

# 2 EIGENWERTPROBLEME

Viele numerische Lösungsansätze von z.B. physikalischen Problemen erfordern die Lösung eines sogenannten Eigenwertproblems. Zu einer Matrix  $A \in \mathbb{R}^{n \times n}$  finde man eine Zahl  $\lambda \in \mathbb{C}$  und einen Vektor  $v \in \mathbb{C}^n$ ,  $v \neq 0$ , sodass die *Eigenwertgleichung*

$$Av = \lambda v$$

erfüllt ist. Die Zahl  $\lambda$  heißt *Eigenwert* und der Vektor  $v$  *Eigenvektor* zum Eigenwert  $\lambda$ . Betrachten wir nun zunächst einige konkrete Beispiele, die auf ein Eigenwertproblem führen.

**Beispiel 2.0.1 (Sturm-Liouville-Problem)** Die mathematische Modellierung zur Beschreibung der Überlagerung von Schwingungsvorgängen – wie sie etwa beim Brückenbau durchgeführt wird, um Resonanzen zu vermeiden, die einen Brückeneinsturz verursachen könnten – führt auf das sogenannte Sturm-Liouville Problem: Zu einer bekannten stetigen Funktion  $r(x) > 0$ ,  $x \in [0, 1]$ , finde man die Zahl  $\lambda$  und die Funktion  $u(x)$ , die die Differentialgleichung

$$-u''(x) - \lambda r(x) u(x) = 0, \quad x \in (0, 1) \quad (2.1)$$

mit den Randbedingungen

$$u(0) = u(1) = 0$$

erfüllen. Um die Lösung über ein Diskretisierungsverfahren numerisch anzunähern, betrachten wir Gitterpunkte

$$x_j = jh, \quad j = 0, \dots, n, \quad h = \frac{1}{n}$$

und ersetzen  $u''(x_j)$  durch den zweiten zentrierten Differenzenquotienten

$$\frac{u(x_j + h) - 2u(x_j) + u(x_j - h))}{h^2}, \quad j = 1, \dots, n - 1.$$

Es ergibt sich so ein System

$$Au - \lambda Ru = 0 \quad (2.2)$$

für die Unbekannten  $\lambda$  und  $u_i \approx u(x_i)$ ,  $i = 1, \dots, n - 1$ , wobei

$$A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \quad R = \begin{pmatrix} r(x_1) & & & & 0 \\ & r(x_2) & & & \\ & & \ddots & & \\ 0 & & & & r(x_{n-1}) \end{pmatrix}. \quad (2.3)$$

Mit  $R^{1/2} = \text{diag}(\sqrt{r(x_1)}, \dots, \sqrt{r(x_{n-1})})$ ,  $R^{-1/2} = (R^{1/2})^{-1}$ ,  $B := R^{-1/2}AR^{-1/2}$  sowie  $v = R^{1/2}u$  erhält man aus (2.2) die transformierte Gleichung

$$Bv = \lambda v,$$

also ein Eigenwertproblem.

**Bemerkung 2.0.2** Da  $R$  regulär ist, hätte man die Gleichung (2.2) auch von links mit  $R^{-1}$  multiplizieren und so das Eigenwertproblem  $R^{-1}Au = \lambda u$  erhalten können. Im Gegensatz zur Matrix  $B$  ist die Matrix  $R^{-1}A$  jedoch nicht symmetrisch.

**Beispiel 2.0.3 (Wellengleichung auf  $[0, 1]$ )** Wir betrachten die Wellengleichung

$$\frac{1}{c^2} u_{tt}(t, x) - u_{xx}(t, x) = 0, \quad t > 0, x \in (0, 1) \quad (2.4)$$

mit den Anfangs- und Randbedingungen

$$u(t, 0) = u(t, 1) = 0, \quad t > 0 \quad (2.5)$$

$$u(0, x) = u_0(x), \quad x \in [0, 1] \quad (2.6)$$

$$u_t(0, x) = u_1(x), \quad x \in [0, 1]. \quad (2.7)$$

Wendet man nun den Separationsansatz  $u(t, x) = T(t) \cdot X(x)$  auf (2.4) an, so ergibt sich

$$\frac{1}{c^2} T''(t)X(x) - T(t)X''(x) = 0 \quad \Leftrightarrow \quad \frac{1}{c^2} \frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)} = \text{const.} = -\lambda.$$

Es folgt

$$-T''(t) - \lambda c^2 T(t) = 0, \quad (2.8)$$

$$-X''(x) - \lambda X(x) = 0, \quad (2.9)$$

$$X(0) = X(1) = 0. \quad (2.10)$$

Es ist also als Teilproblem das Sturm-Liouville-Problem (2.9) – (2.10) bzgl.  $X$  zu lösen. Zu gegebenem  $\lambda \in \mathbb{R}$  muss die Funktion  $T$ , welche (2.8) genügt, von der Form

$$T(t) = a \cos(\sqrt{\lambda} ct) + b \sin(\sqrt{\lambda} ct)$$

sein. Kann man nun  $m$  Eigenpaare  $(\lambda_k, X_k)$ ,  $k = 1, \dots, m$  des Sturm-Liouville-Problems näherungsweise numerisch bestimmen, so ist für beliebige  $a_k, b_k \in \mathbb{R}$

$$\hat{u}(t, x) := \sum_{k=1}^m \left( a_k \cos(\sqrt{\lambda_k} ct) + b_k \sin(\sqrt{\lambda_k} ct) \right) X_k(x)$$

eine Näherungslösung von (2.4) – (2.5). Durch geeignete Wahl der Koeffizienten  $a_k$  und  $b_k$  kann man erreichen, dass auch die Anfangswerte (2.6) und (2.7) näherungsweise angenommen werden. Hierbei fordert man die Übereinstimmungen

$$\begin{aligned} \hat{u}(0, x_i) &= \sum_{k=1}^m a_k X_k(x_i) \stackrel{!}{=} u_0(x_i), \\ \hat{u}_t(0, x_i) &= \sum_{k=1}^m b_k \sqrt{\lambda_k} c X_k(x_i) \stackrel{!}{=} u_1(x_i), \quad i = 0, \dots, n, \end{aligned}$$

wobei  $x_i$  die Gitterpunkte aus der Diskretisierung des Sturm-Liouville-Problems seien. Dies ist ein lineares Gleichungssystem, aus dem man die Koeffizienten bestimmen kann.

**Bemerkung 2.0.4 (Fourierreihen)** Man kann zeigen, dass  $\lambda_k = k^2\pi^2$  für  $k \in \mathbb{N}$  die Eigenwerte des Sturm-Liouville-Problems (2.9) – (2.10) sind mit den zugehörigen Eigenfunktionen  $X_k(x) = \sin(k\pi x)$ . Beachtet man nun, dass für  $N \in \mathbb{N}$  die endliche Linearkombination

$$u_N(t, x) := \sum_{k=1}^N \left( a_k \cos(k\pi ct) + b_k \sin(k\pi ct) \right) \sin(k\pi x)$$

die Anfangswerte

$$u_N(0, x) = \sum_{k=1}^N a_k \sin(k\pi x)$$

$$\frac{\partial}{\partial t} u_N(0, x) = \sum_{k=1}^N b_k k\pi c \sin(k\pi x)$$

annimmt und wählt  $a_k$  als die Fourierkoeffizienten von  $u_0$

$$a_k = 2 \int_0^1 u_0(s) \sin(k\pi s) ds$$

sowie  $b_k$  gemäß  $b_k = \frac{\beta_k}{k\pi c}$  mit den Fourierkoeffizienten  $\beta_k$  von  $u_1$

$$\beta_k = 2 \int_0^1 u_1(s) \sin(k\pi s) ds,$$

so löst unter bestimmten Differenzierbarkeitsanforderungen an  $u_0$  und  $u_1$  die Funktion

$$u(t, x) := \lim_{N \rightarrow \infty} u_N(t, x) = \sum_{k=1}^{\infty} \left( a_k \cos(k\pi ct) + b_k \sin(k\pi ct) \right) \sin(k\pi x)$$

die Wellengleichung auf  $[0, 1]$ .

Eigenwerte beschreiben nicht nur physikalische Eigenschaften, sondern spielen auch eine wichtige Rolle in der Mathematik, beispielsweise in der Numerik, um die Norm oder Konditionszahl einer Matrix zu bestimmen.

**Beispiel 2.0.5 (Norm und Konditionszahl)** i) Für eine beliebige Matrix  $A \in \mathbb{R}^{n \times n}$  gilt

$$\|A\|_2 = \sqrt{\varrho(A^T A)},$$

wobei  $\varrho$  wiederum den Spektralradius einer Matrix, also den Betrag des betragsgrößten Eigenwertes, bezeichnet.

ii) Für eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  gilt

$$\|A\|_2 = \varrho(A).$$

iii) Für eine nichtsinguläre symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  gilt

$$\kappa_2(A) = \varrho(A) \varrho(A^{-1}),$$

wobei  $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$  die Konditionszahl von  $A$  bzgl. der 2-Norm bezeichnet.

## 2.1 GRUNDLAGEN AUS DER LINEAREN ALGEBRA

Bevor wir uns der Abschätzung sowie der numerischen Berechnung von Eigenwerten widmen, werden an dieser Stelle zunächst elementare Ergebnisse aus der linearen Algebra zusammengefasst, welche in Zusammenhang mit der Eigenwertberechnung stehen.

Sei  $A \in \mathbb{C}^{n \times n}$ ,  $\lambda \in \mathbb{C}$  ein Eigenwert der Matrix  $A$  und  $v \in \mathbb{C} \setminus \{0\}$  ein zugehöriger Eigenvektor:

$$Av = \lambda v. \quad (2.11)$$

Zusätzlich zur Definition (2.11) eines (rechtsseitigen) Eigenvektors  $v \in \mathbb{C}^n \setminus \{0\}$  kann man auch linksseitige Eigenvektoren betrachten.  $x \in \mathbb{C}^n \setminus \{0\}$  heißt linksseitiger Eigenvektor von  $A$ , falls  $x^H A = \lambda x^H$  gilt, wobei  $x^H$  denjenigen Vektor bezeichnet, der aus  $x$  durch Bildung der konjugiert komplexen Einträge und Transponieren hervorgeht.

**Satz 2.1.1** Die Eigenwerte einer Matrix  $A \in \mathbb{C}^{n \times n}$  sind gerade die  $n$  Nullstellen des zugehörigen charakteristischen Polynoms

$$P_A(\lambda) := \det(A - \lambda I).$$

*Beweis.* Die Eigenwertgleichung (2.11) besagt offensichtlich, dass  $(A - \lambda I)v = 0$  gilt, also die Matrix  $A - \lambda I$  singularär ist. Letzteres ist zu  $\det(A - \lambda I) = 0$  äquivalent.  $\square$

**Definition 2.1.2** Die Menge aller Eigenwerte einer Matrix  $A \in \mathbb{C}^{n \times n}$  wird als Spektrum bezeichnet und man verwendet die Schreibweise

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid \lambda \text{ ist Eigenwert von } A\}.$$

**Satz 2.1.3** Sei  $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ . Es besteht folgender Zusammenhang zwischen der Determinante bzw. der Spur von  $A$  und den Eigenwerten  $\lambda_i$ :

$$\det(A) = \prod_{i=1}^n \lambda_i,$$

$$\operatorname{tr}(A) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i.$$

*Beweis.* Die zweite Identität erhält man z.B. dadurch, dass man den Koeffizienten vor  $\lambda^{n-1}$  im charakteristischen Polynom untersucht.

$$P_A(\lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \\ \vdots & & \ddots & \\ a_{n1} & & & a_{nn} - \lambda \end{pmatrix}$$

$$= (a_{11} - \lambda) \cdots (a_{nn} - \lambda) + \text{Terme mit weniger als } (n-1) \text{ } (a_{ii} - \lambda)\text{-Termen}$$

$$= (-\lambda)^n + (-\lambda)^{n-1}(a_{11} + a_{22} + \dots + a_{nn}) + \sum_{k=0}^{n-2} c_k \lambda^k$$

$$\stackrel{!}{=} (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

$\square$

Ein Unterraum  $S \subseteq \mathbb{C}^n$  mit der Eigenschaft

$$x \in S \Rightarrow Ax \in S$$

wird als invariant bzgl.  $A$  bezeichnet. Der Aufspan eines (rechtsseitigen) Eigenvektors bildet somit einen eindimensionalen Unterraum, der invariant bleibt unter der Multiplikation mit  $A$ . Es sei

$$AX = XB, \quad B \in \mathbb{C}^{k \times k}, \quad X \in \mathbb{C}^{n \times k},$$

dann ist  $\operatorname{Bild}(X)$  invariant bzgl.  $A$  und aus  $By = \lambda y, y \in \mathbb{C}^k \setminus \{0\}$  folgt

$$A(Xy) = (AX)y = (XB)y = \lambda Xy.$$

Falls  $X$  vollen Spaltenrang hat, dann impliziert  $AX = XB$  also  $\sigma(B) \subseteq \sigma(A)$ . Für den Fall einer regulären quadratischen Matrix  $X$  gilt für das Spektrum von  $A$  und  $B = X^{-1}AX$  folgendes Ergebnis.

**Lemma 2.1.4** Sei  $A \in \mathbb{C}^{n \times n}$  beliebig sowie  $X \in \mathbb{C}^{n \times n}$  eine reguläre Matrix. Dann gilt

$$\sigma(A) = \sigma(X^{-1}AX).$$

Ähnliche Matrizen haben also das gleiche Spektrum.

*Beweis.* Die Aussage folgt direkt mit Satz 2.1.1 aus der Tatsache, dass ähnliche Matrizen das gleiche charakteristische Polynom haben:

$$\det(X^{-1}AX) = \det(X^{-1}(A - \lambda I)X) = \det(X^{-1}) \det(A - \lambda I) \det(X) = \det(A - \lambda I).$$

□

Multipliziert man eine hermitesche Matrix  $A \in \mathbb{C}^{n \times n}$  mit einer nichtsingulären Matrix  $X \in \mathbb{C}^{n \times n}$  wie folgt

$$X^H A X = B \in \mathbb{C}^{n \times n},$$

so hat  $B$  im Allgemeinen nicht die gleichen Eigenwerte wie  $A$ . Der Sylvester'sche Trägheitssatz sagt aus, dass wenigstens die Vorzeichen der Eigenwerte sich nicht ändern.

**Satz 2.1.5 (Sylvester'scher Trägheitssatz)** Es sei  $A \in \mathbb{C}^{n \times n}$  hermitesch und  $X \in \mathbb{C}^{n \times n}$  regulär. Dann haben  $A$  und  $X^H A X$  den gleichen Rang sowie die gleichen Anzahlen positiver und negativer Eigenwerte.

*Beweis.* Z.B. Fischer (Lineare Algebra)

□

Wir kommen nun zu einer wichtigen Matrixfaktorisierung, der Schur-Zerlegung. Im Gegensatz zu der ebenfalls aus der linearen Algebra bekannten Jordan'schen Normalform führt ihre Berechnung auf numerisch stabile Algorithmen, da man hierfür nur orthogonale bzw. unitäre Transformationen benötigt und diese normerhaltend sind.

Diese Faktorisierung spielt beim  $QR$ -Verfahren zur Berechnung von Eigenwerten eine zentrale Rolle.

**Satz 2.1.6 (Satz von Schur)** Sei  $A \in \mathbb{C}^{n \times n}$ . Dann gibt es eine unitäre Matrix  $Q \in \mathbb{C}^{n \times n}$  ( $Q Q^H = Q^H Q = I$ ), so dass

$$Q^H A Q = R$$

ist, wobei  $R \in \mathbb{C}^{n \times n}$  eine obere Dreiecksmatrix ist und die Diagonaleinträge  $r_{11}, \dots, r_{nn}$  von  $R$  die Eigenwerte von  $A$  sind.

Ist  $A$  reell, so sollte man versuchen mit reeller Arithmetik auszukommen. Allerdings ist  $R$  dann nur eine obere Block-Dreiecksmatrix mit  $1 \times 1$  oder  $2 \times 2$  Blöcken.

**Satz 2.1.7 (Reelle Schur-Zerlegung)** Für jedes  $A \in \mathbb{R}^{n \times n}$ , gibt es eine orthogonale Matrix  $Q \in \mathbb{R}^{n \times n}$ , so dass  $Q^T A Q$  in quasi-oberer Dreiecksform

$$Q^T A Q = \begin{pmatrix} R_{11} & \cdots & R_{1m} \\ & \ddots & \vdots \\ & & R_{mm} \end{pmatrix}$$

ist, wobei jedes  $R_{ii}$  entweder ein  $1 \times 1$  oder  $2 \times 2$  Block ist. Dabei sind die  $1 \times 1$  Blöcke unter  $R_{11}, \dots, R_{mm}$  die reellen Eigenwerte von  $A$  und die  $2 \times 2$  Blöcke enthalten die Paare von komplex-konjugierten Eigenwerten von  $A$ .

Zum Schluss dieses Abschnitts betrachten wir noch Matrixfaktorisierungen für eine besondere Klasse von Matrizen.

**Definition 2.1.8**  $A \in \mathbb{C}^{n \times n}$  heißt normal, wenn gilt

$$A^H A = A A^H.$$

**Bemerkung 2.1.9** Alle hermiteschen, schieferhermiteschen ( $A^H = -A$ ), Diagonal- und unitären Matrizen sind Beispiele für normale Matrizen.

**Korollar 2.1.10 (Schur-Zerlegung von normalen Matrizen)** Eine Matrix  $A \in \mathbb{C}^{n \times n}$  ist genau dann normal, wenn eine unitäre Matrix  $Q \in \mathbb{C}^{n \times n}$  existiert mit  $Q^H A Q = D = \text{diag}(\lambda_1, \dots, \lambda_n)$ .

*Beweis.* Aus der unitären Ähnlichkeit von  $A$  zu einer Diagonalmatrix, folgt offensichtlich, dass  $A$  normal ist. Andererseits sei  $A$  normal und  $Q^H A Q = R$  sei die zugehörige Schur'sche Normalform. Dann ist auch  $R$  normal.

$$R^H R = Q^H A^H Q Q^H A Q = Q^H A^H A Q = Q^H A A^H Q = Q^H A Q Q^H A^H Q = R R^H$$

Die Behauptung folgt nun aus der Tatsache, dass eine normale, obere Dreiecksmatrix eine Diagonalmatrix ist.  $\square$

**Korollar 2.1.11 (Schur-Zerlegung von symmetrischen Matrizen)** Jede reelle, symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  lässt sich mittels einer orthogonalen Matrix  $Q \in \mathbb{R}^{n \times n}$  auf Diagonalgestalt transformieren

$$Q^T A Q = D = \text{diag}(\lambda_1, \dots, \lambda_n),$$

wobei  $\lambda_1, \dots, \lambda_n$  die reellen Eigenwerte von  $A$  sind.

## 2.2 ABSCHÄTZUNGEN UND GEOMETRISCHE LAGE DER EIGENWERTE

Im Folgenden werden einige Abschätzungen für das Spektrum einer Matrix  $A \in \mathbb{C}^{n \times n}$  vorgestellt. Solche Abschätzungen können später bei der Wahl geeigneter Startwerte für lokal konvergente Iterationsverfahren zur Bestimmung der Eigenwerte nützlich sein. Eine erste Eingrenzung der Lage der Eigenwerte von  $A$  liefert der folgende Satz.

**Satz 2.2.1 (Abschätzung mittels Matrixnorm)** Sei  $\|\cdot\|$  eine konsistente Matrixnorm, d.h. es gebe eine Vektornorm  $\|\cdot\|$  auf  $\mathbb{C}^n$ , sodass  $\|Av\| \leq \|A\| \|v\|$ , dann gilt

$$|\lambda| \leq \|A\| \quad \forall \lambda \in \sigma(A).$$

*Beweis.* Seien  $\lambda$  ein Eigenwert von  $A$  und  $v \neq 0$  ein dazugehöriger Eigenvektor. Da  $\|\cdot\|$  konsistent ist, erhält man

$$|\lambda| \|v\| = \|\lambda v\| = \|Av\| \leq \|A\| \|v\|,$$

sodass  $|\lambda| \leq \|A\|$  folgt.  $\square$

**Definition 2.2.2 (Wertebereich einer Matrix)** Sei  $A \in \mathbb{K}^{n \times n}$  ( $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ ). Die Menge aller Rayleigh-Quotienten  $\frac{x^H A x}{x^H x}$  mit  $x \in \mathbb{C}^n \setminus \{0\}$ , also

$$W(A) := \left\{ \frac{x^H A x}{x^H x} : x \in \mathbb{C}^n \setminus \{0\} \right\}$$

wird als Wertebereich der Matrix  $A$  bezeichnet.

**Bemerkungen 2.2.3** i) Selbst für eine reelle Matrix  $A \in \mathbb{R}^{n \times n}$  werden beim Wertebereich alle komplexen Vektoren  $x \in \mathbb{C}^n \setminus \{0\}$  durchlaufen.

ii) Der Wertebereich  $W(A)$  enthält alle Eigenwerte von  $A$ . Ist  $x \in \mathbb{C}^n \setminus \{0\}$  ein Eigenvektor von  $A$ , dann ist der zugehörige Rayleigh-Quotient gerade der Eigenwert.

**Lemma 2.2.4 (Eigenschaften des Wertebereichs)** Sei  $A \in \mathbb{K}^{n \times n}$ ,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ .

- i)  $W(A)$  ist zusammenhängend.
- ii) Ist  $A$  hermitesch, dann ist  $W(A)$  das reelle Intervall  $[\lambda_{\min}, \lambda_{\max}]$ .
- iii) Ist  $A$  schiefhermitesch ( $A = -A^H$ ), dann ist  $W(A)$  eine rein imaginäre Menge, nämlich die konvexe Hülle der Eigenwerte von  $A$ .

**Satz 2.2.5 (Bendixson)** Das Spektrum von  $A \in \mathbb{K}^{n \times n}$  ist in dem Rechteck

$$R = W\left(\frac{A + A^H}{2}\right) + W\left(\frac{A - A^H}{2}\right)$$

enthalten, wobei die Summe zweier Menge folgendermaßen zu verstehen ist

$$A + B = \{a + b : a \in A, b \in B\}.$$

*Beweis.* Sei  $\lambda \in \mathbb{C}$  ein Eigenwert von  $A$  zum Eigenvektor  $x \in \mathbb{C}^n$  mit  $\|x\|_2 = 1$ . Dann gilt

$$\begin{aligned} \lambda &= \lambda x^H x = x^H A x = x^H \left( \frac{A + A^H}{2} + \frac{A - A^H}{2} \right) x \\ &= x^H \frac{A + A^H}{2} x + x^H \frac{A - A^H}{2} x \in W\left(\frac{A + A^H}{2}\right) + W\left(\frac{A - A^H}{2}\right) \end{aligned}$$

□

**Bemerkung 2.2.6** Da  $\frac{A+A^H}{2}$  hermitesch und  $\frac{A-A^H}{2}$  schiefhermitesch sind, besagt der Satz von Bendixson, dass für jedes  $\lambda \in \sigma(A)$  gilt

$$\begin{aligned} \lambda_{\min}\left(\frac{A + A^H}{2}\right) &\leq \operatorname{Re}(\lambda) \leq \lambda_{\max}\left(\frac{A + A^H}{2}\right) \\ \min\left\{\operatorname{Im}(\mu) \mid \mu \in \sigma\left(\frac{A - A^H}{2}\right)\right\} &\leq \operatorname{Im}(\lambda) \leq \max\left\{\operatorname{Im}(\mu) \mid \mu \in \sigma\left(\frac{A - A^H}{2}\right)\right\}. \end{aligned}$$

Eine weitere *a priori* Schranke wird durch das folgende Resultat gegeben.

**Satz 2.2.7 (Gerschgorin Kreise)** Sei  $A \in \mathbb{C}^{n \times n}$ . Dann gilt

$$\sigma(A) \subseteq \mathcal{S}_{\mathcal{R}} := \bigcup_{i=1}^n \mathcal{R}_i, \quad \mathcal{R}_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right\}. \quad (2.12)$$

Die Mengen  $\mathcal{R}_i$  werden Gerschgorin Kreise genannt.

*Beweis.* Sei  $\lambda \in \sigma(A)$  und  $Ax = \lambda x$  für ein  $x \neq 0$ . Dann existiert ein  $x_i$  mit  $|x_j| \leq |x_i| \forall i \neq j$ .  $(Ax)_i$  sei die  $i$ -te Komponente von  $Ax$ , dann ist

$$\begin{aligned} \lambda x_i &= (Ax)_i = \sum_{j=1}^n a_{ij} x_j \\ \Leftrightarrow \lambda - a_{ii} &= \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} \frac{x_j}{x_i} \\ \Rightarrow |\lambda - a_{ii}| &= \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \end{aligned}$$

also

$$\lambda \in \mathcal{R}_i \subset \bigcup_{i=1}^n \mathcal{R}_i.$$

□

Da  $A$  und  $A^T$  das gleiche Spektrum haben, gilt Satz 2.2.7 auch in der Form

$$\sigma(A) \subseteq \mathcal{S}_{\mathcal{C}} := \bigcup_{j=1}^n \mathcal{C}_j, \quad \mathcal{C}_j = \left\{ z \in \mathbb{C} : |z - a_{jj}| \leq \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}| \right\}. \quad (2.13)$$

Die Aussagen (2.12) und (2.13) liefern zusammen das erste Gerschgorin Theorem.

**Satz 2.2.8 (Erstes Gerschgorin Theorem)** Für eine Matrix  $A \in \mathbb{C}^{n \times n}$  erfüllen die Eigenwerte folgende Eigenschaft

$$\forall \lambda \in \sigma(A) : \lambda \in \mathcal{S}_{\mathcal{R}} \cap \mathcal{S}_{\mathcal{C}}.$$

Das zweite Gerschgorin Theorem liefert in bestimmten Fällen eine Aussage über die Verteilung der Eigenwerte auf die verschiedenen Gerschgorin Kreise.

**Satz 2.2.9 (Zweites Gerschgorin Theorem)** Seien

$$\mathcal{M}_1 := \bigcup_{j=1}^k \mathcal{R}_{i_j}, \quad \mathcal{M}_2 := \bigcup_{j=k+1}^n \mathcal{R}_{i_j}.$$

Ist die Vereinigung  $\mathcal{M}_1$  von  $k$  Kreisen disjunkt von der Vereinigung  $\mathcal{M}_2$  der übrigen  $n - k$  Kreise, also  $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$ , so enthält  $\mathcal{M}_1$  genau  $k$  und  $\mathcal{M}_2$  die übrigen  $n - k$  Eigenwerte, jeder entsprechend seiner algebraischen Vielfachheit gezählt.

*Beweis.* Sei  $D := \text{diag}(a_{11}, \dots, a_{nn})$  sowie für  $t \in [0, 1]$

$$A_t := D + t(A - D).$$

Dann ist  $A_0 = D$  und  $A_1 = A$ . Die Eigenwerte von  $A_t$  sind stetige Funktionen in  $t$ . Wendet man Satz 2.2.7 auf  $A_t$  an, so erhält man für  $t = 0$ , dass genau  $k$  Eigenwerte von  $A_0 = D$  in  $\mathcal{M}_1$  liegen und die restlichen  $n - k$  in  $\mathcal{M}_2$ , wobei mehrfache Eigenwerte entsprechend ihrer algebraischen Vielfachheit gezählt werden. Für  $0 \leq t \leq 1$  müssen alle Eigenwerte von  $A_t$  ebenfalls in den Kreisen liegen. Daher und aufgrund der Stetigkeit der Eigenwerte folgt, dass auch  $k$  Eigenwerte von  $A_1 = A$  in  $\mathcal{M}_1$  und die übrigen  $n - k$  in  $\mathcal{M}_2$  liegen. □

Um das dritte Gerschgorin Theorem formulieren zu können, erinnern wir uns folgende Eigenschaft von Matrizen, die wir bereits aus Numerik I kennen.

**Definition 2.2.10** Eine Matrix  $A \in \mathbb{C}^{n \times n}$  heißt *reduzibel*, wenn eine Permutationsmatrix  $P$  existiert, sodass

$$PAP^T = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}$$

mit quadratischen Matrizen  $B_{11}$  und  $B_{22}$  gilt.  $A$  ist *irreduzibel*, wenn  $A$  nicht reduzibel ist.

**Satz 2.2.11 (Drittes Gerschgorin Theorem)** Sei  $A \in \mathbb{C}^{n \times n}$  eine irreduzible Matrix. Ein Eigenwert  $\lambda \in \sigma(A)$  kann nicht auf dem Rand von  $\mathcal{S}_{\mathcal{R}}$  liegen, es sei denn er liegt auf dem Rand eines jeden Kreises  $\mathcal{R}_i$  für  $i = 1, \dots, n$ .



**Beispiel 2.2.12** Wir betrachten die Matrix

$$A = \begin{pmatrix} 4 & 0 & -3 \\ 0 & -1 & 1 \\ -1 & 1 & -2 \end{pmatrix} \quad \text{mit} \quad A^H = A^T = \begin{pmatrix} 4 & 0 & -1 \\ 0 & -1 & 1 \\ -3 & 1 & -2 \end{pmatrix}$$

und werden zunächst den Satz von Bendixson, Satz 2.2.5, anwenden, um die Eigenwerte der Matrix  $A$  abzuschätzen. Wir erhalten

$$\frac{A + A^H}{2} = \begin{pmatrix} 4 & 0 & -2 \\ 0 & -1 & 1 \\ -2 & 1 & -2 \end{pmatrix}, \quad \frac{A - A^H}{2} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Wendet man auf diese beiden Matrizen Satz 2.2.7 an, so ergeben sich folgende erste Abschätzungen eines Eigenwertes  $\lambda$  von  $A$

$$-5 \leq \operatorname{Re}(\lambda) \leq 6, \quad |\operatorname{Im}(\lambda)| \leq 1.$$

Die Gerschgorin Kreise von  $A$  ergeben sich zu

$$\mathcal{R}_1 = \{z \in \mathbb{C} : |z - 4| \leq 3\}$$

$$\mathcal{R}_2 = \{z \in \mathbb{C} : |z + 1| \leq 1\}$$

$$\mathcal{R}_3 = \{z \in \mathbb{C} : |z + 2| \leq 2\}$$

sowie diejenigen von  $A^T$  zu

$$\mathcal{C}_1 = \{z \in \mathbb{C} : |z - 4| \leq 1\}$$

$$\mathcal{C}_2 = \{z \in \mathbb{C} : |z + 1| \leq 1\}$$

$$\mathcal{C}_3 = \{z \in \mathbb{C} : |z + 2| \leq 4\}.$$

Die Abschätzung nach Bendixson und die Gerschgorin Kreise sind in Abb. 2.1 dargestellt. Hierbei ist zu beachten, dass man die Menge  $\mathcal{R}_1 \cap \mathcal{C}_3$  durch Anwenden des zweiten Gerschgorin Theorems auf  $A$  und  $A^T$  als in Frage kommenden Bereich für die Eigenwerte von  $A$  ausschließen kann.

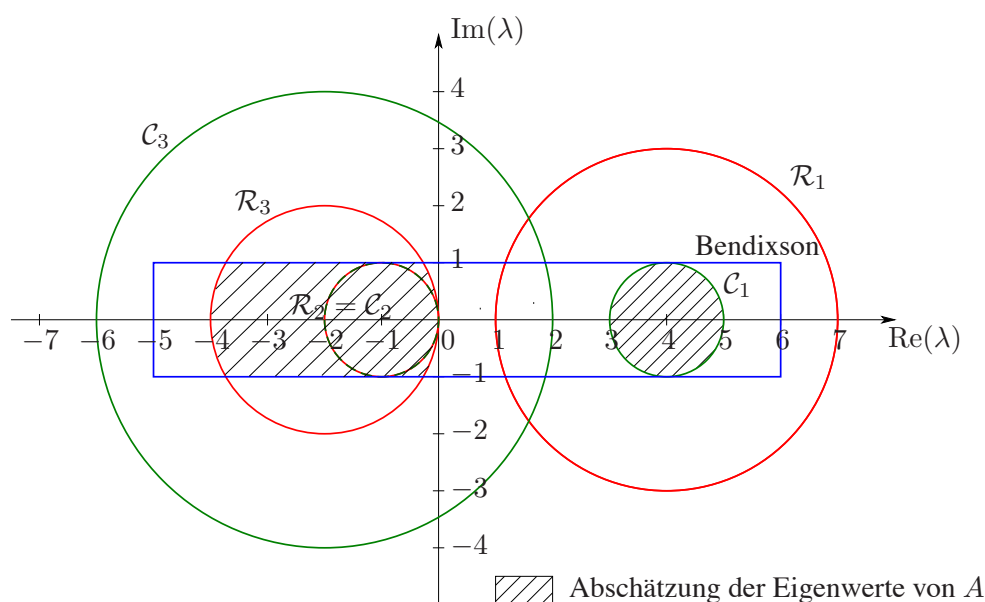


Abb. 2.1: Abschätzung nach Bendixson und Gerschgorin Kreise

## 2.3 POTENZMETHODE

Die Potenzmethode oder *Vektoriteration* ist eine sehr einfache, aber dennoch effektive Methode zur Bestimmung des betragsmäßig größten Eigenwertes.

Um die Grundidee des Verfahrens zu verstehen, nehmen wir zunächst an, dass die Matrix  $A \in \mathbb{C}^{n \times n}$  diagonalisierbar ist. Dann gibt es eine Basis  $v_1, \dots, v_n$  von  $\mathbb{C}^n$  aus Eigenvektoren  $v_i$  von  $A$  mit  $\|v_i\| = 1$ . Des Weiteren seien die Eigenwerte von  $A$  in der Form

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| \quad (2.14)$$

geordnet, wobei  $\lambda_1$  die algebraische Vielfachheit 1 besäße. Gehen wir nun von einem Startvektor  $x^{(0)}$  aus, so lässt sich dieser als Linearkombination der Eigenvektoren  $v_i$  schreiben:

$$x^{(0)} = \sum_{i=1}^n \alpha_i v_i. \quad (2.15)$$

Definiert man nun die Iterierten  $a^{(k)}$ ,  $k \in \mathbb{N}$ , gemäß  $a^{(k)} := A^k x^{(0)}$ , so ergibt sich

$$a^{(k)} = A^k x^{(0)} = A^k \left( \sum_{i=1}^n \alpha_i v_i \right) = \sum_{i=1}^n \alpha_i \lambda_i^k v_i = \lambda_1^k \sum_{i=1}^n \alpha_i \left( \frac{\lambda_i}{\lambda_1} \right)^k v_i. \quad (2.16)$$

Falls der Koeffizient  $\alpha_1$  von Null verschieden ist, d.h.  $x^{(0)}$  nicht auf  $v_1$  senkrecht steht, wird sich dieser Ausdruck dem Summanden mit dem dominanten Eigenwert  $\lambda_1$  annähern, also

$$a^{(k)} = A^k x^{(0)} \approx \lambda_1^k \alpha_1 v_1.$$

Um in der Praxis einen Over- oder Underflow zu vermeiden, normiert man in jedem Iterationsschritt den Iterationsvektor. Insgesamt ergibt sich für die Potenzmethode folgender Algorithmus.

---

### Algorithmus 2.3.1: Potenzmethode

**Input:**  $x^{(0)}$  mit  $v_1^T x^{(0)} \neq 0$  und  $\|x^{(0)}\| = 1$   
**for**  $k = 0, 1, \dots$

$$a^{(k)} = Ax^{(k)}$$

$$\rho^{(k)} = x^{(k)T} a^{(k)} \quad \% \text{ Rayleigh-Quotient}$$

$$x^{(k+1)} = \frac{a^{(k)}}{\|a^{(k)}\|}$$

**end**

---

Wir fassen die bisherigen Ergebnisse der Potenzmethode in einem Satz zusammen.

**Satz 2.3.1** Sei  $\lambda_1$  ein einfacher Eigenwert der diagonalisierbaren Matrix  $A \in \mathbb{C}^{n \times n}$  mit (2.14) und  $x^{(0)} \in \mathbb{C}^n$  sei ein Vektor, der nicht senkrecht auf dem Eigenraum von  $\lambda_1$  steht und  $\|x^{(0)}\| = 1$  erfüllt. Dann konvergiert die Folge  $x^{(k+1)} = a^{(k)} / \|a^{(k)}\|$  mit  $a^{(k)} = Ax^{(k)}$  gegen einen normierten Eigenvektor von  $A$  zum Eigenwert  $\lambda_1$ .

*Beweis.* Wir wissen aus (2.16), dass

$$a^{(k)} = A^k x^{(0)} = \sum_{i=1}^n \alpha_i \lambda_i^k v_i = \alpha_1 \lambda_1^k \underbrace{\left( v_1 + \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} \left( \frac{\lambda_i}{\lambda_1} \right)^k v_i \right)}_{=: z_k}$$

für eine normierte Basis  $v_1, \dots, v_n$  des  $\mathbb{C}^n$  aus Eigenvektoren von  $A$  gilt, wobei  $\alpha_1 \neq 0$  ist, da  $x^{(0)} \not\perp v_1$  vorausgesetzt wurde. Da  $|\lambda_i| < |\lambda_1|$  für alle  $i = 2, \dots, n$  gilt, ist  $\lim_{k \rightarrow \infty} z_k = v_1$  und daher

$$x^{(k)} = \frac{a^{(k)}}{\|a^{(k)}\|} = \frac{z_k}{\|z_k\|} \rightarrow \pm v_1 \quad \text{für } k \rightarrow \infty.$$

□

**Bemerkungen 2.3.2** i) Der Rayleigh-Quotient  $\rho^{(k)}$  in Algorithmus 2.3.1 liefert im Grenzwert den betragsgrößten Eigenwert:

$$\lim_{k \rightarrow \infty} \rho^{(k)} = \lambda_1.$$

ii) Der Beweis des Satzes 2.3.1 lässt sich auch auf diagonalisierbare Matrizen mit eindeutig bestimmten betragsgrößtem Eigenwert  $\lambda_1$ , welcher aber nicht einfach zu sein braucht, d.h.

$$\begin{aligned} \lambda_1 &= \lambda_2 = \dots = \lambda_r \\ |\lambda_1| &= \dots = |\lambda_r| > |\lambda_{r+1}| \geq \dots \geq |\lambda_n| \end{aligned}$$

übertragen. Hierbei muss der Startvektor der Vektoriteration die Bedingung  $\alpha_i \neq 0$  für ein  $i \in \{1, \dots, r\}$  erfüllen. Man beachte, dass für  $r = 1$  ( $\lambda_1$  ist also einfacher Eigenwert) der Grenzvektor  $v_1$  ist und somit nicht von der Wahl von  $x^{(0)}$  abhängt, sofern nur  $\alpha_1 \neq 0$  ist. Ist  $\lambda_1$  ein mehrfacher dominanter Eigenwert,  $r > 1$ , so hängt der gefundene Eigenvektor von den Verhältnissen  $\alpha_1 : \alpha_2 : \dots : \alpha_r$  und damit vom Startvektor  $x^{(0)}$  ab.

iii) Man sieht in (2.16), dass für die Potenzmethode lineare Konvergenz mit Konvergenzfaktor  $|\lambda_2/\lambda_1|$  bzw.  $|\lambda_{r+1}/\lambda_1|$  vorliegt. Das Verfahren konvergiert also umso besser, je mehr die Beträge der Eigenwerte von  $A$  getrennt sind.

Falls  $A$  symmetrisch ist, sind die Eigenvektoren  $v_i$  orthogonal. Mit Hilfe dieser Orthogonalitätseigenschaft kann man zeigen, dass in diesem Fall der Konvergenzfaktor sogar  $|\lambda_2/\lambda_1|^2$  bzw.  $|\lambda_{r+1}/\lambda_1|^2$  beträgt.

iv) Allgemein konvergiert das Verfahren nicht gegen  $\lambda_1$  und einem zu  $\lambda_1$  gehörigen Eigenvektor, sondern gegen  $\lambda_k$  und einem Eigenvektor zu  $\lambda_k$ , sofern in der Zerlegung von  $x^{(0)}$  gilt

$$\alpha_1 = \alpha_2 = \dots = \alpha_{k-1}, \quad \alpha_k \neq 0$$

und es keinen von  $\lambda_k$  verschiedenen Eigenwert gleichen Betrags gibt.

Praktisch konvergiert jedoch auch im Falle  $\alpha_1 = 0$  das Verfahren gegen  $\lambda_1$  und einem zugehörigen Eigenvektor, da infolge von Rundungsfehlern  $\alpha_1^{(1)} \neq 0$  gilt ( $\alpha_1^{(1)}$  der Koeffizient von  $x^{(1)}$ ).

v) Da dieses Verfahren in jeder Iteration nur eine Matrix-Vektor-Multiplikation erfordert, ist der Aufwand der Potenzmethode nach  $k$  Iterationen  $\mathcal{O}(kn^2)$ .

Sei nun  $A \in \mathbb{C}^{n \times n}$  eine nicht diagonalisierbare Matrix mit eindeutig bestimmten betragsgrößtem Eigenwert  $\lambda_1$ , d.h. aus  $|\lambda_1| = |\lambda_i|$  folgt  $\lambda_1 = \lambda_i$ . Ersetzt man die Darstellung (2.15) des Startvektors  $x^{(0)}$  durch eine als Linearkombination von Eigen- und Hauptvektoren von  $A$ , so kann man auf dieselbe Weise wie im diagonalisierbaren Fall zeigen, dass unter analogen Voraussetzungen an  $x^{(0)}$  in der Potenzmethode 2.3.1 der Rayleigh-Quotient  $\rho^{(k)}$  gegen  $\lambda_1$  und  $x^{(k)}$  gegen einen zu  $\lambda_1$  gehörigen Eigenvektor konvergieren.

**Beispiel 2.3.3** Wir betrachten das Eigenwertproblem aus Beispiel 2.0.1 mit  $R = I$ , also  $Ax = \lambda x$  mit  $A \in \mathbb{R}^{(n-1) \times (n-1)}$  wie in (2.3). Die Eigenwerte der Matrix  $A$  lassen sich explizit angeben:

$$\lambda_{n-k} = \frac{4}{h^2} \sin^2\left(\frac{1}{2}k\pi h\right), \quad k = 1, 2, \dots, n-1, \quad h := \frac{1}{n}.$$

Die Nummerierung ist hierbei so gewählt, dass  $\lambda_1 > \lambda_2 > \dots > \lambda_{n-1}$  gilt. Wegen

$$\begin{aligned} \left|\frac{\lambda_2}{\lambda_1}\right| &= \frac{\lambda_2}{\lambda_1} = \frac{\sin^2(\frac{1}{2}(n-2)\pi h)}{\sin^2(\frac{1}{2}(n-1)\pi h)} = \frac{\sin^2(\frac{1}{2}\pi - \pi h)}{\sin^2(\frac{1}{2}\pi - \frac{1}{2}\pi h)} \\ &= \frac{\cos^2(\pi h)}{\cos^2(\frac{1}{2}\pi h)} = \frac{(1 - \frac{1}{2}(\pi h)^2)^2}{(1 - \frac{1}{2}(\frac{1}{2}\pi h)^2)^2} + \mathcal{O}(h^4) = 1 - \frac{3}{4}\pi^2 h^2 + \mathcal{O}(h^4) \end{aligned}$$

ist für  $h \ll 1$  eine sehr langsame Konvergenz mit dem Konvergenzfaktor  $\left|\frac{\lambda_2}{\lambda_1}\right|^2 = 1 - \frac{3}{2}\pi^2 h^2$  (quadratisch, da die Matrix symmetrisch ist) zu erwarten. Dies wird bestätigt von den Ergebnissen in Tabelle 2.1 und Abbildung 2.2, die aus der Anwendung der Vektoriteration auf die Matrix  $A$  mit  $h = \frac{1}{30}$  und dem Startvektor  $x^{(0)} = y^{(0)} / \|y^{(0)}\|_2$ ,  $y^{(0)} = (1, 2, \dots, 29)^T$  resultieren.

| $k$ | $ \rho^{(k)} - \lambda_1 $ | $\frac{ \rho^{(k)} - \lambda_1 }{ \rho^{(k-1)} - \lambda_1 }$ |
|-----|----------------------------|---|
| 1   | 1.79e+3                    | 0.5117  |
| 5   | 4.81e+2                    | 0.8151  |
| 15  | 1.64e+2                    | 0.9319  |
| 50  | 43.60                      | 0.9758  |
| 100 | 17.01                      | 0.9844  |
| 150 | 8.12                       | 0.9857  |
| 200 | 3.90                       | 0.9852  |

Tab. 2.1: Konvergenz der Vektoriteration angewandt auf Beispiel 2.0.1

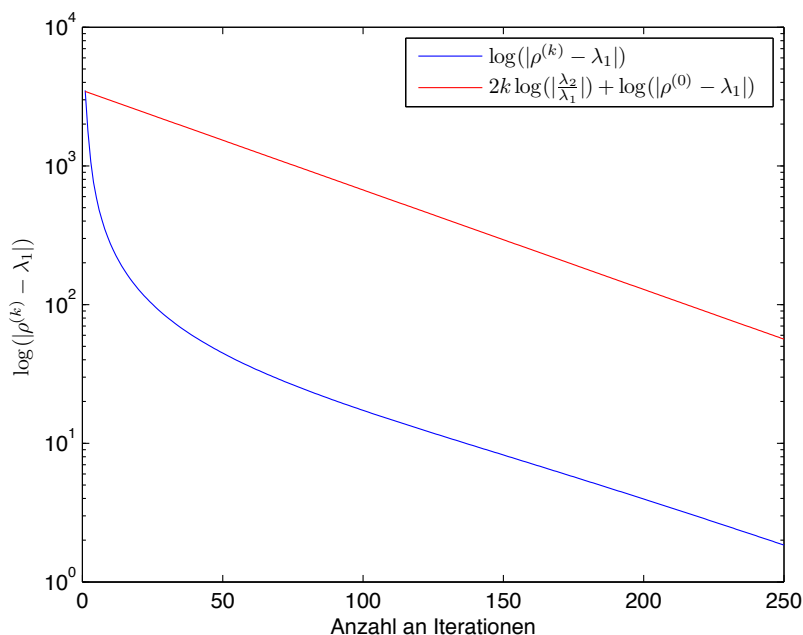


Abb. 2.2: Halblogarithmische Darstellung der Entwicklung des Fehlers

**Beispiel 2.3.4** Es soll der betragsgrößte Eigenwert der Matrix

$$A = \begin{pmatrix} 7 & 4 & 9 & 3 \\ 2 & 2 & 4 & 9 \\ 3 & 1 & 0 & 4 \\ 2 & 7 & 5 & 1 \end{pmatrix}$$

bestimmt werden.  $A$  besitzt das Spektrum  $\sigma(A) = \{15.2806, 4.4616, -2.7330, -7.0092\}$ . Da  $A$  nicht symmetrisch ist, erwarten wir somit eine lineare Konvergenz der Potenzmethode mit Konvergenzfaktor

$$\left| \frac{\lambda_2}{\lambda_1} \right| = \frac{7.0092}{15.2806} \approx 0.4587.$$

Dieses theoretische Ergebnis wird von den konkreten Resultaten der Vektoriteration angewandt auf die Matrix  $A$  mit Startvektor  $x^{(0)} = (0.5, 0.5, 0.5, 0.5)^T$  in Tabelle 2.2 bestätigt.

| $k$ | $ \rho^{(k)} - \lambda_1 $ | $\frac{ \rho^{(k)} - \lambda_1 }{ \rho^{(k-1)} - \lambda_1 }$ |
|-----|----------------------------|---|
| 0   | 0.4694                     | –   |
| 1   | 0.2625                     | 0.5592  |
| 5   | 5.987e-4                   | 0.1095  |
| 10  | 1.916e-5                   | 0.5453  |
| 15  | 3.608e-7                   | 0.4504  |
| 20  | 7.385e-9                   | 0.4596  |
| 25  | 1.498e-10                  | 0.4586  |
| 30  | 3.043e-12                  | 0.4587  |

Tab. 2.2: Konvergenz der Vektoriteration angewandt auf Beispiel 2.3.4

## 2.4 INVERSE ITERATION NACH WIELANDT

Sei nun  $A \in \mathbb{C}^{n \times n}$  regulär. Um den in vielen technischen Anwendungen gesuchten betragskleinsten Eigenwert  $\lambda_n$

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n| > 0$$

zu finden, kann man die Tatsache ausnutzen, dass der betragskleinste Eigenwert von  $A$  der inverse betragsgrößte Eigenwert von  $A^{-1}$  ist, d.h. für die Eigenwerte  $\nu_i$  von  $A^{-1}$  gilt

$$|\nu_n| > |\nu_{n-1}| \geq \dots \geq |\nu_1|, \quad \text{wobei} \quad \nu_i = \frac{1}{\lambda_i}, i = 1, \dots, n.$$

Die Anwendung der Potenzmethode auf  $A^{-1}$  liefert also eine Möglichkeit,  $\nu_n = \frac{1}{\lambda_n}$ , also auch den betragskleinsten Eigenwert von  $A$  zu bestimmen. In jeder Iteration ist dann der Vektor  $a^{(k)} = A^{-1}x^{(k)}$  zu berechnen; dies entspricht dem Lösen des linearen Gleichungssystems  $Aa^{(k)} = x^{(k)}$ . Dies ist die sogenannte inverse Iteration nach Wielandt.

Mit Hilfe einer Verschiebung kann man auch die Berechnung der anderen Eigenwerte erreichen. Hierbei wird vorausgesetzt, dass man eine gute Näherung  $\mu$  für einen Eigenwert  $\lambda_j$  von  $A$  kennt, sodass

$$|\mu - \lambda_j| < |\mu - \lambda_i| \quad \forall i \neq j$$

gilt. Dann hat die Matrix  $(A - \mu I)^{-1}$  den betragsgrößten Eigenwert  $(\lambda_j - \mu)^{-1}$  und die gleichen Eigenvektoren wie  $A$ . Somit liefert die inverse Iteration angewandt auf  $A - \mu I$  eine Approximation zu  $\frac{1}{\lambda_j - \mu}$ , woraus sich  $\lambda_j$  ergibt.

**Algorithmus 2.4.1: Inverse Iteration nach Wielandt mit Spektralverschiebung**

**Input:**  $\mu \approx \lambda_j$ ,  $x^{(0)}$  mit  $v_j^T x^{(0)} \neq 0$  und  $\|x^{(0)}\| = 1$   
**for**  $k = 0, 1, \dots$

**Löse**  $(A - \mu I) a^{(k)} = x^{(k)}$   
 $\rho^{(k)} = (x^{(k)})^T a^{(k)} \quad \% \text{ Rayleigh-Quotient}$   
 $x^{(k+1)} = \frac{a^{(k)}}{\|a^{(k)}\|}$

**end**

**Bemerkung 2.4.1** Pro Iteration muss also ein lineares Gleichungssystem gelöst werden, dessen Koeffizientenmatrix aber konstant ist. Bestimmt man also einmal eine LR-Zerlegung von  $A - \mu I$ , so sind pro Iterationsschritt zwei Dreieckssysteme zu lösen, was einem Aufwand von  $\mathcal{O}(n^2)$  entspricht.

Mit Satz 2.3.1 können wir nun folgern, dass der Rayleigh-Quotient  $\rho^{(k)}$

$$\rho^{(k)} \rightarrow \frac{1}{\lambda_j - \mu} \quad \text{für } k \rightarrow \infty$$

erfüllt. Gemäß der Konvergenzanalyse der Potenzmethode in Abschnitt 2.3 ergibt sich die Konvergenzgeschwindigkeit aus dem Verhältnis zwischen  $\frac{1}{\lambda_j - \mu}$  und dem betragsmäßig zweitgrößten Eigenwert von  $(A - \mu I)^{-1}$ , also durch den Faktor

$$\frac{\max_{i \neq j} \frac{1}{|\lambda_i - \mu|}}{\frac{1}{|\lambda_j - \mu|}} = \frac{\min_{i \neq j} |\lambda_i - \mu|}{|\lambda_j - \mu|} = \frac{|\lambda_j - \mu|}{\min_{i \neq j} |\lambda_i - \mu|}$$

bestimmt wird.

**Bemerkungen 2.4.2** i) Ist  $\mu$  eine gute Schätzung von  $\lambda_j$ , so gilt

$$\frac{|\lambda_j - \mu|}{\min_{i \neq j} |\lambda_i - \mu|} \ll 1$$

und das Verfahren konvergiert in diesem Fall sehr rasch.

ii) Die Kondition von  $A - \mu I$  strebt für „immer besser“ gewähltes  $\mu$  gegen unendlich, die Matrix ist für  $\mu \approx \lambda_j$  fast singular. Daraus entstehen aber keine numerischen Schwierigkeiten, da nur die Richtung des Eigenvektors gesucht wird. Man ersetzt im Gauß-Verfahren ein auftretendes Pivotelement  $\epsilon = 0$  durch die relative Maschinengenauigkeit  $\text{eps}$ .

Durch geeignete Wahl des Spektralverschiebungsparameters  $\mu$  kann man also mit der inversen Vektoriteration 2.4.1 einzelne Eigenwerte und Eigenvektoren der Matrix  $A$  bestimmen. In der Praxis ist aber oft nicht klar, wie man für einen beliebigen Eigenwert  $\lambda_j$  diesen Parameter  $\mu$  geeignet wählen kann.

Die Konvergenzgeschwindigkeit der Methode kann man noch erheblich verbessern, wenn man den Parameter  $\mu$  nach jedem Schritt auf die aktuelle Annäherung  $\lambda^{(k)} := \frac{1}{\rho^{(k)}} + \mu$  von  $\lambda_j$  setzt. Da die LR-Zerlegung dann aber in jedem Schritt neu berechnet werden muss, steigt damit der Rechenaufwand sehr stark an.

**Beispiel 2.4.3** Wir betrachten die Matrix aus Beispiel 2.3.4 und wenden zur Berechnung des Eigenwerts  $\lambda_3 = 4.4616$  dieser Matrix den Algorithmus 2.4.1 mit  $\mu = 3.5$  und Startvektor  $x^{(0)} = (0.5, 0.5, 0.5, 0.5)^T$  an. Die Resultate in Tabelle 2.3 bestätigen die obigen theoretischen Betrachtungen zur Konvergenz des Wielandt-Verfahrens, nach denen wir lineare Konvergenz mit einem Konvergenzfaktor in der Größenordnung

$$\frac{|\lambda_3 - 3.5|}{\min_{i \neq 3} |\lambda_i - 3.5|} = \frac{|4.4616 - 3.5|}{|-2.7330 - 3.5|} = \frac{0.9616}{6.2330} \approx 0.1543.$$

erwarten.

| $k$ | $ \lambda^{(k)} - \lambda_3 $ | $\frac{ \lambda^{(k)} - \lambda_3 }{ \lambda^{(k-1)} - \lambda_3 }$ |
|-----|-------------------------------|---|
| 0   | 0.8734                        | –   |
| 1   | 3.716e-3                      | 0.0042  |
| 2   | 2.737e-3                      | 0.7364  |
| 3   | 1.902e-4                      | 0.0695  |
| 4   | 4.580e-5                      | 0.2407  |
| 5   | 5.665e-6                      | 0.1237  |
| 6   | 9.805e-7                      | 0.1731  |
| 7   | 1.419e-7                      | 0.1448  |
| 8   | 2.260e-8                      | 0.1592  |
| 9   | 3.424e-9                      | 0.1515  |
| 10  | 5.329e-10                     | 0.1556  |

Tab. 2.3: Konvergenz des Wielandt-Verfahrens mit  $\mu = 3.5$  angewandt auf Beispiel 2.3.4

Für die inverse Vektoriteration, wobei man den Parameter  $\mu$  nach jedem Schritt auf die jeweils aktuelle Annäherung  $\lambda^{(k)}$  vom  $\lambda_3$  setzt,

$$\mu_0 = 3.5 \quad \mu_k = \lambda^{(k-1)} \quad \text{für } k \geq 1$$

sind einige Ergebnisse in Tabelle 2.4 dargestellt. Die Resultate zeigen, dass die Konvergenzgeschwindigkeit wesentlich schneller (genauer: quadratisch statt linear) ist.

| $k$ | $ \lambda^{(k)} - \lambda_3 $ | $\lambda^{(k)}$ |
|-----|-------------------------------|-----------------|
| 0   | 0.8735                        | 5.3351          |
| 1   | 7.245e-3                      | 4.4544          |
| 2   | 1.490e-5                      | 4.4616          |
| 3   | 5.819e-12                     | 4.4616          |

Tab. 2.4: Konvergenz des Wielandt-Verfahrens mit  $\mu_k = \lambda^{(k-1)}$  angewandt auf Beispiel 2.3.4

**Beispiel 2.4.4** Gegeben sei die Matrix

$$A = \begin{pmatrix} -1 & 3 \\ -2 & 4 \end{pmatrix}$$

mit dem Spektrum  $\sigma(A) = \{1, 2\}$ . Gehen wir von einer Approximation  $\mu = 1 - \epsilon$  von  $\lambda_2 = 1$  mit  $0 < |\epsilon| \ll 1$  aus, so ist die Matrix

$$(A - \mu I) = \begin{pmatrix} -2 + \epsilon & 3 \\ -2 & 3 + \epsilon \end{pmatrix}$$

fast singular; ihre Inverse ist gegeben durch

$$(A - \mu I)^{-1} = \frac{1}{\epsilon^2 + \epsilon} \begin{pmatrix} 3 + \epsilon & -3 \\ 2 & -2 + \epsilon \end{pmatrix}.$$

Da sich der Faktor  $\frac{1}{\epsilon^2 + \epsilon}$  bei der Normierung  $x^{(k+1)} = \frac{a^{(k)}}{\|a^{(k)}\|}$  in der Wielandt-Iteration herauskürzt, ist die Berechnung der Richtung einer Lösung von  $(A - \mu I)a^{(k)} = x^k$  gut konditioniert.

## 2.5 QR-VERFAHREN

Die in den vorherigen Abschnitten vorgestellten Methoden der Vektoriteration und der inversen Iteration nach Wielandt haben den schwerwiegenden Nachteil, dass man mit ihnen nur bestimmte Eigenwerte bestimmen kann. So liefert die Potenzmethode nur den betragsmäßig größten Eigenwert und bei der Wielandt-Iteration benötigt man zur Bestimmung eines Eigenwertes  $\lambda_i$  zunächst einen geeignet gewählten Parameter  $\mu_i \approx \lambda_i$ . In diesem Abschnitt werden wir uns nun einem effizienteren Verfahren zuwenden, mit dessen Hilfe man nicht nur einen, sondern *gleichzeitig alle* Eigenwerte einer Matrix  $A \in \mathbb{R}^{n \times n}$  approximieren kann – dem QR-Verfahren.

### 2.5.1 Das Basisverfahren der QR-Iteration

Könnte man die Schur-Zerlegung  $Q^H A Q = R$  aus Satz 2.1.6 einer Matrix  $A \in \mathbb{R}^{n \times n}$  auf direkte Weise, also mit einer endlichen Anzahl an Operationen berechnen, so hätte man das Eigenwertproblem gelöst, da die Eigenwerte  $\lambda_i(A)$  dann gerade durch die Diagonalelemente  $r_{ii}$  der oberen Dreiecksmatrix  $R$  gegeben wären. Leider ist die direkte Bestimmung der Matrix  $Q$  für  $n \geq 5$  nicht möglich – dies folgt aus dem Abelschen Theorem (vgl. [Quarteroni et. al., Übung 8, Seite 258]). Daher kann das Eigenwertproblem nur durch den Übergang zu iterativen Methoden gelöst werden. Den Basisalgorithmus für das Verfahren, welches wir im Weiteren vorstellen und untersuchen möchten, liefert die *QR-Iteration*.

---

#### Algorithmus 2.5.1: QR-Basisiteration

```

Input:  $A \in \mathbb{R}^{n \times n}$ 
Setze  $A^{(0)} = A$ .
for  $k = 1, 2, \dots$ 
     $A^{(k-1)} = Q_k R_k$     % Bestimme QR-Zerlegung
     $A^{(k)} = R_k Q_k$ 
end

```

---

Pro Iterationsschritt fällt also der Aufwand einer QR-Zerlegung  $\mathcal{O}(n^3)$  und einer Matrixmultiplikation  $\mathcal{O}(n^2)$  an. Insgesamt benötigt man also für dieses Verfahren  $\mathcal{O}(k_{\max} n^3)$  Operationen, wobei  $k_{\max}$  die Anzahl der Iterationen im Verfahren ist. Hierbei ist zu beachten, dass für die Iterierten  $A^{(k)}$  folgende Eigenschaften erfüllt sind.



**Lemma 2.5.1** *Es sei  $A \in \mathbb{R}^{n \times n}$ , sowie  $A^{(k)}$  für  $k \in \mathbb{N}_0$  wie in Algorithmus 2.5.1 definiert. Dann gilt:*

- i) *Die Matrizen  $A^{(k)}$  sind orthogonal ähnlich zu  $A$ .*
- ii) *Ist  $A$  symmetrisch, so sind es auch die Matrizen  $A^{(k)}$ .*
- iii) *Ist  $A$  symmetrisch und tridiagonal, so haben auch die Matrizen  $A^{(k)}$  diese Gestalt.*

*Beweis.* ad i) Sei  $A \in \mathbb{R}^{n \times n}$ . Mit den Bezeichnungen aus Algorithmus 2.5.1 gilt dann

$$\begin{aligned} A^{(k)} &= R_k Q_k = Q_k^T Q_k R_k Q_k = Q_k^T A^{(k-1)} Q_k = \dots \\ &= Q_k^T Q_{k-1}^T \dots Q_1^T A^{(0)} Q_1 \dots Q_k = (Q_1 \dots Q_k)^T A (Q_1 \dots Q_k) \end{aligned}$$

ad ii) Sei  $A \in \mathbb{R}^{n \times n}$  nun zusätzlich symmetrisch. Mit i) folgt, dass es für alle  $k \in \mathbb{N}_0$  eine orthogonale Matrix  $P_k \in \mathbb{R}^{n \times n}$  gibt, sodass

$$A^{(k)} = P_k^T A P_k$$

gilt. Daraus folgt

$$(A^{(k)})^T = (P_k^T A P_k)^T = P_k^T A^T P_k = P_k^T A P_k = A^{(k)}.$$

Für den Beweis von iii) sei auf [Deuffhard/Hohmann, Lemma 5.9] verwiesen.  $\square$

Für die QR-Basisiteration erhalten wir folgendes Konvergenzresultat, das Wilkinson<sup>1</sup> im Jahr 1965 in seiner Arbeit [Wilkinson65] beschrieben hat.

**Satz 2.5.2 (Wilkinson)** *Es sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch mit Eigenwerten*

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$$

*und  $A^{(k)}$ ,  $Q_k$  sowie  $R_k$  seien wie in Algorithmus 2.5.1 definiert. Dann gilt:*

- i)  $\lim_{k \rightarrow \infty} Q_k = I,$
- ii)  $\lim_{k \rightarrow \infty} R_k = \text{diag}(\lambda_1, \dots, \lambda_n),$
- iii)  $a_{ij}^{(k)} = \mathcal{O}\left(\left|\frac{\lambda_i}{\lambda_j}\right|^k\right)$  für  $i > j$ , wobei  $A^{(k)} = (a_{ij}^{(k)})_{i,j=1}^n.$

**Bemerkungen 2.5.3** i) *Sofern  $A$  keine symmetrische Matrix ist, aber die Eigenwerte von  $A$  immer noch getrennt sind, d.h.*

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|,$$

*kann man zeigen, dass die Folge  $A^{(k)}$  anstatt gegen eine Diagonal- gegen eine Dreiecksmatrix konvergiert:*

$$\lim_{k \rightarrow \infty} A^{(k)} = \begin{pmatrix} \lambda_1 & * & * & \dots & * \\ 0 & \lambda_2 & * & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}.$$

<sup>1</sup>James H. Wilkinson, \* 27. September 1919 in Strood, Kent; † 5. Oktober 1986 in London, war ein britischer Mathematiker, der die numerische Mathematik vor allem durch Arbeiten zur Rückwärtsanalyse von Rundungsfehlern bereichert hat. 1970 wurde ihm der Turing-Preis verliehen.

ii) Haben wir nun ein Matrix  $A \in \mathbb{R}^{n \times n}$  mit einem Paar konjugiert komplexer Eigenwerte, also  $\lambda_{r+1} = \bar{\lambda}_r$  für ein  $r \in \{1, \dots, n-1\}$  und somit

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_r| = |\lambda_{r+1}| > \dots > |\lambda_n|,$$

so konvergiert die Folge  $A^{(k)}$  gegen eine Matrix der Form

$$\lim_{k \rightarrow \infty} A^{(k)} = \begin{pmatrix} \lambda_1 & * & \dots & \dots & \dots & \dots & \dots & * \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & \lambda_{r-1} & * & * & * & & \vdots \\ \vdots & & 0 & a_{rr}^{(k)} & a_{r,r+1}^{(k)} & * & & \vdots \\ \vdots & & 0 & a_{r+1,r}^{(k)} & a_{r+1,r+1}^{(k)} & * & & \vdots \\ \vdots & & 0 & 0 & 0 & \lambda_{r+2} & \ddots & * \\ \vdots & & & & & \ddots & \ddots & * \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

wobei die Matrix  $\begin{pmatrix} a_{rr}^{(k)} & a_{r,r+1}^{(k)} \\ a_{r+1,r}^{(k)} & a_{r+1,r+1}^{(k)} \end{pmatrix}$  für  $k \rightarrow \infty$  im Allgemeinen divergiert, aber deren Eigenwerte konvergieren gegen  $\lambda_r$  und  $\bar{\lambda}_r = \lambda_{r+1}$ .

*Beweis von Satz 2.5.2.* Wir zeigen zunächst induktiv, dass für  $k \in \mathbb{N}$  gilt:

$$A^k = \underbrace{Q_1 \dots Q_k}_{=: P_k} \underbrace{R_1 \dots R_k}_{=: U_k}.$$

Für  $k = 1$  ist die Behauptung klar. Aus der Konstruktion der  $A^{(k)}$  folgt wie im Beweis zu Lemma 2.5.1, dass

$$Q_{k+1} R_{k+1} = A^{(k)} = Q_1^T \dots Q_k^T A Q_1 \dots Q_k = P_k^T A P_k$$

gilt und somit auch der Induktionsschritt

$$A^{k+1} = A A^k \stackrel{IH}{=} A P_k U_k = P_k A^{(k)} U_k = P_k Q_{k+1} R_{k+1} U_k = P_{k+1} U_{k+1}.$$

Da  $P_k \in \mathbb{R}^{n \times n}$  orthogonal ist und  $U_k$  eine obere Dreiecksmatrix, können wir die QR-Zerlegung  $A^k = P_k U_k$  von  $A^k$  durch die QR-Zerlegung der  $A^{(1)}, \dots, A^{(k)}$  ausdrücken. Ferner ist  $A$  diagonalisierbar, da  $A$  symmetrisch ist, und daher folgt

$$A^k = Q^T \Lambda^k Q \quad \text{mit } \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Da man durch eine geeignete Permutation von  $A$  immer erreichen kann, dass  $Q$  eine LR-Zerlegung  $Q = LR$  mit einer unipotenten unteren Dreiecksmatrix  $L$  und einer oberen Dreiecksmatrix  $R$  besitzt, nehmen dies wir im Folgenden o.B.d.A. an. Damit gilt

$$A^k = Q^T \Lambda^k Q = Q^T \Lambda^k L R = Q^T (\Lambda^k L \Lambda^{-k}) (\Lambda^k R).$$

Für die unipotente untere Dreiecksmatrix  $(\Lambda^k L \Lambda^{-k})$  gilt

$$(\Lambda^k L \Lambda^{-k})_{ij} = l_{ij} \left( \frac{\lambda_i}{\lambda_j} \right)^k,$$

insbesondere verschwinden alle Nicht-Diagonalelemente für  $k \rightarrow \infty$ , d.h.

$$(\Lambda^k L \Lambda^{-k}) = I + E_k \quad \text{mit } E_k \rightarrow 0 \quad \text{für } k \rightarrow \infty.$$

Damit folgt mit einer QR-Zerlegung von  $I + E_k = \tilde{Q}_k \tilde{R}_k$ , dass

$$A^k = Q^T (I + E_k) \Lambda^k R = (Q^T \tilde{Q}_k) (\tilde{R}_k \Lambda^k R), \quad (2.17)$$

wobei alle Diagonaleinträge von  $\tilde{R}_k$  positiv gewählt seien, was diese Zerlegung eindeutig macht. Aus  $\lim_{k \rightarrow \infty} E_k = 0$  folgt

$$\lim_{k \rightarrow \infty} \tilde{Q}_k = I, \quad \lim_{k \rightarrow \infty} \tilde{R}_k = I.$$

Wir haben in (2.17) eine weitere QR-Zerlegung von  $A^k$  gefunden, daher gilt bis auf Vorzeichen der Diagonale

$$P_k = Q^T \tilde{Q}_k, \quad U_k = \tilde{R}_k \Lambda^k R.$$

Für den Grenzübergang  $k \rightarrow \infty$  folgt dann schließlich

$$\begin{aligned} Q_k &= P_{k-1}^T P_k = \tilde{Q}_{k-1}^T Q Q^T \tilde{Q}_k = \tilde{Q}_{k-1}^T \tilde{Q}_k \rightarrow I \\ R_k &= U_k U_{k-1}^{-1} = \tilde{R}_k \Lambda^k R R^{-1} \Lambda^{-(k-1)} \tilde{R}_{k-1}^{-1} = \tilde{R}_k \Lambda \tilde{R}_{k-1}^{-1} \rightarrow \Lambda \end{aligned}$$

und

$$\lim_{k \rightarrow \infty} A^{(k)} = \lim_{k \rightarrow \infty} Q_k R_k = I \cdot \Lambda = \Lambda.$$

□

Der QR-Algorithmus 2.5.1 in seiner Grundform ist sehr aufwendig. Pro Iterationsschritt  $A^{(k)} \rightarrow A^{(k+1)}$  benötigt man bei vollbesetzten Matrizen  $\mathcal{O}(n^3)$  Operationen. Zudem besitzt das Verfahren in dieser Form auch den Nachteil, den schon die inverse Iteration nach Wielandt aufwies: Sind einige Eigenwerte von  $A$  betragsmäßig nur schlecht getrennt, d.h.  $|\lambda_j / \lambda_k| \approx 1$  für  $j \neq k$ , so ist die Konvergenz des Verfahrens nur sehr langsam, vgl. Satz 2.5.2 iii).

Im Folgenden werden für beide Probleme Lösungsansätze vorgestellt. Eine Möglichkeit, den Aufwand des Verfahrens zu verringern, besteht darin, die Matrix  $A$  auf *einfachere Gestalt* zu transformieren und auf diese transformierte Matrix die QR-Iteration mit einem geringeren Aufwand anzuwenden. Sogenannte *Shift-Techniken* oder *Verschiebungen* sorgen schließlich für eine Verbesserung des Konvergenzverhaltens des Verfahrens.

## 2.5.2 Transformation auf Hessenberg-Form

## 2.5.3 Das QR-Verfahren für Matrizen in Hessenberg-Form

## 2.5.4 Das QR-Verfahren mit Verschiebungen

# 2.6 VERFAHREN FÜR SYMMETRISCHE MATRIZEN

## 2.6.1 Die Methode der Sturmschen Ketten

## 2.6.2 Das Lanczos-Verfahren

# 2.7 SINGULÄRWERTZERLEGUNG