \dots, n,$$

und explizit angegeben lauten

$$\ell_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^n (x - x_i) \bigg/ \prod_{\substack{i=0 \\ i \neq j}}^n (x_j - x_i), \quad j = 0, \dots, n. \quad (5 \text{ Punkte})$$

Die Interpolationsaufgabe läßt sich auffassen als die Interpolation der Datenpaare

$$\begin{array}{c|ccccc} x_i & -2 & -1 & 0 & 1 & 2 \\ \hline y_i & 1 & -1 & 1 & -1 & 1 \end{array}.$$

Das Interpolationspolynom p_4 hat sicherlich den Grad vier und aufgrund der Symmetrie der Daten bezüglich der y -Achse werden alle Koeffizienten der ungeraden Monome gleich Null sein. Da der konstante Koeffizient gleich eins ist, da ja $p_4(0) = 1$ gilt, sind nur noch zwei Koeffizienten

$$p_4(x) = 1 + ax^2 + bx^4$$

zu bestimmen. Dieses kann schnell über ein lineares Gleichungssystem mit einer 2×2 Matrix geschehen, z.B. muss ja gelten

$$\begin{aligned} p_4(1) &= 1 + 1 \cdot a + 1 \cdot b = -1, \\ p_4(2) &= 1 + 4 \cdot a + 16 \cdot b = 1, \end{aligned}$$

also

$$\begin{pmatrix} 1 & 1 \\ 4 & 16 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

mit der Lösung

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -8 \\ 2 \end{pmatrix}.$$

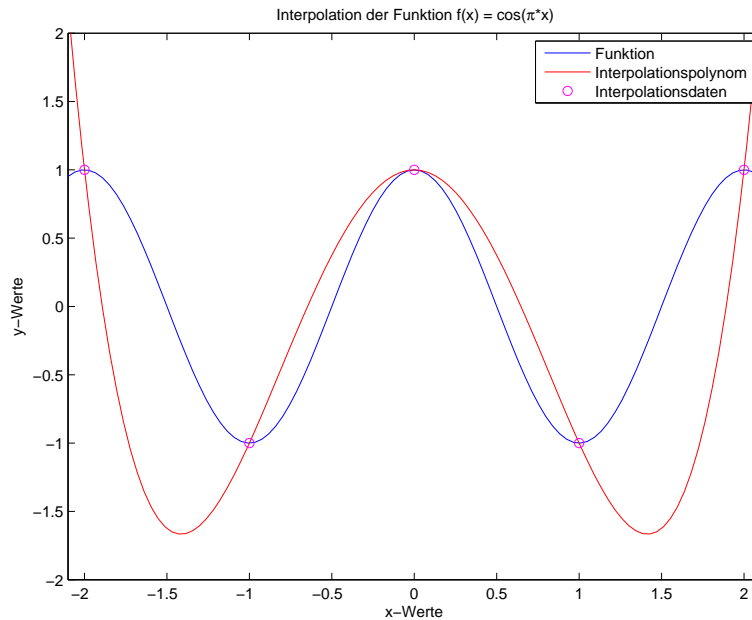


Abbildung 1: Graph der Funktion und des Interpolationspolynomes.

Damit hat p_4 die Gestalt aus Abbildung 1 und ist gegeben durch

$$p_4(x) = 1 - \frac{8}{3}x^2 + \frac{2}{3}x^4. \quad (5 \text{ Punkte})$$

Natürlich konnte man p_4 auch mittels der Lagrangeschen oder der Newtonschen Form der Polynominterpolation berechnen.

Aufgabe 3: (6 Punkte)

Was ist der Vorteil der Newtonschen Form der Polynominterpolation gegenüber der Lagrangeschen Form der Polynominterpolation?

Lösung zu Aufgabe 3: Bei der Lagrangeschen Form der Interpolation muss bei der Hinzunahme eines zusätzlichen Datenpaares jedes Basispolynom neu berechnet werden. Die Newtonsche Form der Polynominterpolation vermeidet dieses durch eine Aufdatierung des bereits bestehenden Interpolationspolynomes. Überdies ist die Newtonsche Form der Polynominterpolation weniger aufwändig als die Lagrangesche. (6 Punkte)

Aufgabe 4: (4+5 Punkte)

Was ist eine Quadraturformel? Wie ist der genaue Zusammenhang der im Skript behandelten Quadraturformeln mit der Interpolation?

Lösung zu Aufgabe 4: Eine Quadraturformel ist eine Formel zur numerischen

Annäherung eines Integrales, welche nur auf Funktionsauswertungen der Funktion unter dem Integral basiert und allgemein die Gestalt

$$\int_0^1 f(x) dx \approx \sum_{j=0}^n \omega_j f(x_j)$$

hat. Die Angabe der Auswertungsstellen, den Stützstellen $\{x_j\}_{j=0}^n$ und der sogenannten Gewichte $\{\omega_j\}_{j=0}^n$ bestimmt in eindeutiger Weise eine Quadraturformel. (4 Punkte)

Die im Skript angegebenen Quadraturformeln sind alle interpolatorisch, d.h., sie lassen sich (re)konstruieren bei gegebenen Stützstellen, da sie auf der Idee basieren, die (schwierige) Integration einer beliebigen Funktion durch die (einfache) Integration eines zugeordneten Interpolationspolynomes zu ersetzen. Die Quadraturformeln erhält man bei gegebenen Stützstellen $\{x_j\}_{j=0}^n$ gemäß

$$\int_0^1 f(x) dx \approx \int_0^1 p_n(x) dx = \int_0^1 \sum_{j=0}^n f(x_j) \ell_j(x) dx = \sum_{j=0}^n f(x_j) \int_0^1 \ell_j(x) dx,$$

und die Gewichte lassen sich in eindeutiger Weise aus der Integration der Lagrangeschen Basispolynome berechnen als

$$\omega_j = \int_0^1 \ell_j(x) dx. \quad (5 \text{ Punkte})$$

Aufgabe 5: (5+5 Punkte)

Was ist eine QR-Zerlegung? Welcher Varianten der QR-Zerlegung bezüglich des nötigen Speicherplatzaufwandes kennen Sie?

Lösung zu Aufgabe 5: Eine QR-Zerlegung ist eine Zerlegung einer gegebenen Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ (meist mit $m \geq n$) in ein Produkt einer orthogonalen oder orthonormalen Matrix und einer oberen Dreiecksmatrix. (5 Punkte)

Es gibt im Wesentlichen zwei Zerlegungen dieser Art, hier bezeichnet mit

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1 = \mathbf{Q}_2 \mathbf{R}_2,$$

wobei in der ersten Variante $\mathbf{Q}_1 \in \mathbb{R}^{m \times m}$ orthogonal (und damit quadratisch) ist, also

$$\mathbf{Q}_1^T \mathbf{Q}_1 = \mathbf{E}_m$$

gilt, und $\mathbf{R}_1 \in \mathbb{R}^{m \times n}$ eine (rechteckige) obere Dreiecksmatrix ist, also eine Matrix, deren Elemente r_{ij} gleich Null sind für alle in Frage kommenden $j > i$.

In der zweiten Variante hat die rechteckige Matrix $\mathbf{Q}_2 \in \mathbb{R}^{m \times n}$ orthonormale Spalten und besteht aus den ersten n Spalten der oben genannten Matrix \mathbf{Q}_1 , es gilt also

$$\mathbf{Q}_2^T \mathbf{Q}_2 = \mathbf{E}_n$$

Die Matrix $\mathbf{R}_2 \in \mathbb{R}^{n \times n}$ ist quadratisch und enthält nur den oberen Block der Matrix \mathbf{R}_1 . (5 Punkte)

Die zweite Variante der QR-Zerlegung bekommt man leicht über das (klassische oder modifizierte) Gram-Schmidt-Verfahren, die erste Variante resultiert (implizit) aus der QR-Zerlegung mit Householder-Reflektoren nach Erweiterung der einzelnen Reflektoren auf den Gesamtraum und Produktbildung oder Givens-Rotationen mit einer ähnlichen Erweiterung und Produktbildung. Selbstverständlich werden diese Produkte normalerweise niemals gebildet.

Aufgabe 6: (2+4+4+6 Punkte)

Berechnen Sie die QR-Zerlegungen (unter Angabe des resultierenden \mathbf{Q} und \mathbf{R}) der folgenden Matrizen:

$$\mathbf{A} = (42), \quad \mathbf{B} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 5 & -3 \\ 0 & 2 \\ 0 & 1 \end{pmatrix}.$$

Lösung zu Aufgabe 6: Die Matrix \mathbf{A} enthält nur eine Spalte, welche aber noch nicht normiert ist. Die entsprechende Spalte von \mathbf{Q} ist die normierte Spalte, in \mathbf{R} steht in der Diagonale die Normierungskonstante,

$$\mathbf{A} = (42) = \underbrace{\begin{pmatrix} 1 \\ 42 \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} 42 \end{pmatrix}}_{\mathbf{R}}. \quad (2 \text{ Punkte})$$

Auch die Matrix \mathbf{B} ist von diesem Typ. Hier gilt

$$\mathbf{B} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} \sqrt{2} \end{pmatrix}}_{\mathbf{R}}. \quad (4 \text{ Punkte})$$

Die Matrix \mathbf{C} hat zwei Spalten. Diese stehen bereits senkrecht aufeinander, also muss in \mathbf{R} nur noch die Normalisierung der Spalten in den Diagonalelementen stehen und die Spalten normalisiert werden. Damit gilt

$$\mathbf{C} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}}_{\mathbf{R}}. \quad (4 \text{ Punkte})$$

Bei der letzten Matrix sind die Spalten zwar noch nicht senkrecht, dafür ist \mathbf{D} fast eine obere Dreiecksmatrix. Das (2, 3) Element d_{23} von \mathbf{D} läßt sich schnell über eine Givens-Rotation wegrehen:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

Diese Givens-Rotation wird noch auf den Gesamttraum erweitert gemäß

$$\mathbf{G}_{2,3} = \frac{1}{\sqrt{5}} \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Es gilt nach Konstruktion

$$\mathbf{G}_{2,3}\mathbf{D} = \begin{pmatrix} 5 & -3 \\ 0 & \sqrt{5} \\ 0 & 0 \end{pmatrix},$$

also, da $\mathbf{G}_{2,3}$ orthogonal ist,

$$\mathbf{D} = \underbrace{\mathbf{G}_{2,3}^T}_{\mathbf{Q}} \underbrace{\begin{pmatrix} 5 & -3 \\ 0 & \sqrt{5} \\ 0 & 0 \end{pmatrix}}_{\mathbf{R}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2/\sqrt{5} & -1/\sqrt{5} \\ 0 & 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 5 & -3 \\ 0 & \sqrt{5} \\ 0 & 0 \end{pmatrix}. \quad (6 \text{ Punkte})$$

Aufgabe 7: (5+7 Punkte)

Lösen Sie die Ausgleichsprobleme

$$\|\mathbf{A}_i \mathbf{x}_i - \mathbf{b}_i\|_2 = \min, \quad i \in \{1, 2\},$$

wobei

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}$$

und

$$\mathbf{A}_2 = \begin{pmatrix} -1 & -2 & 2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix}.$$

Lösung zu Aufgabe 7: Das erste Ausgleichsproblem wird durch den Vektor

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

gelöst, wie man entweder sofort sieht oder mittels Normalgleichungen, QR-Zerlegung oder SVD schnell berechnen kann. Dass diese Lösung die Pseudonormallösung ist, resultiert aus dem vollen Rang der Matrix \mathbf{A}_1 . (5 Punkte)

Das zweite Ausgleichsproblem hat eine Matrix mit orthogonalen Spalten (und Zeilen), wie man schnell sehen kann. Damit hat diese Matrix auch vollen Rang

3 und das Ausgleichsproblem ist eigentlich ein lineares Gleichungssystem. Die Normierung der Spalten erfolgt mit dem Faktor $1/3$, es gilt also

$$\underbrace{\frac{1}{3} \begin{pmatrix} -1 & -2 & 2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}}_{\tilde{\mathbf{A}}_2} \mathbf{x}_2 = \frac{1}{3} \begin{pmatrix} 9 \\ 9 \\ 9 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix},$$

wobei die skalierte Matrix $\tilde{\mathbf{A}}_2$ eine symmetrische orthogonale Matrix ist, also das Gleichungssystem schnell mittels (nicht direkt vorhandener) Transponation gelöst wird:

$$\mathbf{x}_2 = \frac{1}{3} \begin{pmatrix} -1 & -2 & 2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 & -2 & 2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix}. \quad (7 \text{ Punkte})$$

Auch hier konnte man alternativ mit Normalgleichungen (wobei das Lösen des Gleichungssystems mit $\mathbf{A}_2^T \mathbf{A}_2 = 9 \cdot \mathbf{E}_3$ sehr einfach war), mit der QR-Zerlegung (wobei der Q-Faktor die oben erwähnte Matrix $\tilde{\mathbf{A}}_2$ und der R-Faktor $3 \cdot \mathbf{E}_3$ war) oder der SVD (z.B. $\mathbf{U} = \tilde{\mathbf{A}}_2$, $\mathbf{\Sigma} = 3 \cdot \mathbf{E}_3$, $\mathbf{V} = \mathbf{E}_3$) arbeiten. Das Ergebnis sollte immer dasselbe sein.

Aufgabe 8: (8+4+4 Punkte)

Beschreiben Sie die Potenzmethode. Ist es möglich, dass die Potenzmethode den Eigenwert 2 der Matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

approximiert? Wenn nein, warum nicht? Wenn ja, ist dieses für einen zufällig gewählten Startvektor überhaupt wahrscheinlich?

Lösung zu Aufgabe 8: Die Potenzmethode geht aus von einer quadratischen Matrix \mathbf{A} und einem gegebenen Startvektor \mathbf{u}^0 . Für eine (gewünschte) Anzahl von Indizes $k = 1, 2, \dots$ wird dann rekursiv (z.B.) berechnet:

$$\mathbf{v} = \mathbf{A}\mathbf{u}^{k-1}, \quad \mathbf{u}^k = \mathbf{v}/\ell^H \mathbf{v},$$

wobei ℓ ein fest vorgegebener Vektor ist. Unter geeigneten Voraussetzungen konvergiert dann die Folge $\{\mathbf{u}^k\}_{k=0}^\infty$ gegen einen (meist den zum betragsgrößten Eigenwert gehörigen) Eigenvektor und die Skalierungsgröße $\ell^H \mathbf{v}$ gegen den zugehörigen Eigenwert. (8 Punkte)

Es ist möglich, dass die Potenzmethode nicht den betragsgrößten Eigenwert und zugehörigen Eigenvektor zurückgibt. Für geeignete Startvektoren kann jedes

Eigenpaar zurückgegeben werde, z.B. wenn der Startvektor gleich dem (gesuchten) Eigenvektor ist. (4 Punkte)

Konvergenz gegen den Eigenwert 2 ist der Fall, wenn der Startvektor keinen Anteil am Eigenvektor zum Eigenwert 3 enthält, aber einen nichttrivialen Anteil am Eigenvektor zum Eigenwert 2 hat. Ein zufällig gewählter Startvektor hat mit großer Wahrscheinlichkeit Anteile an jedem Eigenvektor, damit ist die Konvergenz gegen den Eigenwert 2 unter diesen Umständen nahezu (aber nicht völlig) ausgeschlossen. (4 Punkte)

Die Anteile des Startvektors lassen sich in diesem Fall leicht beschreiben. Ein Eigenvektor zum Eigenwert 1 ist gegeben durch

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

ein Eigenvektor zum Eigenwert 2 durch

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

und ein Eigenvektor zum Eigenwert 3 läßt sich herleiten zu

$$\mathbf{v}_3 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}.$$

Damit ist eine Matrix von Eigenvektoren und ihre Inverse gegeben als

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \quad \mathbf{V}^{-1} = \frac{1}{2} \begin{pmatrix} 2 & -2 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Anteile $\{\alpha_i\}_{i=1}^3$ von \mathbf{u}^0 an den einzelnen Eigenvektoren erhält man aus

$$\mathbf{u}^0 = \sum_{i=1}^3 \mathbf{v}_i \alpha_i = \mathbf{V} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix},$$

also gemäß

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \mathbf{V}^{-1} \mathbf{u}^0 = \frac{1}{2} \begin{pmatrix} 2 & -2 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_{10} \\ u_{20} \\ u_{30} \end{pmatrix}.$$

Um den Eigenwert 2 zu approximieren, muss $\alpha_3 = u_{30}/2$ gleich Null sein und (unter der Voraussetzung, dass bereits u_{30} gleich Null ist) $\alpha_2 = u_{20}$ ungleich

Null sein, mit anderen Worten, jeder Startvektor der Form

$$\mathbf{u}^0 = \begin{pmatrix} \star \\ c \\ 0 \end{pmatrix}, \quad c \neq 0$$

mit beliebigem Eintrag in der ersten Komponente führt zur Konvergenz der Potenzmethode gegen den Eigenwert 2. Ein zufällig gewählter Startvektor wird mit an Sicherheit grenzender Wahrscheinlichkeit einen Eintrag ungleich Null in der letzten Komponente haben, womit die Potenzmethode in jedem Fall den Eigenvektor zum Eigenwert 3 berechnet.

Aufgabe 9: (6 Punkte)

Beschreiben Sie das klassische Gram–Schmidt-Verfahren zur Berechnung der QR-Zerlegung einer Matrix.

Lösung zu Aufgabe 9: Das klassische Gram–Schmidt-Verfahren berechnet zu einem gegebenen Satz von linear unabhängigen Vektoren $\{a_j\}_{j=1}^n$ eine Orthonormalbasis $\{q_j\}_{j=1}^n$ und zugehörige Orthonormalisierungskonstanten $\{r_{ij}\}_{i \leq j=1}^n$. Wenn die Vektoren $\{a_j\}_{j=1}^n$ die Spalten von einer Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ darstellen, so ist die Matrix $\mathbf{Q} \in \mathbb{R}^{m \times n}$ gebildet aus den orthonormalen Basisvektoren $\{q_j\}_{j=1}^n$ der Q-Faktor der QR-Zerlegung von \mathbf{A} und der R-Faktor, die quadratische Matrix $\mathbf{R} \in \mathbb{R}^{n \times n}$ enthält auf der Diagonalen die Normalisierungskonstanten, wohingegen die Orthogonalisierungskonstanten im oberen Dreieck von \mathbf{R} enthalten sind.

Das klassische Gram–Schmidt-Verfahren berechnet die orthonormalen Basisvektoren $\{q_j\}_{j=1}^n$ sukzessive. Zuerst wird der erste Vektor orthonormalisiert, da er schon gegen alle (also keine) vorherigen orthonormalen Basisvektoren orthogonalisiert wurde, muss er nur noch normalisiert werden:

$$r_{11} = \|a_1\|_2, \tag{1}$$

$$q_1 = a_1/r_{11}. \tag{2}$$

Alle weiteren orthonormalen Basisvektoren werden durch Orthogonalisierung und anschließende Normalisierung aus den Spalten von \mathbf{A} erzeugt. Sei dazu die Orthonormalbasis bereits gegeben bis zum Index $j - 1$. Dann wird a_j , die j te Spalte von \mathbf{A} orthogonalisiert gemäß

$$r_{ij} = \langle q_i, a_j \rangle, \quad i < j, \tag{3}$$

$$\tilde{q}_j = a_j - \sum_{i=1}^{j-1} q_i r_{ij}. \tag{4}$$

Anschließend wird der so erhaltene Vektor $\tilde{q}_j \in \mathbb{R}^n$ normalisiert,

$$r_{jj} = \|\tilde{q}_j\|_2, \tag{5}$$

$$q_j = \tilde{q}_j/r_{jj}. \tag{6}$$

Im klassischen Gram–Schmidt-Verfahren iteriert man die Orthogonalisierung und anschließende Normalisierung bis zur letzten Spalte und erhält damit die QR-Zerlegung von \mathbf{A} . (6 Punkte)

Dass das (klassische) Gram–Schmidt-Verfahren die QR-Zerlegung liefert, sieht man wie folgt. Schreibt man die Normalisierung (6) um zu

$$\tilde{q}_j = q_j r_{jj}$$

und setzt diesen Ausdruck in die Gleichung (4) ein, so erhält man

$$q_j r_{jj} = a_j - \sum_{i=1}^{j-1} q_i r_{ij}$$

oder, umgeformt,

$$a_j = \sum_{i=1}^j q_i r_{ij}.$$

Das ist aber gerade die j te Spalte der Matrixgleichung

$$\mathbf{A} = \mathbf{QR},$$

und da dieses für alle $j \leq n$ gilt, berechnet das (klassische) Gram–Schmidt-Verfahren die QR-Zerlegung.

Aufgabe 10: (6+3 Punkte)

Beschreiben Sie das modifizierte Gram–Schmidt-Verfahren zur Berechnung der QR-Zerlegung einer Matrix. Worin liegt die Modifikation?

Lösung zu Aufgabe 10: Das modifizierte Gram–Schmidt-Verfahren ist eigentlich nur eine spezielle Art der Implementation des klassischen Gram–Schmidt-Verfahrens, Unterschiede gibt es nur in endlicher Genauigkeit, da die Rundungsfehler sich unterschiedlich stark auswirken.

Da auf dem Rechner Daten oftmals überschrieben werden, was sich mit einer strengen mathematischen Notation nicht immer verträgt, werden die Zuweisungen an eine Variable (einen Skalar oder einen Vektor) mit einem Pfeil anstatt einem Gleichheitszeichen versehen. Die Gleichung (4) des klassischen Gram–Schmidt-Verfahrens wird dann auf dem Rechner meist ausgeführt wie folgt. Sei wieder der Fall der j ten Spalte von \mathbf{A} gegeben. Dann wird der (temporären) Variable \tilde{q}_j zuerst a_j , die j te Spalte von \mathbf{A} zugewiesen,

$$\tilde{q}_j \leftarrow a_j.$$

Danach wird für $i = 1, \dots, j - 1$ *der Reihe nach* der jeweilige Anteil in Richtung des i ten bereits orthonormalen Basisvektors abgezogen, d.h., für $i = 1, \dots, j - 1$

werden die folgenden beiden Gleichungen (7) und (8) iteriert,

$$r_{ij} \leftarrow \langle q_i, a_j \rangle \quad (7)$$

$$\tilde{q}_j \leftarrow \tilde{q}_j - q_i r_{ij}. \quad (8)$$

Im modifizierten Gram–Schmidt-Verfahren berechnet man die Orthogonalisierungskonstanten nicht mittels innerer Produkte mit dem Vektor a_j , sondern mit dem (sich in jedem Schritt ändernden) Vektor \tilde{q}_j ,

$$r_{ij} \leftarrow \langle q_i, \tilde{q}_j \rangle \quad (9)$$

$$\tilde{q}_j \leftarrow \tilde{q}_j - q_i r_{ij}. \quad (10)$$

Die Verwendung der Gleichungen (9) und (10) anstatt (7) und (8) im Iterationsprozess stellt das modifizierte Gram–Schmidt-Verfahren dar (6 Punkte), die Modifikation ist die Verwendung von dem (sich in jedem Schritt ändernden) Vektor \tilde{q}_j anstatt dem (immer gleich bleibenden) Vektor a_j . (3 Punkte)

Um zu sehen, dass das modifizierte Gram–Schmidt-Verfahren zu dem klassischen Gram–Schmidt-Verfahren mathematisch äquivalent ist, zeigt man, dass in exakter Rechnung dieselben Orthogonalisierungskonstanten r_{ij} herauskommen. Seien also Indices i und j mit $i < j$ gegeben. Da sich \tilde{q}_j abhängig von i ändert, fügen wir einen oberen Index (i) hinzu. Aus $\tilde{q}_j^{(1)} = a_j$ und der Kombination der Gleichungen (9) und (10) folgt sofort

$$\tilde{q}_j^{(i)} = a_j - \sum_{\ell=1}^{i-1} q_\ell r_{\ell j},$$

und nach Bildung des inneren Produktes mit q_i erhält man

$$\langle q_i, \tilde{q}_j^{(i)} \rangle = \langle q_i, a_j \rangle - \sum_{\ell=1}^{i-1} \underbrace{\langle q_i, q_\ell \rangle}_{=0} r_{\ell j} = \langle q_i, a_j \rangle.$$

also gilt (mit der kurzen Notation kGS für das klassische Gram–Schmidt-Verfahren und mGS für das modifizierte Gram–Schmidt-Verfahren)

$$r_{ij}^{\text{mGS}} = \langle q_i, \tilde{q}_j^{(i)} \rangle = \langle q_i, a_j \rangle = r_{ij}^{\text{kGS}}.$$

Dass das klassische und modifizierte Gram–Schmidt-Verfahren unterschiedliche Ergebnisse liefern sieht man am im Skript gegebenen Beispiel. Man kann beweisen, dass das modifizierte Gram–Schmidt-Verfahren numerisch stabiler ist als das klassische Gram–Schmidt-Verfahren.

Aufgabe 11: (4+4+5 Punkte)

Berechnen Sie die Pseudoinversen der Matrizen

$$\mathbf{H} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Lösung zu Aufgabe 11: Die Pseudoinverse ist die eindeutige lineare Abbildung, die einer gegebenen rechten Seite die Pseudonormallösung eines Ausgleichsproblems zuordnet. Da im Falle einer Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ mit $m \geq n$ mit vollem (Spalten-)Rang n die Pseudonormallösung die eindeutige Lösung der Normalgleichungen ist, gilt im Falle eines vollen Ranges

$$\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T.$$

Die beiden Matrizen \mathbf{H} und \mathbf{K} haben vollen Spaltenrang. Die Pseudoinverse von \mathbf{H} ist gegeben als

$$\mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T = (\|\mathbf{H}\|_2^2)^{-1} \mathbf{H}^T = \frac{1}{30} \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}, \quad (4 \text{ Punkte})$$

die von \mathbf{K} als

$$\begin{aligned} \mathbf{K}^\dagger &= (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix}. \end{aligned} \quad (4 \text{ Punkte})$$

Die Matrix \mathbf{M} hat nicht vollen Rang 2, sie hat nur den Rang 1, da sie zweimal dieselbe Spalte enthält. Damit läßt sie sich als äußeres Produkt auffassen und die (ökonomische) SVD schnell angeben:

$$\begin{aligned} \mathbf{M} &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{1}{2} \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}_{\mathbf{U}_1} \underbrace{\begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}}_{\Sigma_1} \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix}}_{\mathbf{V}_1^T} \\ &= \frac{1}{6} \underbrace{\begin{pmatrix} 3 & 3 & 3 & 3 \\ 3 & -5 & 1 & 1 \\ 3 & 1 & -5 & 1 \\ 3 & 1 & 1 & -5 \end{pmatrix}}_{\mathbf{U}} \underbrace{\begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_{\mathbf{V}^T}. \end{aligned}$$

Die volle SVD wird nicht benötigt, sie ist nur der Vollständigkeit halber einmal angegeben. Mit Kenntnis der (ökonomischen) SVD kann die Pseudoinverse schnell berechnet werden:

$$\begin{aligned} \mathbf{M}^\dagger &= \mathbf{V}_1 (\Sigma_1)^{-1} \mathbf{U}_1^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \frac{1}{2\sqrt{2}} \cdot \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{8} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \end{aligned} \quad (5 \text{ Punkte})$$

Aufgabe 12: (6+6 Punkte)

Berechnen Sie die Kondition der Matrix

$$\mathbf{N} = \begin{pmatrix} 1 & 1.001 \\ 1.001 & 1 \end{pmatrix}$$

bezüglich der Spektralnorm $\|\cdot\|_2$. Welche relative Genauigkeit in \mathbf{x} erwarten Sie bei der Lösung von

$$\mathbf{N}\mathbf{x} = \mathbf{b},$$

wenn \mathbf{b} aus Messwerten resultiert und einen relativen Fehler von 10^{-4} hat?

Lösung zu Aufgabe 12: Die Kondition einer regulären Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist definiert bezüglich einer Matrixnorm $\|\cdot\|$ durch

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|.$$

Die Spektralnorm einer Matrix ist gegeben durch den größten Singulärwert, die Spektralnorm einer inversen Matrix durch den Kehrwert des kleinsten Singulärwertes. Im Falle einer symmetrischen reellen Matrix (allgemeiner einer normalen Matrix) sind die Singulärwerte gegeben durch die Beträge der Eigenwerte. Symmetrische reelle Matrizen (allgemein normale Matrizen) ermöglichen ein System von orthogonalen Eigenvektoren (unitären Eigenvektoren).

Die Matrix \mathbf{N} hat, wie leicht ersichtlich ist, den Vektor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

als Eigenvektor zum Eigenwert $\lambda_1 = 1 + 1.001 = 2.001$. Der andere Eigenvektor muss auf diesem senkrecht stehen, bis auf Skalierung ist er also gleich

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

und gehört zum Eigenwert $\lambda_2 = 1 - 1.001 = -0.001$. Damit sind die Singulärwerte gegeben als $\sigma_1 = 2.001$ und $\sigma_2 = 0.001$. Die Kondition berechnet sich zu

$$\kappa_2(\mathbf{N}) = \|\mathbf{N}\|_2 \|\mathbf{N}^{-1}\|_2 = \sigma_1 \cdot \frac{1}{\sigma_2} = \frac{\sigma_1}{\sigma_2} = \frac{2.001}{0.001} = 2001. \quad (6 \text{ Punkte})$$

Sei das eigentlich zu lösende Gleichungssystem gegeben als

$$\mathbf{N}\hat{\mathbf{x}} = \hat{\mathbf{b}},$$

und \mathbf{b} die mit einem relativen Fehler von 10^{-4} gemessene Lösung. Der relative Fehler in der Lösung eines gestörten Gleichungssystemes

$$\mathbf{N}(\hat{\mathbf{x}} + \Delta\mathbf{x}) = \mathbf{b} = \hat{\mathbf{b}} + \Delta\mathbf{b}$$

läßt sich allgemein durch

$$\frac{\|\Delta \mathbf{x}\|_2}{\|\hat{\mathbf{x}}\|_2} \leq \frac{\kappa_2(\mathbf{N})}{1 - \kappa_2(\mathbf{N}) \frac{\|\Delta \mathbf{N}\|_2}{\|\mathbf{N}\|_2}} \left(\frac{\|\Delta \mathbf{N}\|_2}{\|\mathbf{N}\|_2} + \frac{\|\Delta \mathbf{b}\|_2}{\|\hat{\mathbf{b}}\|_2} \right)$$

beschreiben. Im gegebenen Fall wird die Matrix nicht gestört, und der relative Fehler in \mathbf{b} ist gegeben als

$$\frac{\|\Delta \mathbf{b}\|_2}{\|\hat{\mathbf{b}}\|_2} = 10^{-4}.$$

Damit ist die relative Abweichung von der “korrekten” Lösung beschränkt durch

$$\frac{\|\Delta \mathbf{x}\|_2}{\|\hat{\mathbf{x}}\|_2} \leq \kappa_2(\mathbf{N}) \cdot 10^{-4} = 2001 \cdot 10^{-4} \approx 2 \cdot 10^{-1}. \quad (6 \text{ Punkte})$$