# Kapitel 6

# Eigenwertprobleme

## 6.1 Eigenwertprobleme bei Matrizen

Eigenwertprobleme sind neben den linearen Gleichungssystemen die zweite Grundaufgabe der numerischen linearen Algebra. Wir wollen in diesem Abschnitt zunächst einige Tatsachen zusammenstellen.

**Definition 6.1.1** *Sei $A$ eine komplexe $(n, n)$-Matrix. $\lambda \in \mathbb{C}$ heißt Eigenwert von $A$, wenn es $x \in \mathbb{C}^n$, $x \neq 0$ gibt mit $Ax = \lambda x$. $x$ heißt dann Eigenvektor von $A$ zum Eigenwert $\lambda$.*

Als notwendige und hinreichende Bedingung dafür, daß $\lambda$ Eigenwert von $A$ ist, hat man also

$$\varphi(\lambda) = \det\ (\lambda I - A) = 0\ .$$

$\varphi(\lambda)$ heißt "charakteristisches Polynom von $A$". $\varphi(\lambda)$ ist ein Polynom genau vom Grade $n$ in $\lambda$:

$$\varphi(\lambda) = \lambda^n - \left(\sum_{i=1}^{n} a_{ii}\right) \lambda^{n-1} + \cdots + (-1)^n \det\ (A)\ .$$

**Definition 6.1.2** *Jedem Eigenwert $\lambda$ von $A$ ordnen wir zwei Vielfachheiten zu:*

*Seine algebraische Vielfachheit $\sigma(\lambda)$ = Vielfachheit von $x$ als Nullstelle von $\varphi(\lambda)$.*

*Seine geometrische Vielfachheit $\rho(\lambda)$ = Anzahl der linear unabhängigen Eigenvektor zu $\lambda$.*

Sind also $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwert von $A$ und sind $\sigma_k = \sigma(\lambda_k)$ ihre algebraischen Vielfachheiten, so gilt

$$\varphi(\lambda) = \prod_{k=1}^{m} (\lambda - \lambda_k)^{\sigma_k} \ , \quad \sum_{k=1}^{m} \sigma_k = n \ .$$

Für die geometrischen Vielfachheiten $\rho_k = \rho(\lambda_k)$ gilt nur

$$\sum_{k=1}^{m} \rho_k \leq n \ .$$

**Definition 6.1.3** *Die $(n, n)$-Matrizen $A$, $B$ heißen ähnlich, wenn es eine nichtsinguläre $(n, n)$-Matrix $x$ gibt mit*

$$A = XBX^{-1} \ .$$

**Satz 6.1.1** *Seien $A$, $B$ ähnlich. Dann haben $A$, $B$ die gleichen Eigenwerte mit übereinstimmenden algebraischen und geometrischen Vielfachheiten.*

**Beweis:** Sei $A = XBX^{-1}$. Dann ist

$$\begin{aligned} \det\,(\lambda I - A) &= \det\,(X(\lambda I - B)X^{-1}) \\ &= \det\,(X)\det(\lambda I - B)\det\,(X^{-1}) \\ &= \det\,(\lambda I - B) \ . \end{aligned}$$

Die charakteristischen Polynome stimmen also überein, also auch die Eigenwerte samt ihrer algebraischen Vielfachheiten. Ist nun $\lambda$ ein Eigenwert von $A$ der geometrischen Vielfachheit $\rho$, so gibt es $\rho$ l.u. Eigenvektoren $x_1, \ldots, x_\rho$ zu $\lambda$, also

$$Ax_k = \lambda x_k \ , \quad k = 1, \ldots, \rho \ .$$

Mit $y_k = X^{-1}x_k$ gilt

$$\begin{aligned} By_k &= X^{-1}AX\,X^{-1}x_k = X^{-1}Ax_k = \lambda X^{-1}x_k \\ &= \lambda y_k\ , \end{aligned}$$

also sind $y_1, \ldots, y_\rho$ l.u. Eigenvektoren von $B$ zu $\lambda$. Die geometrische Vielfachheit von $\lambda$ als Eigenwert von $A$ ist also nicht größer als die geometrische Vielfachheit von $\lambda$ als Eigenwert von $B$. Da die Voraussetzungen in $A$, $B$ symmetrisch sind, müssen die geometrischen Vielfachheiten übereinstimmen.

□

**Beispiele:**

1) $A = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$. Dann ist $\det(\lambda I - A) = \prod\limits_{k=1}^{n}(\lambda - d_k)$, also $\lambda_k = d_k$ mit Eigenvektor $e_k = k$-tem Einheitsvektor. Offenbar ist

$$\rho(\lambda_k) = \sigma(\lambda_k)\ , \quad k = 1, \ldots, n\ .$$

2) $A = XDX^{-1}$ mit $D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$. Nach Satz 6.1.1 und Beispiel 1) ist $\lambda_k = d_k$, $\rho(\lambda_k) = \sigma(\lambda_k)$, $k = 1, \ldots, n$. Dem Beweis von Satz 6.1.1 und Beispiel 1) entnimmt man die Eigenvektoren $x_k = Xe_k = k$-te Spalte von $X$. Matrizen dieser Art, welche also ähnlich zu einer Diagonalmatrix sind, nennt man diagonalisierbar.

3) $J(\mu) = \begin{pmatrix} \mu & 1 & & \\ & \mu & \ddots & \\ & & \ddots & 1 \\ & & & \mu \end{pmatrix}$. Es ist $\det(\lambda I - J(\mu)) = (\lambda - \mu)^n$, also ist $\lambda = \mu$ der einzige Eigenwert von $J(\mu)$, und er hat die algebraische

Vielfachheit $n$. Ist $x$ ein Eigenvektor zum Eigenwert $\mu$ von $J(\mu)$, so gilt

$$(J(\mu)-\mu I)x = \begin{pmatrix} 0 & 1 & & \mathrm{O} \\ & & \ddots & \\ & & \ddots & 1 \\ \mathrm{O} & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ 0 \end{pmatrix} = 0 \,,$$

und der einzige l.u. Eigenvektor ist $x = e_1$. Also ist die geometrische Vielfachheit von $\mu$ für $n > 1$ verschieden von seiner algebraischen Vielfachheit, nämlich 1.

Bis auf Ähnlichkeiten sind die Matrizen $J(\mu)$ bereits die allgemeinsten Matrizen, soweit das Eigenwertproblem betroffen ist. Es gilt nämlich der

**Satz 6.1.2** *(Jordan'sche Normalform). Jede komplexe $(n,n)$-Matrix ist ähnlich einer Matrix*

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} , \quad J_\ell = \begin{pmatrix} \lambda_\ell & 1 & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_\ell \end{pmatrix} .$$

*Die $J_\ell$ sind bis auf die Reihenfolge eindeutig bestimmt.*

**Beweis:** Siehe etwa F. Lorenz, Lineare Algebra II, Kap. IX, §4.

**Bemerkungen zu Satz 1.2:**

1) Jedes $\lambda_\ell$, $\ell = 1,\ldots,r$ ist Eigenwert.

2) $\rho(\lambda)$ = Anzahl der $J_\ell$ mit $\lambda$ auf der Hauptdiagonalen. Die $\lambda_\ell$ sind also die nach ihrer geometrischen Vielfachheit gezählten Eigenwerte.

3) $\sigma(\lambda)$ = Summe der Längen sämtlicher $J_\ell$ mit $\lambda$ auf der Hauptdiagonalen, denn

$$\begin{aligned} \det\ (\lambda I - J) &= \prod_{\ell=1}^{r} \det\ (\lambda I - J_\ell) = \prod_{\ell=1}^{r} (\lambda - \lambda_\ell)^{\nu_\ell} \,, \\ \nu_\ell &= \text{Länge } (J_\ell) \,. \end{aligned}$$

**Beispiel:**

$$J = \begin{pmatrix} 1 &&&&&&& \\ & 1 &&&&&& \\ && 2 & 1 &&&& \\ &&& 2 &&&& \\ &&&& 2 &&& \\ &&&&& i & 1 & \\ &&&&&& i & 1 \\ &&&&&&& i \end{pmatrix}$$

| $\lambda$ | $\rho(\lambda)$ | $\sigma(\lambda)$ |
|---|---|---|
| 1 | 2 | 2 |
| 2 | 2 | 3 |
| $i$ | 1 | 3 |

Sei nun $A$ eine Matrix mit Jordan'scher Normalform $J$, also

$$A = XJX^{-1}\ , \quad J = \begin{pmatrix} J_1 && \\ & \ddots & \\ && J_r \end{pmatrix}$$

mit einer geeigneten Matrix $X$. Wir wollen uns die Bedeutung von $X$ klarmachen. Sei $\nu_\ell$ die Länge von $J_\ell$. Wir spalten $X$ auf entsprechend der Aufspaltung von $J$:

$$X = (X_1, \ldots, X_r)$$

mit $(n, \nu_\ell)$-Matrizen $X_\ell$. Dann haben wir

$$AX_\ell = X_\ell J_\ell\ , \quad \ell = 1, \ldots, r\ .$$

Insbesondere ist der von den Spalten von $X_\ell$ aufgespannte Teilraum ein invarianter Unterraum von $A$. Wir untersuchen diese Gleichung für ein $\ell$ und setzen $\nu = \nu_\ell$, $\lambda = \lambda_\ell$. Seien $x_1, \ldots, x_\nu$ die Spalten von $X_\ell$, also $X_\ell = (x_1, \ldots, x_\nu)$. Dann lautet die Gleichung

$$(Ax_1, \ldots, Ax_\nu) = (\lambda x_1, \lambda x_2 + x_1, \ldots, \lambda x_\nu + x_{\nu-1})$$

oder

$$\begin{aligned} Ax_1 &= \lambda x_1 \qquad , \\ Ax_i &= \lambda x_i + x_{i-1}\ , \quad i = 2, \ldots, \nu\ . \end{aligned}$$

$x_1$ ist also Eigenvektor zum Eigenwert $\lambda$, wie wir schon aus Beispiel 3 und Satz 6.1.1 wissen. Für die weiteren Vektoren $x_i$ gilt

$$(A - \lambda I)x_i = x_{i-1}\ , \quad i = 2, \ldots, \nu\ ,$$

also

$$(A-\lambda I)^{i-1}x_i = x_1 \ , \quad (A-\lambda)^i x_i = 0 \ .$$

**Definition 6.1.4** *Ein Vektor mit* $(A-\lambda I)^{i-1}x \neq 0$, $(A-\lambda I)^i x = 0$ *heißt Hauptvektor der Stufe* $i$ *von* $A$ *zum Eigenwert* $\lambda$. *Der von allen Hauptvektoren zu einem Eigenwert* $\lambda$ *von* $A$ *aufgespannte Teilraum heißt invarianter Unterraum von* $A$ *zum Eigenwert* $\lambda$.

**Bemerkungen:**

1. Hauptvektoren der Stufe 1 sind gerade die Eigenvektoren, der von ihnen aufgespannte Teilraum heißt Eigenraum zu $\lambda$.
2. Die algebraische Vielfachheit eines Eigenwertes ist gleich der Dimension des zugehörigen invarianten Unterraumes.
3. Eine komplexe $(n,n)$-Matrix besitzt $n$ l.u. Hauptvektoren, nämlich die Spalten einer Matrix $X$, welche sie auf Jordan-Form transformiert.

**Definition 6.1.5** *Eine* $(n,n)$*-Matrix* $A$ *heißt hermitesch, wenn* $A = A^*$ *mit* $A^* = \overline{A}^T$.

**Beispiel:** Folgende Matrizen sind hermitesch:

$$\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix} \ , \quad \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} .$$

Insbesondere sind reell-symmetrische Matrizen hermitesch.

**Satz 6.1.3** *Sei* $A$ *hermitesch. Dann sind alle Eigenwerte von* $A$ *reell.* $A$ *besitzt* $n$ *l.u. paarweise orthonormale Eigenvektoren.*

**Bemerkung:** Die Jordan'sche Normalform einer hermiteschen Matrix ist also eine reelle Diagonalmatrix. Die Matrix $X$ kann unitär gewählt werden, also $XX^* = I$. $A$ heißt positiv definit, wenn alle Eigenwerte $> 0$ sind. $(Ax, x) = 1$ stellt dann ein Ellipsoid dar mit Halbachsen $1/\sqrt{\lambda_\ell}$ in Richtung des $\ell$-ten Eigenvektors.

**Beweis:**

1) Realität der Eigenwerte. Ist $Ax = \lambda x$, so ist $(x, Ax) = \lambda(x, x)$ und $(x, x) > 0$ genügt es zu zeigen, daß $(x, Ax)$ reell ist. Dies folgt aus

$$\overline{(x, Ax)} = \overline{(A^* x, x)} = \overline{(Ax, x)} = (x, Ax) \ .$$

2) Orthonormalität der Eigenvektoren.

   Sind $x_1$, $x_2$ Eigenvektoren zu den verschiedenen Eigenwerten $\lambda_1$, $\lambda_2$, so gilt

$$0 = (Ax_1, x_2) - (Ax_2, x_1) = (\lambda_1 - \lambda_2)(x_1, x_2) \ ,$$

   also sind $x_1$, $x_2$ orthogonal. Sind $x_1, \ldots, x_2$ die Eigenvektoren zu dem Eigenwert $\lambda$, so kann man diese orthonormalisieren.

3) Es genügt zu zeigen, daß keine Hauptvektoren der Stufe 2 auftreten können. Ist $x$ ein solcher, so ist

$$((A - \lambda I)x, (A - \lambda I)x) = (x, (A - \lambda I)^2 x) = 0 \ ,$$

   also $(A - \lambda I)x = 0$, im Widerspruch zu $(A - \lambda I)x \neq 0$.

□

## 6.2 Die Potenzmethode

Sei $A$ eine komplexe $(n, n)$-Matrix. Wir wollen die Eigenwerte $\lambda_i$ von $A$ berechnen. Das einfachste Verfahren ist die Potenzmethode. Ausgehend von einem Vektor $x^{(0)}$ bildet sie der Reihe nach die Vektoren

$$x^{(k+1)} = Ax^{(k)} = A^{k+1} x^{(0)} \ , \qquad k = 0, 1, \ldots \ .$$

Wir analysieren die Potenzmethode zunächst in dem Fall

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n| \; .$$

Dann hat $A$ $n$ l.u. Eigenvektoren $x_1, \ldots, x_n$, und es gilt

$$\begin{aligned}
x^{(0)} &= \sum_{i=1}^{n} c_i x_i \quad , \\
x^{(k)} &= A^k x^{(0)} = \sum_{i=1}^{n} c_i A^k x_i = \sum_{i=1}^{n} c_i \lambda_i^k x_i \\
&= \lambda_1^k (c_1 x_1 + r_k) \;\; , \\
r_k &= \sum_{i=2}^{n} c_i \left( \frac{\lambda_i}{\lambda_1} \right)^k x_i \;\; .
\end{aligned}$$

Offenbar geht $r_k \to 0$ mit $k \to \infty$, und zwar gilt

$$\|r_k\| = 0 \left( \left( \frac{\lambda_2}{\lambda_1} \right)^k \right) \quad .$$

Zur Berechnung von $\lambda_1$ wählt man einen komplexen Vektor $d$ und bildet

$$(x^{(k)}, d) = \lambda_1^k (c_1 (x_1, d) + (r_k, d)) \; .$$

Ist $c_1(x_1, d) \neq 0$, so gilt für $k \to \infty$

$$\frac{(x^{(k+1)}, d)}{(x^{(k)}, d)} = \lambda_1 \frac{c_1(x_1, d) + (r_{k+1}, d)}{c_1(x_1, d) + (r_k, d)} \to \lambda_1 \; .$$

Genauer gilt

$$\frac{(x^{(k+1)}, d)}{(x^{(k)}, d)} = \lambda_1 + \; \mathrm{O} \; \left( \left( \frac{\lambda_2}{\lambda_1} \right)^k \right) \quad ,$$

d.h. die Konvergenzgeschwindigkeit hängt von $\left| \frac{\lambda_2}{\lambda_1} \right|$ ab.

Einen Eigenvektor $x_1$ zu $\lambda_1$ bekommt man als Grenzwert der Folge $x^{(k)}/x_j^{(k)}$ für geeignet gewähltes $j$ ($j$-te Komponente von $x_1$ nicht 0!).

**Beispiel:**

$$A = \begin{pmatrix} 90 & 231 & 70 \\ 110 & 336 & 110 \\ 70 & 231 & 90 \end{pmatrix} \quad , \quad x^{(0)}, x^{(1)}, x^{(2)} = \begin{matrix} 1 & 391 & 190756 \\ 1 & 556 & 272836 \\ 1 & 391 & 190756 \end{matrix}$$

Mit $d = (1,1,1)^T$ erhält man

$$\frac{(x^{(1)}, d)}{(x^{(0)}, d)} = 446 \quad , \quad \frac{(x^{(2)}, d)}{(x^{(1)}, d)} = 489.05 \ .$$

Für $j = 2$ lauten die normierten Vektoren $x^{(k)}/x_2^{(k)}$:

$$\begin{matrix} 1 & 0.703237 & 0.699160 \\ 1 & 1 & 1 \\ 1 & 0.703237 & 0.699160 \end{matrix}$$

Die exakten Werte sind

$$\lambda_1 = 490 \quad , \quad \lambda_2 = 20 \quad , \quad x_1 = \begin{pmatrix} 0.7 \\ 1 \\ 0.7 \end{pmatrix} \quad .$$

Aus dem kleinen Verhältnis $\frac{\lambda_2}{\lambda_1} = 0.04$ erklärt sich die schnelle Konvergenz.

Zur Berechnung der weiteren Eigenwerte bildet man die Matrix $T = (A - \mu I)^{-1}$. Diese hat die Eigenwerte $(\lambda_i - \mu)^{-1}$ mit den Eigenvektoren $x_i$. Zur Berechnung von $\lambda_2$ wählt man $\mu$ so, daß

$$|\lambda_2 - \mu| < |\lambda_i - \mu| \quad , \quad i \neq 2 \quad .$$

Dann ist $(\lambda_2 - \mu)^{-1}$ betragsgrößter Eigenwert von $T$. Diesen kann man nach der Potenzmethode berechnen. Zur Bildung von

$$\begin{aligned} x^{(k+1)} &= T x^{(k)} \ , \\ (A - \mu I) x^{(k+1)} &= x^{(k)} \end{aligned}$$

muß man bei jedem Schritt ein Gleichungssystem mit ein und derselben Matrix lösen. Man braucht also die LR-Zerlegung nur einmal durchzuführen.

Dieses Verfahren heißt "inverse Potenzmethode" oder Wielandt-Iteration.

Sei nun $A$ eine beliebige Matrix mit Jordan'scher Normalform $J$, also

$$A = XJX^{-1} \quad , \quad J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} \quad , \quad J_\ell = \begin{pmatrix} \lambda_\ell & 1 & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_\ell \end{pmatrix} .$$

Die Eigenwerte $\lambda_\ell$ sind also nach geometrischer Vielfachheit gezählt (d.h. hat $\lambda_\ell$ die geometrische Vielfachheit $\rho_\ell$, so tritt $\lambda_\ell$ $\rho_\ell$-mal auf) und nach abnehmenden Beträgen geordnet:

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_r|$$

Seien $x^{(k)} = Ax^{(k-1)}$ die Vektoren der Potenzmethode. Mit $x^{(k)} = Xy^{(k)}$ wird $y^{(k)} = Jy^{(k-1)}$. Spalten wir $y^{(k)}$ auf in Teilvektoren $y_\ell^{(k)}$ der Länge $\nu_\ell$, so entsteht

$$y_\ell^{(k)} = J_\ell y_\ell^{(k-1)} \quad , \quad y_\ell^{(k)} = J_\ell^k y_\ell^{(0)} \quad , \quad \ell = 1, \ldots, r \quad , \quad y^{(k)} = \begin{pmatrix} y_1^{(k)} \\ \vdots \\ y_r^{(k)} \end{pmatrix} .$$

Zur Berechnung von $J_\ell^k$ setzen wir

$$J_\ell = \lambda_\ell I + N_\ell \quad , \quad N_\ell = \begin{pmatrix} 0 & 1 & & O \\ & \ddots & \ddots & \\ & & & 1 \\ O & & & 0 \end{pmatrix} .$$

Wegen $N_\ell^{\nu_\ell} = 0$ wird dann für $k \geq \nu_\ell$

$$\begin{aligned} J_\ell^k &= (\lambda_\ell I + N_\ell)^k = \sum_{\nu=0}^{\nu_\ell - 1} \begin{pmatrix} k \\ \nu \end{pmatrix} \lambda_\ell^{k-\nu} N_\ell^\nu \\ &= \lambda^k \sum_{\nu=0}^{\nu_\ell - 1} \begin{pmatrix} k \\ \nu \end{pmatrix} \lambda_\ell^{-\nu} N_\ell^\nu \\ &= \lambda_\ell^k M_{\ell k} \end{aligned}$$

mit einem Polynom $M_{\ell k}$ vom Grade $< \nu_\ell$ in $k$. Damit haben wir

$$y_\ell^{(k)} = \lambda_\ell^k M_{\ell k} y_\ell^{(0)} \quad , \quad \ell = 1, \ldots, r .$$

Um nun wieder zu den $x^{(k)}$ zurückzukommen, zerlegen wir $X = (X_1, \dots, X_r)$ mit $(n, \nu_\ell)$-Matrizen $X_\ell$ und haben dann

$$x^{(k)} = Xy^{(k)} = \sum_{\ell=1}^{r} X_\ell y_\ell^{(k)} = \sum_{\ell=1}^{r} \lambda_\ell^k X_\ell M_{\ell k} y_\ell^{(0)} .$$

Diese Darstellung von $x^{(k)}$ ist der Ersatz für die oben benutzte Entwicklung nach Eigenvektoren, in welche sie für $\nu_\ell = 1$, $\ell = 1, \dots, r$ übergeht.

Wir untersuchen die Potenzmethode nun für verschiedene Fälle.

**Fall 1:** Es gibt einen Eigenwert $\lambda_1$ maximalen Betrages, und es ist $\rho(\lambda_1) = \sigma(\lambda_1) = \rho$.

Es ist also

$$\begin{aligned} \lambda_1 &= \lambda_2 = \cdots = \lambda_\rho \ , \quad |\lambda_\rho| > |\lambda_{\rho+1}| \geq \cdots \geq |\lambda_r| \ , \\ x^{(k)} &= \sum_{\ell=1}^{\rho} \lambda_\ell^k X_\ell M_{\ell k} y_\ell^{(0)} + \sum_{\ell=\rho+1}^{r} \lambda_\ell^k X_\ell M_{\ell k} y_\ell^{(0)} . \end{aligned}$$

Für $\ell = 1, \dots, \rho$ ist $\nu_\ell = 1$ und $M_{\ell k}$ die $(1,1)$-Matrix 1. Also wird

$$\begin{aligned} x^{(k)} &= \lambda_1^k \left\{ \sum_{\ell=1}^{\rho} X_\ell y_\ell^{(0)} + \sum_{\ell=\rho+1}^{r} \left( \frac{\lambda_\ell}{\lambda_1} \right)^k X_\ell M_{\ell k} y_\ell^{(0)} \right\} \\ &= \lambda_1^k \{ x_1 + r_k \} . \end{aligned}$$

Hier ist $x_1$ ein Eigenvektor zum Eigenwert $\lambda_1$, und $r_k$ hat die Größenordnung

$$\left( \frac{\lambda_{\rho+1}}{\lambda_1} \right)^k k^{\nu-1} \to 0 \quad , \quad k \to \infty ,$$

wo $\nu = \operatorname*{Max}_{\ell > \rho} \nu_\ell$. Damit hat man in diesem Fall die gleichen Verhältnisse wie im oben diskutierten Fall. Die Konvergenz ist im wesentlichen $(\lambda_{\rho+1}/\lambda_\rho)^k$; der Faktor $k^{\nu-1}$ wächst so langsam, daß er numerisch kaum bemerkt wird.

**Fall 2:** Es gibt einen Eigenwert $\lambda_1$ maximalen Betrages, aber es ist $\rho(\lambda_1) < \sigma(\lambda_1)$.

Wir betrachten den einfachsten Spezialfall $\rho(\lambda_1) = 1$, $\sigma(\lambda_1) = 2$. Es ist dann

$$\begin{aligned}
x^{(k)} &= \lambda_1^k X_1 M_{1k} y_1^{(0)} + \sum_{\ell=2}^{r} \lambda_\ell^k X_\ell M_{\ell k} y_\ell^{(0)} \\
&= \lambda_1^k \left\{ X_1 M_{1k} y_1^{(0)} + \sum_{\ell=2}^{r} \left(\frac{\lambda_\ell}{\lambda_1}\right)^k X_\ell M_{\ell k} y_\ell^{(0)} \right\} \\
&= \lambda_1^k \left\{ X_1 M_{1k} y_1^{(0)} + r_k \right\} .
\end{aligned}$$

Ähnlich wie oben ist $r_k$ von der Größenordnung

$$r_k = \left(\frac{\lambda_2}{\lambda_1}\right)^k k^{\nu-1} \to 0 \quad , \quad k \to \infty$$

mit $\nu = \underset{\ell > 1}{\text{Max}}\ \nu_\ell$. $M_{1k}$ ist ein Polynom vom Grade 1 in $k$, d.h.

$$X_1 M_{1k} y_1^{(0)} = a + kb$$

mit geeigneten Vektoren $a, b \in \mathbb{C}^2$. Bildet man nun $(x^k, d)$ und die Quotienten zur Berechnung von $\lambda_1$, so entsteht

$$\begin{aligned}
\frac{(x^{(k+1)}, d)}{(x^{(k)}, d)} &= \lambda_1 \frac{(a, d) + (k+1)(b, d) + (r_{k+1}, d)}{(a, d) + k(b, d) + (r_k, d)} \\
&= \lambda_1 \left(1 + \ \mathrm{O}\left(\frac{1}{k}\right)\right)
\end{aligned}$$

für $k \to \infty$, wenn nur $(a, d) \neq 0$. Man hat also auch in diesem Fall Konvergenz gegen $\lambda_1$, aber sehr langsam.

**Fall 3:** Es gibt verschiedene betragsmaximale Eigenwerte.

Wir behandeln wieder den einfachsten Spezialfall

$$|\lambda_1| = |\lambda_2| > |\lambda_3| \geq \cdots \geq |\lambda_r| \quad , \quad \lambda_1 \neq \lambda_2$$

mit $\sigma(\lambda_1) = \sigma(\lambda_2) = 1$. Es ist dann

$$x^{(k)} = \lambda_1^k X_1 y_1^{(0)} + \lambda_2^k X_2 y_2^{(0)} + \sum_{\ell=3}^{r} \lambda_\ell^k X_\ell M_{\ell k} y_\ell^{(0)}$$

$$= \lambda_1^k \left\{ X_1 y_1^{(0)} + \left(\frac{\lambda_2}{\lambda_1}\right)^k X_2 y_2^{(0)} + \sum_{\ell=3}^{r} \left(\frac{\lambda_\ell}{\lambda_1}\right)^k X_\ell M_{\ell k} y_\ell^{(0)} \right\}$$

$$= \lambda_1^k \left\{ X_1 y_1^{(0)} + \left(\frac{\lambda_2}{\lambda_1}\right)^k X_2 y_2^{(0)} + r_k \right\} .$$

Wie in den früheren Fällen geht $r_k \to 0$ mit $k \to \infty$, und zwar (beinahe) geometrisch. Setzen wir

$$\frac{\lambda_2}{\lambda_1} = e^{i\alpha} \quad , \quad 0 < \alpha < 2\pi \quad ,$$

so ist

$$\left(\frac{\lambda_2}{\lambda_1}\right)^k = e^{i\alpha k} = \cos \alpha k + i \sin \alpha k \ .$$

Der Vektor

$$X_1 y_1^{(0)} + \left(\frac{\lambda_2}{\lambda_1}\right)^k X_2 y_2^{(0)}$$

ist also (i. allg., d.h. für $y_2^{(0)} \neq 0$) nicht konvergent, vielmehr oszillierend. In diesem Fall haben wir also keine Konvergenz.

Zusammenfassend haben wir den

**Satz 6.2.1** *Die Potenzmethode konvergiert, wenn es genau einen betragsgrößten Eigenwert gibt. Stimmen für diesen die algebraische und geometrische Vielfachheit überein, so ist die Konvergenz geometrisch. Gibt es verschiedene betragsgleiche Eigenwerte, so ist die Potenzmethode nicht konvergent.*

## 6.3 Der LR- und der QR-Algorithmus

Wir wollen uns nun überlegen, wie wir alle Eigenwerte einer Matrix durch die Potenzmethode berechnen können. Im Prinzip kann das - wie oben besprochen - durch die inverse Potenzmethode geschehen. Wir werden aber eine sehr viel elegantere Methode finden.

Betrachten wir wieder den Fall $n$ betragsmäßig verschiedener Eigenwerte, also $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$, und es gebe $n$ l.u. Eigenvektoren $x_1, \ldots, x_n$. Wenden wir die Potenzmethode auf $n$ Startvektoren $x_1^{(0)}, \ldots, x_n^{(0)}$ gleichzeitig an, also

$$X_{k+1} = AX_k \quad , \quad X_k = \left(x_1^{(k)}, \ldots, x_n^{(k)}\right) \ ,$$

so passiert nicht viel Interessantes: Alle Spalten von $X_k$ werden von $\lambda_1^k x_1$ dominiert. Um dies zu vermeiden, gehen wir raffinierter vor. In der ersten Spalte machen wir die ganz normale Potenzmethode, normieren $x_1^{(0)}$ allerdings so, daß die erste Komponente 1 ist:

$$r_{11} x_1^{(1)} = A x_1^{(0)} \ .$$

In der zweiten Spalte wollen wir aber möglichst keine Anteile von $x_1$ haben. Daher subtrahieren wir ein geeignetes Vielfaches von $x_1^{(1)}$:

$$r_{22} x_2^{(1)} = A x_2^{(0)} - r_{12} x_1^{(1)} \ .$$

$r_{12}$ bestimmen wir so, daß die erste Komponente von $x_2^{(1)}$ verschwindet. Danach wird $r_{22}$ so bestimmt, daß die zweite Komponente von $x_2^{(1)}$ 1 wird.

Entsprechend geht man in der Spalte $j$ vor: Man möchte, daß $x_j^{(1)}$ möglichst keine Anteile von $x_1, \ldots, x_{j-1}$ hat und subtrahiert dazu Vielfache von $x_1^{(1)}, \ldots, x_{j-1}^{(1)}$ so, daß die ersten $j-1$ Komponenten von $x_j^{(1)}$ verschwinden. Anschließend wird $x_j^{(1)}$ so normiert, daß die $j$-te Komponente 1 ist:

$$r_{jj} x_j^{(1)} = A x_j^{(0)} - r_{1j} x_1^{(1)} - r_{2j} x_2^{(1)} - \ldots - r_{j-1,j} x_{j-1}^{(1)} \ , \quad j = 1, \ldots, n \ . \qquad (3.1)$$

Faßt man die $r_{ij}$ zu der rechten Dreiecksmatrix $R_0$ zusammen, so lautet dies

$$X_1 R_0 = A X_0 \ . \qquad (3.2)$$

$X_1$ ist eine linke Dreiecksmatrix mit Hauptdiagonale 1. Wir haben hier also die LR-Zerlegung von $AX_0$ vorliegen.

Die Potenzmethode läuft nun folgendermaßen: Sei $X_0 = I$.
Ist $X_0$ berechnet, so bilde man die LR-Zerlegung

$$X_{k+1} R_k = A X_k \qquad (3.3)$$

von $AX_k$, wo also $X_{k+1}$ der linke Faktor ist.

Aufgrund der Herleitung erwarten wir, daß mit $k \to \infty$

$$\begin{array}{rcl} x_1^{(k)} & \to & r_{11}x_1 \\ x_2^{(k)} & \to & r_{12}x_1 + r_{22}x_2 \\ & \vdots & \\ x_n^{(k)} & \to & r_{1n}x_1 + r_{2n}x_2 + \ldots + r_{nn}x_n \end{array}$$

mit geeigneten Zahlen $r_{ij}$. Anders ausgedrückt: Mit einer rechten Dreiecksmatrix $R$ gilt

$$X_k \to XR \quad , \quad X = (x_1, \ldots, x_n) \; . \tag{3.4}$$

Nach einer Idee von Rutishauser kann man die Rechnung sehr elegant durchführen: Man setze

$$L_k = X_k^{-1}X_{k+1} \quad , \quad A_k = X_k^{-1}AX_k \; .$$

Dann sind die $L_k$ linke Dreiecksmatrizen mit Diagonale 1, und die $A_k$ sind alle ähnlich zu $A$. Es gilt weiter

$$\begin{array}{rcccccl} A_k & = & X_k^{-1}AX_k & = & (L_kX_{k+1}^{-1})AX_k & = & L_kR_k \quad , \\ A_{k+1} & = & (X_{k+1}^{-1}A)X_{k+1} & = & (R_kX_k^{-1})X_{k+1} & = & R_kL_k \quad . \end{array}$$

Schließlich erwarten wir noch, daß genügt der vermuteten Konvergenz $X_k \to XR$ mit $k \to \infty$

$$A_k = X_k^{-1}AX_k \to R^{-1}X^{-1}AXR = R^{-1}JR \; ,$$

wobei $J$ die Diagonalmatrix mit den Eigenwerten $\lambda_\ell$ auf der Diagonalen ist. Damit sind wir beim LR-Verfahren angelangt:

1) $A_0 = A$.

2) Ist $A_k$ berechnet, so bilde man die LR-Zerlegung

$$A_k = L_kR_k$$

und setze

$$A_{k+1} = R_kL_k \; .$$

3) Für große $k$ gilt

$$A_k \sim \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & x & \\ & & \ddots & & \\ & O & & & \\ & & & & \lambda_n \end{pmatrix}$$

mit irgendwelchen Elementen oberhalb der Diagonalen.

Tatsächlich kann man unter gewissen Voraussetzungen nur zeigen, daß die Diagonale von $A_k$ gegen eine Permutation der Eigenwerte konvergiert. Wir werden das LR-Verfahren jedoch nicht weiter verfolgen. Es hat den Nachteil, daß die LR-Zerlegung ja nicht immer durchführbar ist und numerische Stabilität nur durch Pivotisierung, also Zeilenvertauschungen, erreicht wird.

Durch eine leichte Modifikation des LR-Algorithmus kommen wir zum QR-Algorithmus: Wir bestimmen $r_{1,j}, \ldots, r_{j-1,j}$ in (3.1) so, daß $x_j^{(1)}$ orthogonal zu $x_1^{(1)}, \ldots, x_{j-1}^{(1)}$ ist und danach $r_{j,j}$ so, daß $x_j^{(1)}$ die Länge 1 hat. Dann hat man wieder (3.2), aber $X_1$ ist jetzt eine unitäre Matrix. Die Potenzmethode lautet wieder wie in (3.3), nur daß jetzt statt der LR-Zerlegung eine QR-Zerlegung mit unitärem Faktor $X_{k+1}$ durchgeführt wird. Wieder erwarten wir (3.4). Mit den Matrizen

$$Q_k = X_k^{-1} X_{k+1} \quad , \quad A_k = X_k^{-1} A X_k$$

finden wir genau wie beim LR-Verfahren

$$A_k = Q_k R_k \quad , \quad A_{k+1} = R_k Q_k \ .$$

Damit haben wir den QR-Algorithmus gefunden:

1) $A_0 = A$

2) Ist $A_k$ berechnet, so bilde man die QR-Zerlegung

$$A_k = Q_k R_k$$

und setze

$$A_{k+1} = R_k Q_k \ .$$

3) Wir erwarten

$$A_k \to \begin{pmatrix} \lambda_1 & & x \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} .$$

**Lemma 6.3.1** *Sei $A_k \to A$ und $Q_k R_k = A_k$ die QR-Zerlegung von $A_k$. Dann gilt $Q_k \to Q$, $R_k \to R$, wo $QR = A$ die QR-Zerlegung von $A$ ist.*

**Beweis:** Die erste Spalte von $Q_k R_k = A_k$ lautet

$$r_{11}^{(k)} q_1^{(k)} = a_1^{(k)} .$$

Wegen $r_{11}^{(k)} > 0$, $\|q_1^{(k)}\|_2 = 1$ und $a_1^{(k)} \to a_1$ konvergiert $r_{11}^{(k)}$, etwa gegen $r_{11}$. Damit konvergiert auch $q_1^{(k)}$, etwa gegen $q_1$. Die zweite Spalte von $Q_k R_k = A_k$ lautet

$$r_{12}^{(k)} q_1^{(k)} + r_{22}^{(k)} q_2^{(k)} = a_2^{(k)} .$$

Weil $Q_k$ unitär ist, gilt $r_{12}^{(k)} = (a_2^{(k)}, q_1^{(k)}) \to r_{12}$. Also konvergiert auch $r_{22}^{(k)} q_2^{(k)}$ und wegen $\|q_2^{(k)}\|_2 = 1$ auch $r_{22}^{(k)} \to r_{22}$, mithin auch $q_2^{(k)} \to q_2$.

Es ist klar, daß man so fortfahren kann und

$$R_k \to R \quad , \quad Q_k \to Q$$

bekommt mit einer rechten Dreiecksmatrix $R$ und einer unitären Matrix $Q$. Dies muß die QR-Zerlegung von $A$ sein.

□

**Satz 6.3.1** *$A$ besitze $n$ betragsmäßig verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$. Sei $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$ und $A = XJ^{-1}X^{-1}$ die Jordan'sche Normalform von $A$. $X^{-1}$ besitze eine LR-Zerlegung. Dann gilt für die Matrix $A_k$ des QR-Algorithmus' für $k \to \infty$*

$$(A_k)_{i,j} \to \begin{cases} \lambda_i & , \quad j = i \\ 0 & \quad\;\; j < i \end{cases} .$$

**Beweis:** Sei $X^{-1} = L_- R_-$ die LR-Zerlegung von $X^{-1}$. Der Beweis beruht auf dem Vergleich zweier QR-Zerlegungen von $A^k$. Die erste dieser QR-Zerlegungen bekommen wir aus

$$A^k = XJ^kX^{-1} = XJ^kL_-R_- = XJ^kL_-J^{-k}J^kR_- \ .$$

Wegen unserer Numerierung der Eigenwerte ist $J^kL_-J^{-k}$ von der Form $I+F_k$ mit $F_k \to 0$. Also haben wir

$$A^k = X(I+F_k)J^kR_- \ .$$

Sei nun $X = QR$ die QR-Zerlegung von $X$. Dann gilt

$$\begin{aligned} A^k &= QR(I+F_k)J^kR_- \\ &= Q(I+G_k)RJ^kR_- \end{aligned}$$

mit $G_k = RF_kR^{-1} \to 0$. Sei $P_kS_k = I + G_k$ die QR-Zerlegung von $I + G_k$. Nach dem Lemma gilt $P_k \to I$, $S_k \to I$. Wir haben also

$$A^k = QP_kS_kRJ^kR_- \tag{3.5}$$

mit unitären Matrizen $P_k \to I$ und rechten Dreiecksmatrizen $S_k \to I$. Die zweite QR-Zerlegung von $A^k$ ist

$$A^k = Q_0 \cdots Q_{k-1}R_{k-1} \cdots R_0 \ . \tag{3.6}$$

Dies ist der Fall $\ell = k$ der Identität

$$Q_0 \cdots Q_{k-\ell-1}A^\ell_{k-\ell}R_{k-\ell-1} \cdots R_0 = Q_0 \cdots Q_{k-1}R_{k-1} \cdots R_0 \ ,$$

die man für $\ell = 0, \ldots, k$ folgendermaßen beweist. Für $\ell = 0$ ist die Identität trivial. Es genügt also zu zeigen, daß die linke Seite von $\ell$ unabhängig ist, also

$$Q_{k-\ell-1}A^\ell_{k-\ell}R_{k-\ell-1} = A^{\ell+1}_{k-\ell-1} \ .$$

Wegen $A_{k-\ell} = R_{k-\ell-1}Q_{k-\ell-1}$, $A_{k-\ell-1} = Q_{k-\ell-1}R_{k-\ell-1}$ bestätigt man dies unmittelbar.

Aus dem Vergleich der QR-Zerlegungen (3.5), (3.6) bekommen wir

$$Q_0 \cdots Q_{k-1} = QP_k \quad , \quad R_{k-1} \cdots R_0 = S_kRJ^kR_- \ .$$

Aus der ersten dieser Beziehungen folgt

$$Q_k = (Q_0 \cdots Q_{k-1})^{-1} Q_0 \cdots Q_k = (QP_k)^{-1} QP_{k+1} \to I \ , \qquad (3.7)$$

aus der zweiten

$$\begin{aligned} R_k = R_k R_{k-1} \cdots R_0 (R_{k-1} \cdots R_0)^{-1} &= S_{k+1} R J^{k+1} R_- R_-^{-1} J^{-k} R^{-1} S_k^{-1} \\ &= S_{k+1} R J R^{-1} S_k^{-1} \to R J R^{-1} \ . \end{aligned} \qquad (3.8)$$

Also gilt

$$A_k = Q_k R_k \to R J R^{-1} \ ,$$

und dies ist eine Matrix der behaupteten Art.

□

**Bemerkungen:** Sind die Voraussetzungen des Satzes nicht erfüllt, so treten folgende Änderungen ein.

**I.** Ist $\lambda_i$ mehrfacher Eigenwert mit $\sigma(\lambda_i) = \rho(\lambda_i)$, so gilt der Satz unverändert.

Dann ist nämlich die Matrix $J^k L_- J^{-k}$ von der Form $L + F_k$ mit einer linken Dreiecksmatrix $L$, und (3.5) gilt nach wie vor, wobei jetzt $P_k \to P$, $S_k \to S$ mit einer unitären Matrix $P$ und einer rechten Dreiecksmatrix $S$ ist, welche nicht mehr notwendig $I$ sind. (3.7) gilt dann nach wie vor, während in (3.8) $R$ durch $SR$ ersetzt werden muß.

**II.** Wir lassen jetzt die Vorraussetzungen, daß die algebraische mit der geometrischen Vielfachheit übereinstimmt, fallen. $A$ besitze also $r$ Eigenwerte $\lambda_1, \ldots, \lambda_r$ mit $|\lambda_1| > |\lambda_2| > \cdots |\lambda_r| > 0$, und zu jedem Eigenwert $\lambda_\ell$ gehören ein oder mehrere Jordankästchen, die wir zu $J_\ell$ zusammenfassen. $J_\ell$ hat dann die Gestalt

$$J_\ell = \begin{pmatrix} \lambda_\ell & \theta_1 & & \\ & \ddots & \ddots & \\ & & & \theta_\ell \\ & & & \lambda_\ell \end{pmatrix} \quad , \quad J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} \ ,$$

wo $\theta_\ell = 0$ oder $\theta_\ell = 1$. $J_\ell$ ist ein $\sigma(\lambda_\ell) \times \sigma(\lambda_\ell)$-Matrix. Spaltet man $L_-$ gemäß $J$ auf, $L_- = (L_{i,j})$, so wird

$$\begin{aligned} \left(J^k L_- J^{-k}\right)_{i,j} &= J_i^k L_{i,j} J_j^{-k} \\ &= \left(\frac{\lambda_i}{\lambda_j}\right)^k M_{i,k} L_{ij} M_{j,k}^{-1} \; , \quad i > j \quad , \end{aligned}$$

wobei $M_{i,k}$, $M_{i,k}^{-1}$ Polynome höchstens vom Grade $< \sigma(\lambda_i)$ in $k$ sind. Also gilt für $k \to \infty$ nach wie vor

$$\left(J^k L_- J^{-k}\right)_{i,j} \to 0 \quad , \quad i > j \; .$$

Die weitere (recht mühsame) Untersuchung zeigt nun, daß $A_k$ entsprechend $J$ aufgeteilt werden kann in Blöcke $A_{ijk}$, wobei für $k \to \infty$

$$A_{ijk} \to 0 \qquad \text{für} \quad i > j \; ,$$

Alle $\sigma(\lambda_i)$ Eigenwerte von $A_{iik}$ konvergieren gegen $\lambda_i$.

**III.** Den Fall betragsgleicher aber verschiedener Eigenwerte brauchen wir nicht zu betrachten, da wir ihn durch *shifts* vermeiden.