7.5.1 Eigenwerte und Diagonalisierung

Zwei $n \times n$ Matrizen A und B heißen *ähnlich*, falls es eine invertierbare Matrix P mit $A = P^{-1}BP$ gibt. Ist die Matrix P auch unitär, also gilt $P^{-1} = P^{\top}$, so heißen die Matrizen *unitär ähnlich*.

Ist P invertierbar, so ist wegen $(P^{-1})^{-1} = P$ auch P^{-1} invertierbar. Daher ist es unwichtig, ob man P^{-1} links oder rechts von B einsetzt. D. h. A und B sind auch dann ähnlich, falls $A = QBQ^{-1}$ für eine invertierbare Matrix Q gilt.

> Eine geometrische Interpretation dieser Äquivalenzrelation ist die folgende. Jede $n \times n$ Matrix A über $\mathbb{F}$ beschreibt eine lineare Abbildung $f_A(\boldsymbol{x}) = A\boldsymbol{x}$ von $\mathbb{F}^n$ nach $\mathbb{F}^n$. Eine invertierbare Matrix P implementiert einen Basiswechsel. Die Koordinaten sind durch P folgendermaßen »gedreht«, dass man diese Drehung auch umkehren kann, d. h. sie werden nicht in eine kleinere Dimension »gequetscht«. Die Abbildung $f_{P^{-1}BP}$ dreht den Vektorraum, wendet die Funktion f_B an und kehrt in den ursprünglichen Vektorraum zurück. D. h. ähnliche Matrizen beschreiben die gleichen linearen Abbildungen aus verschiedenen Sichtpunkten, die der gewählten Basis entsprechen.

Eine wichtige Eigenschaft ähnlicher Matrizen ist, dass sie denselben Rang, dieselbe Determinante wie auch dasselbe Spektrum (d. h. dieselben Eigenwerte) besitzen.

Satz 7.32: Ähnliche Matrizen

Sind die Matrizen A und B ähnlich, so gilt:

1. $\mathrm{rk}(A) = \mathrm{rk}(B)$;
2. $\det(A) = \det(B)$;
3. $\sigma(A) = \sigma(B)$.

Beweis:

(1) Es gelte $A = P^{-1}BP$ für eine invertierbare $n \times n$ Matrix P. Nach (7.1) gilt $\mathrm{rk}(A) = \mathrm{rk}(P^{-1}BP) \le \mathrm{rk}(B)$ und wegen der Symmetrie auch $\mathrm{rk}(B) \le \mathrm{rk}(A)$.
(2) folgt unmittelbar aus Satz 7.22: $\det(A) = \det(P^{-1})\det(B)\det(P) = \det(B)$.
(3) Wir wollen zeigen, dass die Matrizen A und B dieselben Eigenwerte besitzen. Dazu reicht es zu zeigen, dass die Determinanten der beiden Matrizen $A - \lambda E$ und $B - \lambda E$ für alle $\lambda \in \mathbb{R}$ gleich sind. Diese Gleichheit kann man mit Hilfe der uns bereits bekannten Eigenschaften der Determinate leicht zeigen:

$$
\begin{aligned}
\det(A - \lambda E) &= \det(P^{-1}BP - \lambda P^{-1}P) && \text{wegen } E = P^{-1}P\\
&= \det(P^{-1}(B - \lambda E)P) && \text{Satz 7.1(1)}\\
&= \det(P^{-1}) \cdot \det(B - \lambda E) \cdot \det(P) && \text{Satz 7.22(1)}\\
&= \det(B - \lambda E) && \text{wegen } \det(P^{-1}) = \tfrac{1}{\det(P)}\,.
\end{aligned}
$$

□

Wir wissen bereits, dass man den Rang, die Determinante wie auch die Eigenwerte einer Dreiecksmatrix sehr leicht bestimmen kann. Daher sind die zu Dreiecksmatrizen ähnlichen Matrizen besonders interessant; solche Matrizen nennt man *trigonalisierbar*. Ist die Matrix sogar einer Diagonalmatrix ähnlich, so heißt sie *diagonalisierbar*.

Satz 7.33: Diagonalisierung

Sind alle n Eigenwerte $\lambda_1, \ldots, \lambda_n$ einer reellwertigen $n \times n$ Matrix A verschiedene reelle Zahlen, so ist A der Diagonalmatrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ ähnlich.

Beweis:

Sei $P = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ eine $n \times n$ Matrix, deren Spalten Eigenvektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$ sind. Aus $A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$ für alle i folgt

$$\begin{aligned} A \cdot P = A \cdot [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] &= [A\boldsymbol{v}_1, \ldots, A\boldsymbol{v}_n] = [\lambda_1 \boldsymbol{v}_1, \ldots, \lambda_n \boldsymbol{v}_n] \\ &= P \cdot \operatorname{diag}(\lambda_1, \ldots, \lambda_n)\,. \end{aligned}$$

Da nach Satz 7.28 Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sein müssen, gilt $\operatorname{rk}(P) = n$. Somit existiert das Inverse P^{-1} und es genügt, beide Seiten der Gleichung $A \cdot P = P \cdot \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ mit P^{-1} zu multiplizieren. □

Beispiel 7.34:

Wir betrachten die Matrix $A = \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix}$ mit $a > 0$ aus Beispiel 7.30. Diese Matrix hat zwei Eigenwerte $\lambda = 1 \pm \sqrt{a}$ mit Eigenvektoren $\boldsymbol{v} = (\pm\sqrt{a}, 1)^\top$. Die entsprechende Matrix P hat in diesem Fall die Form

$$P = \begin{bmatrix} \sqrt{a} & -\sqrt{a} \\ 1 & 1 \end{bmatrix}$$

und ihr Inverses ist (siehe Beispiel 7.10)

$$P^{-1} = \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix}.$$

Wenn wir nun das Matrixprodukt $P^{-1}AP$ ausrechnen, dann erhalten wir eine Diagonalmatrix

$$\begin{aligned} P^{-1}AP &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix} \cdot \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{a} & -\sqrt{a} \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 1 & \sqrt{a} \\ -1 & \sqrt{a} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{a} + a & -\sqrt{a} + a \\ \sqrt{a} + 1 & -\sqrt{a} + 1 \end{bmatrix} \\ &= \frac{1}{2\sqrt{a}} \begin{bmatrix} 2\sqrt{a}(1 + \sqrt{a}) & 0 \\ 0 & 2\sqrt{a}(1 - \sqrt{a}) \end{bmatrix} \\ &= \begin{bmatrix} 1 + \sqrt{a} & 0 \\ 0 & 1 - \sqrt{a} \end{bmatrix}. \end{aligned}$$

Für Matrizen, deren Eigenwerte nicht unbedingt verschieden sind, gilt Folgendes.

Satz 7.35: Trigonalisierung – Satz von Schur
Jede reellwertige $n \times n$ Matrix A mit reellen Eigenwerten $\lambda_1, \ldots, \lambda_n$ ist einer oberen $n \times n$ Dreiecksmatrix $\varDelta$ mit $\lambda_1, \ldots, \lambda_n$ auf der Diagonalen unitär ähnlich.

Beweis:

Nach Satz 7.32(3) und Lemma 7.29 reicht es zu zeigen, dass A einer Dreiecksmatrix unitär ähnlich ist.

Induktion über n. Für $n = 1$ ist die Aussage richtig, da jede 1×1 Matrix auch eine Dreiecksmatrix ist. Wir nehmen nun an, dass die Aussage für alle $(n-1) \times (n-1)$ Matrizen gilt und dass A eine $n \times n$ Matrix ist, deren Eigenwerte reelle Zahlen sind.

Sei $\boldsymbol{v}_1$ ein Eigenvektor von A mit Eigenwert λ, also $A\boldsymbol{v}_1 = \lambda \boldsymbol{v}_1$ und $\boldsymbol{v}_1 \neq \mathbf{0}$. Ersetze $\boldsymbol{v}_1$ durch $\boldsymbol{v}_1 / \|\boldsymbol{v}_1\|$, um $\|\boldsymbol{v}_1\| = 1$ zu erreichen. Dann wende den Gram–Schmidt Algorithmus, um $\{\boldsymbol{v}_1\}$ bis zu einer Orthonormalbasis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ von $\mathbb{R}^n$ zu erweitern. Sei $B = V^{-1}AV$ mit $V = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$. Da die Basis orthonormal ist, ist die Matrix V unitär, also gilt $V^{-1} = V^\top$. Wegen $A\boldsymbol{v}_1 = \lambda \boldsymbol{v}_1$, $\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle = 1$ und $\langle \boldsymbol{v}_1, \boldsymbol{v}_i \rangle = 0$ für alle $i = 2, \ldots, n$ hat die Matrix $B = V^{-1}AV$ die Form

$$V^{-1}AV = V^\top [\lambda \boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n] = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{bmatrix}.$$

Nach Induktionsannahme gibt es eine unitäre $(n-1) \times (n-1)$ Matrix Q, so dass $Q^{-1}CQ$ eine obere Dreiecksmatrix ist. Da Q unitär ist, ist auch die Matrix

$$R = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{bmatrix}$$

unitär. Wegen

$$(VR)^{-1} = R^{-1}V^{-1} = R^\top V^\top = (VR)^\top$$

ist daher auch die Matrix $P = VR$ unitär. Die Matrix

$$P^{-1}AP = (VR)^{-1}A(VR) = R^{-1}(V^{-1}AV)R$$

hat die Form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{bmatrix} \cdot \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & Q^{-1}CQ & \\ 0 & & & \end{bmatrix}.$$

Da $Q^{-1}CQ$ nach Induktionsannahme eine obere Dreiecksmatrix ist, ist auch $\varDelta = P^{-1}AP$ eine solche Matrix. Somit sind die Matrizen A und $\varDelta$ unitär ähnlich. □

Korollar 7.36: Diagonalisierung symmetrischer Matrizen
Jede symmetrische reellwertige $n \times n$ Matrix A mit den Eigenwerten $\lambda_1, \ldots, \lambda_n$ ist der Diagonalmatrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ unitär ähnlich.

Beweis:
Nach Satz 7.31 sind alle λ_i reelle Zahlen. Nach dem Satz von Schur gilt $A = P\Delta P^{-1}$ für eine unitäre $n \times n$ Matrix P und eine Dreiecksmatrix Δ mit Einträgen $\lambda_1, \ldots, \lambda_n$ auf der Diagonalen. Wegen $A^\top = A$ und $P^\top = P^{-1}$ gilt

$$P\Delta P^{-1} = A = A^\top = (P\Delta P^{-1})^\top = (P^{-1})^\top \Delta^\top P^\top = P\Delta^\top P^{-1},$$

woraus $\Delta = \Delta^\top$ folgt. Da Δ eine Dreiecksmatrix ist, kann das nur dann der Fall sein, wenn alle Einträge außerhalb der Diagonalen gleich Null sind, d. h. wenn Δ eine Diagonalmatrix ist. □

7.5.2 Eigenwerte, die Spur und die Determinante

Eigenwerte von A sind Nullstellen des charakteristischen Polynoms $\det(A - \lambda E)$. Nach der Leibniz'scher Formel (7.5) für die Determinante hat dieses Polynom die Form

$$\det(A - \lambda E) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

für geeignete Koeffizienten c_i. Für $\lambda = 0$ ergibt dies $c_0 = \det(A - 0 \cdot E) = \det(A)$. Der großte Term λ^n kommt in der Summe (7.5) nur einmal vor, und zwar mit dem Koeffizient $c_n = (-1)^n$. Somit gilt

$$\det(A - \lambda E) = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + \det(A).$$

Sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A, so sind sie Nullstellen dieses Polynoms und es gilt

$$\det(A - \lambda E) = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n). \tag{7.11}$$

Daraus ergeben sich einige Verbindungen zwischen Eigenwerten, »Spur« und Determinante. Die *Spur* (engl. »trace«) einer $n \times n$ Matrix $A = (a_{ij})$ ist einfach die Summe $\operatorname{Tr}(A) = \sum_{i=1}^n a_{ii}$ der Diagonaleinträge.

Satz 7.37:
Ist A eine $n \times n$ Matrix mit Eigenwerten $\lambda_1, \ldots, \lambda_n$, so gilt

$$\det(A) = \prod_{i=1}^n \lambda_i \quad \text{und} \quad \operatorname{Tr}(A) = \sum_{i=1}^n \lambda_i .$$

Beweis:
Um die erste Gleichung zu erhalten, setze einfach $\lambda = 0$ in (7.11).

Um die zweite Gleichung zu erhalten, schreibe die rechte Seite von (7.11) als ein Polynom (in der Variable λ) und beachte, dass der Koeffizient zu λ^{n-1} gleich

$$(-1)^n \sum_{i=1}^{n}(-\lambda_i) = (-1)^{n+1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$$

ist, wobei der Koeffizient zu λ^{n-1} in $\det(A - \lambda E)$ gleich

$$(-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) = (-1)^{n-1}\mathrm{Tr}(A)$$

ist. Warum? Um λ^{n-1} zu erhalten, müssen wir nach der Leibniz'schen Formel (7.5) den Term $-\lambda$ genau $(n-1)$-mal aus den Diagonaleinträgen $a_{ii} - \lambda$ von $A - \lambda E$ auswählen. □

Lemma 7.38: Eigenschaften der Spur

Für alle $n \times n$ Matrizen A, B über $\mathbb{R}$ und alle $c \in \mathbb{R}$ gilt:

1. $\mathrm{Tr}(cA) = c\mathrm{Tr}(A)$;
2. $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$;
3. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$;
4. $\mathrm{Tr}(A) = \mathrm{Tr}(B)$, falls A und B ähnlich sind.

Beweis:

Die ersten zwei Aussagen sind trivial. Die dritte Aussage folgt aus

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} A[i,k]B[k,i] = \sum_{k=1}^{n}\sum_{i=1}^{n} B[k,i]A[i,k] = \mathrm{Tr}(BA)\,.$$

Die letzte Aussage folgt aus (3): Gilt $A = P^{-1}BP$, so gilt auch

$$\mathrm{Tr}(A) = \mathrm{Tr}((P^{-1}B)P) = \mathrm{Tr}(P(P^{-1}B)) = \mathrm{Tr}(EB) = \mathrm{Tr}(B)\,.$$ □

Lemma 7.39: Rang und Spur symmetrischer Matrizen

Für jede reellwertige symmetrische Matrix A gilt

$$\mathrm{rk}(A^2) = \mathrm{rk}(A) \geq \frac{\mathrm{Tr}(A)^2}{\mathrm{Tr}(A^2)}\,.$$

Beweis:

Seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A. Nach Korollar 7.36 ist A der Diagonalmatrix $\Delta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ähnlich. Daher ist auch A^2 der Diagonalmatrix $\Delta^2 = \mathrm{diag}(\lambda_1^2, \ldots, \lambda_n^2)$ ähnlich. Nach Satz 7.32 sind $\mathrm{rk}(A)$ und $\mathrm{rk}(A^2)$ beide gleich der Anzahl $r = |\{i\colon \lambda_i \neq 0\}|$ der von Null verschiedenen Diagonaleinträge dieser Matrizen, woraus $\mathrm{rk}(A^2) = \mathrm{rk}(A)$ folgt. Weiterhin gilt nach Satz 7.37 und Lemma 7.38(4) $\sum_{i=1}^{n} \lambda_i = \mathrm{Tr}(A)$ und $\sum_{i=1}^{n} \lambda_i^2 = \mathrm{Tr}(A^2)$. Die Cauchy–Schwarz-Ungleichung (6.2) angewandt mit $x_i = \lambda_i$ und $y_i = 1$ ergibt dann

$$\mathrm{Tr}(A^2) \cdot r = \Big(\sum_{i=1}^{n} \lambda_i^2\Big) \cdot r \geq \Big(\sum_{i=1}^{n} \lambda_i\Big)^2 = \mathrm{Tr}(A)^2\,,$$

woraus die gewünschte untere Schranke für $r = \mathrm{rk}(A)$ folgt. □

7.6 Aufgaben

Aufgabe 7.1:

Ein Vater und seine beide Söhne sind zusammen hundert Jahre alt, der Vater ist doppelt so alt wie sein jüngster Sohn und dreißig Jahre älter als sein ältester. Wie alt ist der Vater?

Aufgabe 7.2:

Für welche Werte von a hat das folgende Gleichungssystem (i) keine, (ii) genau eine oder (iii) mehrere Lösungen?

$$\begin{array}{rcrcrl} x & + & y & + & az & = 1\,, \\ x & + & ay & + & z & = 1\,, \\ ax & + & y & + & z & = 1\,. \end{array}$$

Aufgabe 7.3:

Löse die folgende Gleichung über $\mathbb{R}$: $\begin{bmatrix} -4 & x \\ -x & 4 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$.

Aufgabe 7.4:

Löse die folgende Gleichung über $\mathbb{R}$: $X^2 - 2X = \begin{bmatrix} -1 & 0 \\ 6 & 3 \end{bmatrix}$. *Hinweis*: Es lohnt sich $X^2 - 2X$ bis zu einem Quadrat zu erweitern, d. h. durch $(X - E)^2 = X^2 - 2X + E$ zu ersetzen.

Aufgabe 7.5: Eigenschaften der linearen Abbildungen

Sei A eine $m \times n$ Matrix und sei $S \subseteq \mathbb{F}^m$, $|S| = n$ die Menge ihrer Spalten. Weiterhin sei $f_A : \mathbb{F}^n \to \mathbb{F}^m$ die durch $f_A(\boldsymbol{x}) = A\boldsymbol{x}$ gegebe lineare Abbildung. Zeige, dass dann folgendes gilt:

1. f_A ist injektiv $\iff$ S ist linear unabhängig;
2. f_A ist surjektiv $\iff$ $\mathrm{span}(S) = \mathbb{F}^m$;
3. f_A ist bijektiv $\iff$ S bildet eine Basis von $\mathbb{F}^m$.

Aufgabe 7.6:

Berechne $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n$, wobei n eine natürliche Zahl ist.