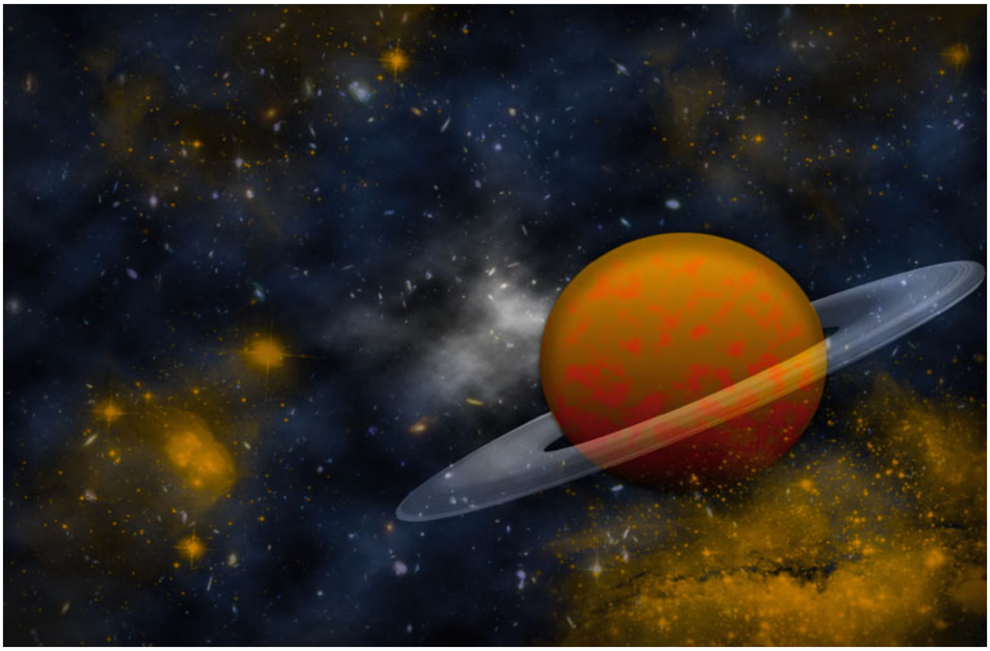
Eigenwerte und Normalformen

14



Wie lassen sich Matrizen vereinfachen?

Wie lassen sich Potenzen von Matrizen berechnen?

Welche Matrizen sind diagonalisierbar?

Nach welchem Prinzip funktionieren Suchmaschinen?

Wie kann man nicht invertierbare Matrizen „invertieren"?

Teil III

L. Göllmann et al., *Mathematik für Ingenieure: Verstehen – Rechnen – Anwenden*, DOI 10.1007/978-3-662-53867-8_14

Matrizen können sehr unterschiedliche Gestalt annehmen, und ihre Gestalt ist von entscheidender Bedeutung dafür, wie leicht oder schwer es ist, mit diesen Matrizen zu arbeiten und zu rechnen. So ist es in der Regel etwa umso einfacher mit einer Matrix zu arbeiten, je weniger von 0 verschiedene Einträge sie hat. Besonders leicht zu behandeln sind dabei Matrizen, die nur entlang der Diagonalen von 0 verschiedene Einträge haben. So ist etwa das Gleichungssystem

$$\begin{aligned} 2x_1 \qquad\qquad &= 1 \\ 3x_2 \qquad &= 2 \\ -x_3 &= 4 \end{aligned}$$

einfacher zu lösen als das Gleichungssystem

$$\begin{aligned} 2x_1 + 4x_2 + 3x_3 &= 1 \\ x_1 \;\; - 3x_2 - 2x_3 &= 2 \\ 2x_1 - 2x_3 + x_3 \;\; &= 4\,. \end{aligned}$$

Bei der Determinantenberechnung ist

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

einfacher zu handhaben als

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

und bei der Multiplikation mit einem Vektor

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

gilt

$$A \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2 \cdot v_1 \\ 3 \cdot v_2 \\ (-1) \cdot v_3 \end{pmatrix},$$

wohingegen $B \cdot \boldsymbol{v}$ nicht nach einer ähnlich einfachen Formel ermittelt werden kann.

Nun kommen natürlich in vielen Fragestellungen Matrizen mit vielen Einträgen vor. In diesem Fall können wir uns jedoch fragen, ob es möglich ist, diese Matrizen durch eine geschickte Transformation zu vereinfachen und auf Diagonalgestalt (oder eine Form, die zumindest nahe an der Diagonalgestalt ist) zu bringen. Dabei stellt sich als Erstes die Frage, was wir unter „geschickter Transformation" verstehen wollen. Der Schlüssel hierzu liegt in der Interpretation von Matrizen als lineare Abbildung. Wir haben bis jetzt die Matrix zu einer linearen Abbildung immer mithilfe der Standardbasis des $\mathbb{R}^n$ gebildet. In diesem Kapitel wollen wir versuchen, durch geeignete Wahl einer Basis eine einfachere Form der beschreibenden Matrix einer linearen Abbildung zu bekommen.

14.1 Eigenwerte und Eigenvektoren

Besonders einfache $n \times n$-Matrizen sind Diagonalmatrizen D. Diese zeichnen sich dadurch aus, dass für sie gilt:

$$D \cdot \mathbf{e}_i = d_i \cdot \mathbf{e}_i \quad \text{für alle } i = 1, \ldots, n\,,$$

wobei $\mathbf{e}_1, \ldots, \mathbf{e}_n$ die Standardbasis des $\mathbb{R}^n$ bezeichnet und d_i das i-te Diagonalelement von D. Bei einer stark besetzten Matrix ist das nicht mehr der Fall; wir können uns aber auch in dieser Situation fragen, ob es Vektoren gibt, die analoge Eigenschaften haben.

Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix}. \tag{14.1}$$

Hier gilt etwa

$$A \cdot \mathbf{e}_1 = \begin{pmatrix} -2 \\ -2 \end{pmatrix},$$

und damit ist $A \cdot \mathbf{e}_1$ kein Vielfaches von $\mathbf{e}_1$. Allerdings erhalten wir

$$A \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix} = 2 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

der Vektor $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ wird also auf ein Vielfaches von sich abgebildet. ◀

Motiviert durch dieses Beispiel betrachten wir eine $n \times n$-Matrix A.

Definition

Ein $\lambda \in \mathbb{R}$ heißt **Eigenwert** von A, wenn es einen Vektor $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v} \neq \mathbf{0}$ gibt mit

$$A \cdot \boldsymbol{v} = \lambda \cdot \boldsymbol{v}\,.$$

In diesem Fall heißt der Vektor $\boldsymbol{v}$ **Eigenvektor** von A zum Eigenwert λ.

Beispiel

Ist

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix},$$

so sind 2 und 3 Eigenwerte von D, $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist ein Eigenvektor von D zum Eigenwert 2, und $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ist ein Eigenvektor von D zum Eigenwert 3.

Ganz allgemein sind die Diagonalelemente einer Diagonalmatrix Eigenwerte dieser Matrix, und die Standardbasisvektoren sind Eigenvektoren dazu. ◀

Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} . \tag{14.2}$$

Dann ist 0 ein Eigenwert von A, und $\boldsymbol{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 0. Ferner ist 3 ein Eigenwert von A, und $\boldsymbol{w} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ist ein Eigenvektor von A zum Eigenwert 3. ◀

Achtung Der Nullvektor $\mathbf{0}$ ist als Eigenvektor nicht zugelassen (schon deshalb nicht, weil ja für jede reelle Zahl r gelten würde: $A \cdot \mathbf{0} = r \cdot \mathbf{0}$). Der Skalar 0 ist aber sehr wohl als Eigenwert möglich, wie wir am Beispiel der Matrix (14.2) gesehen haben. ◀

An der Matrix (14.2) lesen wir auch eine andere wichtige Tatsache ab: Der Wert 0 kommt genau dann als Eigenwert von A vor, wenn es einen Vektor $\boldsymbol{v} \neq \mathbf{0}$ gibt mit $A \cdot \boldsymbol{v} = 0 \cdot \boldsymbol{v} = \mathbf{0}$, also wenn das homogene Gleichungssystem

$$A \cdot \boldsymbol{x} = \mathbf{0}$$

eine nichttriviale Lösung hat, was gleichbedeutend damit ist, dass $\det A = 0$, wie wir in Kap. 13 gesehen haben. Ist $\boldsymbol{v}$ eine solche nichttriviale Lösung, so ist $\boldsymbol{v}$ ein Eigenvektor von A zum Eigenwert 0. Ist also 0 ein Eigenwert einer Matrix A, so können wir das relativ leicht feststellen.

Das charakteristische Polynom einer Matrix bestimmt die Eigenwerte

Die Überlegungen, die wir in Mathematischer Hintergrund 14.1 für das Beispiel der Matrix (14.1) durchführen, lassen sich auf eine beliebige $n \times n$-Matrix A übertragen:

Eigenwerte und lineare Gleichungssysteme

Eine Zahl λ ist genau dann ein Eigenwert von A, wenn das homogene Gleichungssystem

$$(\lambda \cdot \mathrm{E}_n - A) \cdot \boldsymbol{x} = \mathbf{0}$$

eine nichttriviale Lösung hat, also genau dann, wenn $\det(\lambda \cdot \mathrm{E}_n - A) = 0$.

Ist $\det(\lambda \cdot \mathrm{E}_n - A) = 0$, so sind die nichttrivialen Lösungen des homogenen Gleichungssystems

$$(\lambda \cdot \mathrm{E}_n - A) \cdot \boldsymbol{x} = \mathbf{0}$$

die Eigenvektoren von A zum Eigenwert λ.

Diese Aussage motiviert die allgemeine Betrachtung der parameterabhängigen Matrix $\lambda \cdot \mathrm{E}_n - A$ und ihrer Determinante

$$P_A(\lambda) = \det(\lambda \cdot \mathrm{E}_n - A)$$

(als Ausdruck im Parameter λ). Wir benutzen die vollständige Entwicklung von Determinanten, also die Formel

$$\det B = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} ,$$

wobei S_n die symmetrische Gruppe über $\{1, \ldots, n\}$ bezeichnet. Wenden wir das auf die Matrix $\lambda \cdot \mathrm{E}_n - A$ an, so sehen wir, dass

$$\det(\lambda \cdot \mathrm{E}_n - A) = (\lambda - a_{1,1}) \cdots (\lambda - a_{n,n}) + \text{ weitere Terme} ,$$

wobei in jedem der weiteren Terme λ höchstens $(n-1)$-mal auftaucht. Da

$$(\lambda - a_{1,1}) \cdots (\lambda - a_{n,n}) = \lambda^n + \text{ weitere Terme} ,$$

wobei auch hier in jedem weiteren Term λ höchstens $(n-1)$-mal auftaucht, schließen wir:

Das charakteristische Polynom

Ist A eine $n \times n$-Matrix, so ist $P_A(\lambda) = \det(\lambda \cdot \mathrm{E}_n - A)$ ein Polynom vom Grad n in λ von der Form

$$P_A(\lambda) = \lambda^n + r_{n-1} \cdot \lambda^{n-1} + \cdots + r_1 \cdot \lambda + r_0 .$$

Der Ausdruck $P_A(\lambda)$ heißt das **charakteristische Polynom** der Matrix A.

Alle Eigenwerte von A sind also Nullstellen des charakteristischen Polynoms $P_A(\lambda)$. Für große n können diese allerdings nicht mehr exakt bestimmt werden, sondern müssen mit Näherungsverfahren (z. B. dem Newton-Verfahren aus Abschn. 8.2) ermittelt werden. Da ein Polynom vom Grad n höchstens n verschiedene Nullstellen hat, erhalten wir sofort:

14.1 Mathematischer Hintergrund: Eigenwerte und Determinanten

Um allgemein Eigenwerte und Eigenvektoren einer Matrix A zu finden, ist eine Beziehung

$$A \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x}$$

mit einem Vektor $\boldsymbol{x}$ von Unbekannten und einer ebenfalls noch zu bestimmenden Zahl λ zu lösen. Hier erhalten wir ein nichtlineares Gleichungssystem mit n Gleichungen in $n+1$ Unbekannten. Das lässt sich so direkt nicht lösen. Wir wollen daher das Beispiel der Matrix (14.1) noch einmal aufgreifen. Hier wissen wir bereits, dass

$$\begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Wenn wir alles auf die rechte Seite bringen und zusammenfassen, erhalten wir

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= 2 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 2+2 & -6 \\ 2 & 2-5 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} \\ &= \left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} \right) \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \end{aligned}$$

der Vektor $\boldsymbol{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ ist also eine nichttriviale Lösung des homogenen Gleichungssystems

$$(2 \cdot \mathrm{E}_2 - A) \cdot \boldsymbol{x} = \boldsymbol{0}.$$

Umgekehrt gilt für jede nichttriviale Lösung $\boldsymbol{x}$ dieses Gleichungssystems

$$A \cdot \boldsymbol{x} = 2 \cdot \mathrm{E}_2 \cdot \boldsymbol{x} = 2 \cdot \boldsymbol{x},$$

also ist $\boldsymbol{x}$ ein Eigenvektor zum Eigenwert 2. Die Existenz einer solchen nichttrivialen Lösung ist gleichbedeutend mit $\det(2 \cdot \mathrm{E}_2 - A) = 0$.

Dieser Ansatz lässt sich auch verwenden, um zu entscheiden, ob eine Zahl kein Eigenwert von A ist. Betrachten wir dazu etwa $\lambda = 3$. Falls $\lambda = 3$ ein Eigenwert von A ist, so gibt es einen vom Nullvektor verschiedenen Vektor $\boldsymbol{v}$ mit

$$A \cdot \boldsymbol{v} = 3 \cdot \boldsymbol{v}.$$

Führen wir die gleichen Umformungen wie oben durch, so erhalten wir

$$(3 \cdot \mathrm{E}_2 - A) \cdot \boldsymbol{v} = \boldsymbol{0},$$

der Vektor $\boldsymbol{v}$ ist also eine nichttriviale Lösung des homogenen Gleichungssystems mit Koeffizientenmatrix $3 \cdot \mathrm{E}_2 - A$. Damit eine solche Lösung existieren kann, muss die Determinante dieser Koeffizientenmatrix verschwinden. Es gilt jedoch

$$\det(3 \cdot \mathrm{E}_2 - A) = \begin{vmatrix} 3-(-2) & -6 \\ 2 & 3-5 \end{vmatrix} = 2.$$

Daher kann es ein solches $\boldsymbol{v}$ nicht geben, und damit ist gezeigt, dass 3 kein Eigenwert von A ist.

Diese Rechnungen lassen sich für jede Zahl durchführen, und sie zeigen, dass ein λ genau dann ein Eigenwert von A ist, wenn das homogene Gleichungssystem

$$(\lambda \cdot \mathrm{E}_2 - A) \cdot \boldsymbol{x} = \boldsymbol{0}$$

eine nichttriviale Lösung hat, und das ist genau dann der Fall, wenn $\det(\lambda \cdot \mathrm{E}_2 - A) = 0$. Diese Determinante können wir ganz allgemein (in Abhängigkeit von dem Parameter λ) berechnen,

$$\begin{aligned} \det(\lambda \cdot \mathrm{E}_2 - A) &= \begin{vmatrix} \lambda-(-2) & -6 \\ 2 & \lambda-5 \end{vmatrix} \\ &= (\lambda+2) \cdot (\lambda-5) - (-6) \cdot 2 \\ &= \lambda^2 - 3\lambda + 2, \end{aligned}$$

und wir sehen, dass

$$\det(\lambda \cdot \mathrm{E}_2 - A) = 0 \iff \lambda = 1 \text{ oder } \lambda = 2.$$

Neben dem Eigenwert 2, den wir schon kennen, gibt es also noch den Eigenwert 1. Einen zugehörigen Eigenvektor finden wir, indem wir eine nichttriviale Lösung von

$$A \cdot \boldsymbol{x} = 1 \cdot \boldsymbol{x}$$

suchen, also eine nichttriviale Lösung des homogenen Gleichungssystems

$$(1 \cdot \mathrm{E}_2 - A) \cdot \boldsymbol{x} = \boldsymbol{0}.$$

Es ist

$$(1 \cdot \mathrm{E}_2 - A) = \begin{pmatrix} 3 & -6 \\ 2 & -4 \end{pmatrix}$$

Wir gehen vor wie in Abschn. 12.3 und bringen diese Matrix auf gaußsche Normalform

$$\begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix}$$

Also gibt es nichttriviale Lösungen und eine Basis des Lösungsraumes des homogenen Gleichungssystems ist

$$\boldsymbol{w} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Dieser Vektor $\boldsymbol{w}$ ist auch ein Eigenvektor zum Eigenwert 1 (genauso wie jedes von $\boldsymbol{0}$ verschiedene Vielfache von $\boldsymbol{w}$).

Teil III

Anzahl der Eigenwerte

Ist A eine $n \times n$-Matrix, so hat A höchstens n verschiedene Eigenwerte.

Beispiel

Wir wollen nochmals die Matrix (14.1), also

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix},$$

betrachten. Hierfür ist $\lambda \cdot \mathrm{E}_2 - A = \begin{pmatrix} \lambda - 1 & -1 \\ -2 & \lambda - 2 \end{pmatrix}$ und damit

$$P_A(\lambda) = (\lambda - 1)(\lambda - 2) - (-1) \cdot (-2) = \lambda^2 - 3\lambda .$$

Wir sehen also, dass $P_A(\lambda)$ die beiden Nullstellen $\lambda_1 = 0$ und $\lambda_2 = 3$ hat, und dabei handelt es sich um die beiden (einzigen) Eigenwerte von A. ◀

Beispiel

Für die Matrix

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 3 & 2 & -3 \\ 2 & 2 & -3 \end{pmatrix} \tag{14.3}$$

gilt

$$\begin{aligned} P_A(\lambda) &= \begin{vmatrix} \lambda & -1 & 1 \\ -3 & \lambda - 2 & 3 \\ -2 & -2 & \lambda + 3 \end{vmatrix} \\ &= \lambda \cdot \begin{vmatrix} \lambda - 2 & 3 \\ -2 & \lambda + 3 \end{vmatrix} - (-3) \cdot \begin{vmatrix} -1 & 1 \\ -2 & \lambda + 3 \end{vmatrix} \\ &\quad + (-2) \cdot \begin{vmatrix} -1 & 1 \\ \lambda - 2 & 3 \end{vmatrix} \\ &= \lambda^3 + \lambda^2 - \lambda - 1 \\ &= (\lambda + 1)^2 (\lambda - 1) . \end{aligned}$$

Damit hat A die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$, also nur zwei Eigenwerte, einen weniger, als wir maximal bekommen könnten.

Die Eigenvektoren zum Eigenwert -1 sind die nichttrivialen Lösungen des Gleichungssystems

$$\begin{aligned} -x_1 - x_2 + x_3 &= 0 \\ -3x_1 - 3x_2 + 3x_3 &= 0 \\ -2x_1 - 2x_2 + 2x_3 &= 0 \end{aligned}$$

Eine Basis des Lösungsraumes bilden

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$$

und damit sind die Vektoren

$$\boldsymbol{v} = r \cdot \boldsymbol{v}_1 + s \cdot \boldsymbol{v}_2$$

mit $r \neq 0$ oder $s \neq 0$ die allgemeinen Eigenvektoren zum Eigenwert -1.

Die Eigenvektoren zum Eigenwert 1 sind die nichttrivialen Lösungen des Gleichungssystems

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -3x_1 - x_2 + 3x_3 &= 0 \\ -2x_1 - 2x_2 + 4x_3 &= 0 \end{aligned}$$

Eine Basis des zughörigen Lösungsraumes ist der Vektor

$$\boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix},$$

und damit sind die Vektoren $\boldsymbol{v} = r \cdot \boldsymbol{v}_3$ mit $r \neq 0$ die Eigenvektoren zum Eigenwert 1. ◀

Mit A ist auch die transponierte Matrix A^T eine $n \times n$-Matrix. Die Eigenwerte von A^T sind dann die Nullstellen von $P_{A^T}(\lambda)$. Da die transponierte Matrix einer Diagonalmatrix wieder diese Diagonalmatrix und $\lambda \cdot \mathrm{E}_n$ eine Diagonalmatrix ist, erhalten wir

$$\begin{aligned} P_{A^T}(\lambda) &= \det\left(\lambda \cdot \mathrm{E}_n - A^T\right) = \det\left((\lambda \cdot \mathrm{E}_n)^T - A^T\right) \\ &= \det\left((\lambda \cdot \mathrm{E}_n - A)^T\right) = \det\left(\lambda \cdot \mathrm{E}_n - A\right) \\ &= P_A(\lambda) , \end{aligned}$$

wobei wir ausgenutzt haben, dass eine Matrix und ihre Transponierte die gleiche Determinante haben. Die Matrizen A und A^T haben also das gleiche charakteristische Polynom.

Eigenwerte der transponierten Matrix

Die Matrizen A und A^T haben die gleichen Eigenwerte.

Auch komplexe Matrizen haben Eigenwerte und Eigenvektoren

Bei der Untersuchung von Eigenwerten spielen komplexe Zahlen eine besondere Rolle, sogar dann, wenn wir reelle Matrizen betrachten.

Anwendung: Eigenwerte und stabile Systeme

Eigenwerte und Eigenvektoren spielen eine wichtige Rolle in der System- und Regelungstheorie, speziell bei Systemen, in denen die Komponenten ihre Zustände nach festen Regeln verändern. Als Beispiel hierzu wollen wir eine Fahrradvermietung betrachten, die in einer Stadt drei Stationen besitzt. Ein Fahrrad, das an einer Station abgeholt wird, kann an einer beliebigen Station abgegeben werden, sodass die Fahrräder zwischen den Stationen zirkulieren. Dabei lässt sich ein sehr stabiles Muster des Abgabeverhaltens beobachten. Der Prozentsatz der Fahrräder, die an Station i ausgeliehen und am Abend an Station j zurückgegeben werden, ist konstant und wird mit $P_{j,i}$ bezeichnet. Dann ist $p_{j,i} = \frac{P_{j,i}}{100}$ die Wahrscheinlichkeit dafür, dass ein an Station i ausgeliehenes Fahrrad am Abend an Station j zurückgegeben wird, und wir erhalten eine Matrix

$$P = \left(p_{j,i}\right)_{i,j\in\{1,2,3\}}$$

von Übergangswahrscheinlichkeiten. Charakteristisch für solche Systeme sind die Eigenschaften

1. $p_{i,j} \geq 0$ für alle i, j,
2. $\sum\limits_{i=1}^{n} p_{i,j} = 1$.

Die zweite Bedingung besagt in unserem Fall, dass jedes ausgeliehene Fahrrad auch wieder abgegeben wird (kein Schwund durch Verlust, Diebstahl etc.). Ist v_i die Anzahl der Fahrräder, die zu Beginn eines Tages an Station i vorhanden sind, und setzen wir

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = P \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

so ist w_i die Anzahl der Fahrräder, die am Abend an Station i zurückgegeben werden (unter der Annahme, dass alle Fahrräder verliehen werden können). Für den Fahrradverleih ist es nun interessant, einen stabilen Zustand zu erzeugen, d. h., die Fahrräder so aufzuteilen, dass am Abend an jeder Station immer genauso viele sind wie am Morgen. Im Licht obiger Rechnung suchen wir v_1, v_2 und v_3, sodass

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = P \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

also einen Eigenvektor von P zum Eigenwert 1. Zunächst ist zu klären, ob 1 Eigenwert von P ist. Dazu beachten wir, dass die Summe der Einträge einer Spalte den Wert 1 ergibt, also

$$P^T \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{3} p_{i,1} \\ \sum\limits_{i=1}^{3} p_{i,2} \\ \sum\limits_{i=1}^{3} p_{i,3} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

nach den Regeln zum Rechnen mit transponierten Matrizen. Daher ist 1 ein Eigenwert von P^T und damit auch von P. Das System hat also nichttriviale stabile Zustände.

Wir wollen nun eine konkrete Situation betrachten, bei der die Übergangsmatrix P gegeben ist durch

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.1 & 0.2 \end{pmatrix}.$$

Das charakteristische Polynom muss nicht mehr berechnet werden, da wir ja schon wissen, dass der uns interessierende Wert, nämlich 1, ein Eigenwert von P ist. Die zugehörigen stabilen Zustände ergeben sich als Lösung der homogenen Gleichung mit Koeffizientenmatrix

$$\mathrm{E}_3 - P = \begin{pmatrix} 0.6 & -0.2 & -0.4 \\ -0.3 & 0.3 & -0.4 \\ -0.3 & -0.1 & 0.8 \end{pmatrix}.$$

Eine Normalform dieser Matrix ist

$$\begin{pmatrix} 1 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Der Lösungsraum hat also die Dimension 1, und eine Basis des Lösungsraumes ist der Vektor

$$\boldsymbol{v} = \begin{pmatrix} 5 \\ 9 \\ 3 \end{pmatrix}.$$

Die Fahrräder sind deshalb im Verhältnis 5 : 9 : 3 auf die Stationen aufzuteilen. Unter der Voraussetzung, dass sie immer alle ausgeliehen werden (oder dass sie auch in diesem Verhältnis ausgeliehen werden), sind dann stets an jeder Station am Abend genauso viele Fahrräder wie am Morgen.

Anwendung: Input-Output-Analyse (geschlossenes Modell)

In Abschn. 13.3 haben wir eine Volkswirtschaft untersucht, deren Produktion darauf ausgelegt war, eine gegebene externe Nachfrage zu erfüllen. Nun betrachten wir ein ökonomisches System, das vollständig autark ist, in dem also der gesamte Bedarf durch die eigene Produktion abgedeckt wird und in dem weder Güter noch Dienstleistungen ein- oder ausgeführt werden. Als Beispiel betrachten wir ein Modell mit den vier Sektoren R (Roh- und Grundstoffe, einschließlich Energie), I (industrielle Fertigung), S (Service und Dienstleistungen) und H (Haushalte). Mit den entsprechenden Kleinbuchstaben r, i, s und h bezeichnen wir den monetären Wert der Produktion des jeweiligen Sektors (wobei die Produktion der Haushalte aus den Arbeitsstunden ihrer Mitglieder besteht). Für die Produktion gelten die folgenden Regeln:

1. Der Sektor R benötigt für die Produktion einer Einheit 0.2 Einheiten seiner eigenen Produktion r, 0.3 Einheiten von i, 0.2 Einheiten von s und 0.3 Einheiten von h.
2. Der Sektor I benötigt für die Produktion einer Einheit 0.3 Einheiten von r, 0.1 Einheiten seiner eigenen Produktion i, 0.2 Einheiten von s und 0.4 Einheiten von h.
3. Der Sektor S benötigt für die Produktion einer Einheit 0.2 Einheiten von r, 0.1 Einheiten von i, 0.1 Einheiten seiner eigenen Produktion s und 0.6 Einheiten von h.
4. Der Sektor H benötigt für die Produktion einer Einheit 0.4 Einheiten von r, 0.3 Einheiten von i und 0.3 Einheiten von s.

Die Gleichgewichtsbedingung lautet nun, dass die Produktion eines jeden Sektors den Bedarf des Systems exakt abdecken muss. Die Produktion r des Sektors R muss also den Eigenbedarf an Rohstoffen in Höhe von $0.2r$ abdecken, den Rohstoffbedarf des Sektors I in Höhe von $0.3i$, den des Sektors S in Höhe von $0.2s$ und den des Sektors H in Höhe von $0.4h$ (und entsprechend für die anderen Sektoren). Daraus ergibt sich das Gleichungssystem

$$\begin{aligned} 0.2r + 0.3i + 0.2s + 0.4h &= r \\ 0.3r + 0.1i + 0.1s + 0.3h &= i \\ 0.2r + 0.2i + 0.1s + 0.3h &= s \\ 0.3r + 0.4i + 0.6s \phantom{{}+0.3h} &= h\,. \end{aligned} \tag{14.4}$$

Bezeichnen wir mit

$$A = \begin{pmatrix} 0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & 0.1 & 0.1 & 0.3 \\ 0.2 & 0.2 & 0.1 & 0.3 \\ 0.3 & 0.4 & 0.6 & 0.0 \end{pmatrix}$$

die Input-Output-Matrix zu dem System, so kann das Gleichungssystem (14.1) umformuliert werden zu

$$A \cdot \boldsymbol{x} = \boldsymbol{x}\,,$$

d. h., wir haben die Aufgabe, einen Eigenvektor von A zum Eigenwert 1 zu finden. Da die Summe der Einträge in jeder Spalte von A gleich 1 ist, gibt es eine solchen Eigenvektor, wie wir schon im Anwendungsbeispiel „Eigenwerte und stabile Systeme" festgestellt haben. Eine Basis des Lösungsraumes des zugehörigen homogenen Gleichungssystems

$$(\mathrm{E}_4 - A) \cdot \boldsymbol{x} = \boldsymbol{0}$$

ist gegeben durch

$$\boldsymbol{v} = \begin{pmatrix} 472 \\ 362 \\ 351 \\ 497 \end{pmatrix},$$

und jedes Vielfache dieses Vektors beschreibt eine Produktionsaufteilung auf die vier Sektoren, bei der das System im Gleichgewicht ist.

Dieser Ansatz überträgt sich wieder auf Systeme mit beliebig vielen Sektoren. Das System wird beschrieben durch eine Input-Output-Matrix

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$$

wobei $a_{i,j}$ den Anteil an einer Einheit der Produktion des Sektors i bezeichnet, der für die Produktion einer Einheit des Sektors j benötigt wird. Bedingung an die Input-Output-Matrix ist, dass $a_{i,j} \geq 0$ für alle i, j und dass die Spaltensumme in jeder Spalte j den Wert 1 ergibt (d. h., um eine Einheit von j zu produzieren, muss insgesamt auch eine Einheit in die Produktion einfließen, damit ein stabiles geschlossenes System entstehen kann). Die Gleichgewichtsbedingung für dieses System ist dann gegeben durch die Gleichung

$$A \cdot \boldsymbol{x} = \boldsymbol{x}\,.$$

Ein Gleichgewichtszustand ist also ein Eigenvektor von A zum Eigenwert 1.

Wie die Untersuchung offener Systeme geht auch die Analyse geschlossener Systeme im Wesentlichen zurück auf Wassily Leontief.

Teil III

Teil III

Beispiel

Die (reelle) Matrix

$$A = \begin{pmatrix} 1 & -4 \\ 5 & -3 \end{pmatrix}$$

hat das charakteristische Polynom

$$P_A(\lambda) = \begin{vmatrix} \lambda - 1 & 4 \\ -5 & \lambda + 3 \end{vmatrix} = \lambda^2 + 2\lambda + 17$$

und die Eigenwerte $\lambda_1 = -1 + 4\mathrm{i}$ und $\lambda_2 = -1 - 4\mathrm{i}$ (also keine reellen Eigenwerte). Zur Bestimmung eines Eigenvektors zum Eigenwert $\lambda_1 = -1 + 4\mathrm{i}$ haben wir das homogene Gleichungssystem

$$\begin{aligned} (-2 - 4\mathrm{i}) \cdot x_1 + \quad 4x_2 &= 0 \\ -5x_1 + (2 + 4\mathrm{i}) \cdot x_2 &= 0 \end{aligned}$$

zu betrachten. Eine Basis des Lösungsraumes bildet

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 + 4\mathrm{i} \\ 5 \end{pmatrix},$$

und damit sind die Vektoren

$$\boldsymbol{v} = z \cdot \begin{pmatrix} 2 + 4\mathrm{i} \\ 5 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$ die (komplexen) Eigenvektoren zum Eigenwert $-1 + 4\mathrm{i}$. Entsprechend erhalten wir für den Eigenwert $\lambda_2 = -1 - 4\mathrm{i}$ die Eigenvektoren

$$\boldsymbol{v} = z \cdot \begin{pmatrix} 2 - 4\mathrm{i} \\ 5 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$. ◀

Die Eigenwertuntersuchung kann aber auch direkt auf komplexe Matrizen angewendet werden.

Beispiel

Die Matrix

$$A = \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$$

hat das charakteristische Polynom

$$P_A(\lambda) = \begin{vmatrix} \lambda - 1 & -\mathrm{i} \\ -\mathrm{i} & \lambda - 1 \end{vmatrix} = \lambda^2 - 2\lambda + 2,$$

und damit hat A die Eigenwerte $\lambda_1 = 1 + \mathrm{i}$ und $\lambda_2 = 1 - \mathrm{i}$. Die Eigenvektoren zu $\lambda_1 = 1 + \mathrm{i}$ sind wie im reellen Fall die nichttrivialen (komplexen) Lösungen des Gleichungssystems

$$(\lambda_1 \cdot \mathrm{E}_2 - A)\,\boldsymbol{x} = \mathbf{0},$$

also die Lösungen von

$$\begin{aligned} \mathrm{i} \cdot x_1 - \mathrm{i} \cdot x_2 &= 0 \\ -\mathrm{i} \cdot x_1 + \mathrm{i} \cdot x_2 &= 0. \end{aligned}$$

Eine Basis des Lösungsraumes dieses Gleichungssystems ist

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

und damit sind die Vektoren

$$\boldsymbol{v} = z \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$ die Eigenvektoren zum Eigenwert $1 + \mathrm{i}$. Entsprechend erhalten wir für $\lambda_2 = 1 - \mathrm{i}$ die Eigenvektoren

$$\boldsymbol{v} = z \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

mit $z \in \mathbb{C} \setminus \{0\}$. ◀

14.2 Diagonalisierbarkeit

Wie wir gesehen haben, sind Eigenvektoren zu einer Matrix A Vektoren, bzgl. derer sich die Multiplikation von A mit dem Vektor besonders einfach darstellen lässt, nämlich als

$$A \cdot \boldsymbol{v} = \lambda \cdot \boldsymbol{v}.$$

Wie kann uns das aber jetzt bei der Vereinfachung der Matrizen helfen?

Transformationen mit Eigenvektoren vereinfachen die Matrix

Wir betrachten wieder die Matrix (14.1), also

$$A = \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix}.$$

Wie wir schon gesehen haben, gilt für sie und die Vektoren

$$\boldsymbol{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \qquad \boldsymbol{w} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

die Beziehung

$$A \cdot \boldsymbol{v} = 2 \cdot \boldsymbol{v}, \qquad A \cdot \boldsymbol{w} = 1 \cdot \boldsymbol{w}. \tag{14.5}$$

d. h., die Vektoren $\boldsymbol{v}$ und $\boldsymbol{w}$ spielen für A eine ähnliche Rolle wie die Standardbasisvektoren $\mathbf{e}_1$ und $\mathbf{e}_2$ für eine 2×2-Diagonalmatrix.

Schreiben wir die Vektoren $\boldsymbol{v}$ und $\boldsymbol{w}$ als Spalten einer 2×2-Matrix S,

$$S = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix},$$

so schreibt sich Beziehung (14.5) als

$$\begin{aligned} A \cdot S &= \begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 2 \\ 4 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \\ &= S \cdot \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{14.6}$$

Die Vektoren $\boldsymbol{v}$ und $\boldsymbol{w}$ sind linear unabhängig, und damit ist S eine invertierbare Matrix, und

$$S^{-1} = \begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix}.$$

Multplizieren wir (14.6) von links mit S^{-1}, so erhalten wir

$$S^{-1} \cdot A \cdot S = S^{-1} \cdot S \cdot \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \tag{14.7}$$

Durch Transformation mit der Matrix S haben wir also aus A eine Diagonalmatrix gemacht.

In der Terminologie der linearen Abbildungen und ihrer Matrixdarstellungen bedeutet dies, dass die durch die Matrix A definierte lineare Abbildung bzgl. der Basis $\boldsymbol{v}, \boldsymbol{w}$ durch eine Diagonalmatrix dargestellt wird.

Ähnlich können wir mit der Matrix (14.2), also

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix},$$

verfahren. Wie wir gesehen haben, ist 0 ein Eigenwert von A und $\boldsymbol{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ein Eigenvektor von A zum Eigenwert 0. Ferner ist 3 ein Eigenwert von A und $\boldsymbol{w} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ein zugehöriger Eigenvektor von A. Die Eigenvektoren $\boldsymbol{v}$ und $\boldsymbol{w}$ sind auch hier linear unabhängig (Aufgabe 14.10). Daher ist die Matrix

$$S = \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}$$

invertierbar mit Inverser

$$S^{-1} = \frac{1}{3} \cdot \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Transformieren wir A mit S, so erhalten wir hier

$$S^{-1} \cdot A \cdot S = \frac{1}{3} \cdot \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix},$$

also ist auch in diesem Fall $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix.

Matrizen mit hinreichend vielen Eigenvektoren sind diagonalisierbar

Wir haben in den Beispielen oben, ausgehend von den Eigenwerten von A und von zugehörigen Eigenvektoren, eine Transformation gefunden, die aus A eine Diagonalmatrix macht. Wir wollen uns nun damit beschäftigen, wie weit sich diese Ansätze verallgemeinern lassen.

Dazu nehmen wir zunächst an, dass A eine $n \times n$-Matrix ist, die die maximal mögliche Anzahl von n verschiedenen (reellen) Eigenwerten hat, für die das charakteristische Polynom also in n paarweise verschiedene Linearfaktoren zerfällt. Diese Eigenwerte bezeichnen wir mit $\lambda_1, \ldots, \lambda_n$, und zu jedem λ_i wählen wir einen Eigenvektor $\boldsymbol{v}_i$. Dann können wir die Beziehungen

$$A \cdot \boldsymbol{v}_k = \lambda_k \cdot \boldsymbol{v}_k \qquad \text{für } k = 1, \ldots, n$$

als Matrizengleichung schreiben. Dazu betrachten wir die Matrix $S = (\boldsymbol{v}_1\ \boldsymbol{v}_2\ \ldots\ \boldsymbol{v}_n)$, die die Vektoren $\boldsymbol{v}_k$ als k-te Spaltenvektoren hat. Genauso wie in den beiden Beispielen erhalten wir, dass

$$A \cdot S = S \cdot D,$$

wobei

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist. Nun sind nach Aufgabe 14.10 die Spalten von S linear unabhängig, und damit ist S invertierbar. Daher erhalten wir aus dieser Gleichung durch Multiplikation mit S^{-1} wieder eine Beziehung

$$S^{-1} \cdot A \cdot S = D,$$

also eine Transformation von A auf Diagonalgestalt. Wir haben damit die folgende Aussage gezeigt:

Teil III

Matrizen mit n Eigenwerten

Ist A eine $n \times n$-Matrix mit n verschiedenen (reellen) Eigenwerten $\lambda_1, \ldots, \lambda_n$, so gibt es eine invertierbare Matrix S mit

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Definition

Eine quadratische Matrix A heißt **diagonalisierbar**, wenn es eine invertierbare Matrix S gibt, sodass $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix ist.

Teil III

Eine $n \times n$-Matrix mit n verschiedenen Eigenwerten ist also diagonalisierbar. Wie sieht es nun aus mit Matrizen, die weniger als n Eigenwerte haben? Betrachten wir dazu die Matrix (14.3), also

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

Wie wir schon gesehen haben, hat diese Matrix nur zwei Eigenwerte, -1 und 1. Dabei haben wir zum Eigenwert -1 die zwei linear unabhängigen Eigenvektoren

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

und zum Eigenwert 1 den Eigenvektor

$$\boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

gefunden. Wir überzeugen uns, dass die Vektoren $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ linear unabhängig (also eine Basis von $\mathbb{R}^3$) sind, und bilden mit diesen die Matrix

$$S = (\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \boldsymbol{v}_3) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix}.$$

Hierfür rechnen wir unmittelbar nach, dass

$$A \cdot S = S \cdot \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und, da S invertierbar ist, also auch hier

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Für die Diagonalisierbarkeit von A ist es also nicht notwendig, dass A genau n verschiedene Eigenwerte hat. Es reicht, dass A genügend viele linear unabhängige Eigenvektoren hat.

Diagonalisierbarkeit von Matrizen

Genau dann ist eine $n \times n$-Matrix A diagonalisierbar, wenn es n linear unabhängige Eigenvektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ von A gibt. In diesem Fall ist die Matrix S, die $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ als Spalten hat, eine Transformationsmatrix für A.

Sind nämlich $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ linear unabhängige Eigenvektoren von A, so setzen wir $S = (\boldsymbol{v}_1 \ \ldots \ \boldsymbol{v}_n)$, bilden also die Matrix mit diesen Eigenvektoren als Spalten und rechnen dann wie in obigen Vorüberlegungen nach, dass $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix ist.

Gibt es umgekehrt eine invertierbare Matrix S, sodass

$$S^{-1} \cdot A \cdot S = D$$

eine Diagonalmatrix ist, so erhalten wir

$$A \cdot S = S \cdot D.$$

Bezeichnen wir die Spalten von S mit $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, so sind diese linear unabhängig. Betrachten wir dann die obige Matrizengleichung spaltenweise, so folgt

$$A \cdot \boldsymbol{v}_i = d_i \cdot \boldsymbol{v}_i \quad \text{für } i = 1, \ldots, n,$$

wobei d_i das i-te Diagonalelement von D ist, und die Aussage ist gezeigt.

Die Argumente zeigen sogar eine genauere Aussage: Im Komplexen zerfällt das charakteristische Polynom der $n \times n$-Matrix A immer in Linearfaktoren. Wir können also

$$P_A(\lambda) = (\lambda - \lambda_1)^{a_1} \cdots (\lambda - \lambda_k)^{a_k}$$

mit $a_i \geq 1$ schreiben. Die Zahl a_i heißt **algebraische Vielfachheit** des Eigenwertes λ_i. Bezeichnet g_i die Anzahl der linear unabhängigen Eigenvektoren, die wir zu λ_i finden (also $g_i = \mathrm{Nul}(\lambda_i \cdot E_n - A)$), so heißt g_i **geometrische Vielfachheit** des Eigenwertes λ_i. Es gilt:

$$A \text{ ist diagonalisierbar} \iff a_i = g_i \text{ für alle } i \tag{14.8}$$

Es ist nun naheliegend, dass wir uns fragen, ob alle Matrizen (reell) diagonalisierbar sind. Das ist nicht der Fall. Dazu betrachten wir die Matrix

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

Diese hat das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 + 1$$

und damit keinen reellen Eigenwert, also auch keine reellen Eigenvektoren. Damit ist sie auch nicht (über den reellen Zahlen) diagonalisierbar.

Allerdings ist das kein „echtes" Gegenbeispiel, denn das charakteristische Polynom hat zwei komplexe Nullstellen, nämlich $\lambda_1 = \mathrm{i}$ und $\lambda_2 = -\mathrm{i}$, und dazu gibt es in $\mathbb{C}^2$ auch Eigenvektoren: Für $\boldsymbol{v}_1 = \begin{pmatrix} -\mathrm{i} \\ 1 \end{pmatrix}$ ist $A \cdot \boldsymbol{v}_1 = \mathrm{i} \cdot \boldsymbol{v}_1$, also ist $\boldsymbol{v}_1$ ein komplexer Eigenvektor zu $\lambda_1 = \mathrm{i}$, und entsprechend ist $\boldsymbol{v}_2 = \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix}$ ein komplexer Eigenvektor zu $\lambda_2 = -\mathrm{i}$. Die zugehörige komplexe Matrix

$$S := \begin{pmatrix} -\mathrm{i} & \mathrm{i} \\ 1 & 1 \end{pmatrix}$$

ist invertierbar mit inverser Matrix

$$S^{-1} = \frac{1}{2} \cdot \begin{pmatrix} \mathrm{i} & 1 \\ -\mathrm{i} & 1 \end{pmatrix} .$$

Damit gilt dann auch

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} .$$

Die Matrix A ist also nicht über den reellen Zahlen diagonalisierbar, sehr wohl aber über den komplexen.

Beispiel

Betrachten wir die Matrix

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} ,$$

so hat diese das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 - 4\lambda + 4 ,$$

also nur einen Eigenwert $\lambda = 2$. Das zugehörige Gleichungssystem $(2\mathrm{E}_2 - A) \cdot \boldsymbol{x} = \boldsymbol{0}$ schreibt sich explizit als

$$\begin{aligned} 0 \cdot x_1 + 1 \cdot x_2 &= 0 \\ 0 \cdot x_1 + 0 \cdot x_2 &= 0 \end{aligned}$$

und eine Basis des Lösungsraumes ist gegeben durch

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .$$

Damit sind die Eigenvektoren von A zum Eigenwert $\lambda = 2$ die nichttrivialen Vielfachen von $\boldsymbol{v}$. Insbesondere gibt es keine zwei linear unabhängigen Eigenvektoren. Da es keinen weiteren Eigenwert gibt (auch nicht in den komplexen Zahlen), ist A nicht diagonalisierbar.

In der Tat ist A bereits in der optimalen Form, die wir hierfür durch Transformationen erreichen können. ◄

Symmetrische Matrizen sind diagonalisierbar

Es ist eine mühsame Aufgabe, die Diagonalisierbarkeit einer Matrix A von Fall zu Fall durch die Berechnung von Eigenwerten und Eigenvektoren zu überprüfen. In der Regel ist das leider die einzige Möglichkeit; erfreulicherweise gibt es aber zumindest eine Klasse von Matrizen, der wir die Diagonalisierbarkeit unmittelbar ansehen.

Symmetrische Matrizen sind diagonalisierbar

Ist A eine symmetrische $n \times n$-Matrix mit reellen Koeffizienten, so hat A nur reelle Eigenwerte und ist diagonalisierbar, und es gibt eine Orthonormalbasis von $\mathbb{R}^n$, bestehend aus Eigenvektoren von A.

Wir betrachten hierzu zunächst als Beispiel die Matrix

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

Diese hat den Eigenwert 6, und $\boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ist ein Eigenvektor dazu. Um zu einer Orthonormalbasis zu kommen, müssen wir diesen normieren:

$$\boldsymbol{v}_1 = \frac{1}{\sqrt{3}} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Wir ergänzen $\boldsymbol{v}_1$ beliebig zu einer Orthonormalbasis von $\mathbb{R}^3$, etwa durch die Vektoren

$$\boldsymbol{w}_2 = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} , \qquad \boldsymbol{w}_3 = \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

und erhalten somit eine orthogonale Matrix

$$O_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} .$$

Anwendung: Bestimmung von Eigenwerten und Eigenvektoren in MATLAB

Das charakteristische Polynom einer Matrix lässt sich in MATLAB mit dem Befehl `poly` ermitteln:

```
A = [2, 0, 2; 1, 2, 1; 1, 0, 3];
P = poly(A)

P =
     1    -7    14    -8
```

Die Ausgabe erfolgt als Koeffizientenvektor, beginnend mit dem Koeffizienten bei der höchsten Potenz (hier x^3), der bei einem charakteristischen Polynom immer den Wert 1 hat. Die Matrix

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

hat also das charakteristische Polynom

$$P_A(\lambda) = \lambda^3 - 7\lambda^2 + 14\lambda - 8\,,$$

und daraus erhalten wir die Eigenwerte $\lambda_1 = 1$ (durch Ausprobieren oder mithilfe von MATLAB), $\lambda_2 = 2$ und $\lambda_3 = 4$.

Eigenwerte einer Matrix lassen sich mit MATLAB aber auch direkt mit dem Befehl `eig` ermitteln:

```
A =  [0, 1, -2; 3, 2, -3; 2, 2, -3];
w = eig(A)

w =

        0 + 1.0000i
        0 - 1.0000i
  -1.0000
```

Die Matrix

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

hat also die Eigenwerte $\lambda_1 = -1$, $\lambda_2 = \mathrm{i}$ und $\lambda_3 = -\mathrm{i}$.

Auch Eigenvektoren und die Diagonalform einer Matrix werden in MATLAB mit dem Befehl `eig` ermittelt, also mit dem gleichen Befehl, der auch die Eigenwerte bestimmt. Die Syntax des Aufrufes ist dann allerdings etwas anders:

```
A=[0, 8, -4; -2, 8, -2; 0, 1, 1];
[S,D] = eig(A)

S =

    0.8452   -0.8729    0.8165
    0.5071   -0.4364    0.4082
    0.1690   -0.2182    0.4082

D =

    4.0000         0         0
         0    3.0000         0
         0         0    2.0000
```

Die Matrix S ist die Transformationsmatrix, enthält also linear unabhängige Eigenvektoren der übergebenen Matrix A als Spalten; D ist die resultierende Diagonalmatrix, hat daher die Eigenwerte von A in der Diagonalen.

Die Matrix

$$A = \begin{pmatrix} 0 & 8 & -4 \\ -2 & 8 & -2 \\ 0 & 1 & 1 \end{pmatrix}$$

hat also die Eigenwerte $\lambda_1 = 4$, $\lambda_2 = 3$ und $\lambda_3 = 2$ und die Transformationsmatrix

$$S = \begin{pmatrix} \frac{5}{\sqrt{35}} & -\frac{4}{\sqrt{21}} & \frac{2}{\sqrt{6}} \\ \frac{3}{\sqrt{35}} & -\frac{2}{\sqrt{21}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{35}} & -\frac{1}{\sqrt{21}} & \frac{1}{\sqrt{6}} \end{pmatrix}.$$

Zu beachten ist dabei, dass MATLAB die Eigenvektoren auf die Länge 1 normiert; die numerische Ausgabe ist dann natürlich in Dezimalzahlen (und enthält dementsprechend in der Regel Rundungsfehler).

Der Fall von Matrizen, die nicht diagonalisierbar sind, erfordert allerdings eine gesonderte Betrachtung:

```
A=[2,1;0,2];
[S,D] = eig(A)
S =

    1.0000   -1.0000
         0    0.0000
D =
     2     0
     0     2
```

Für die Matrix $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ liefert `eig` als „Transformationsmatrix“ $S = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ und die „Diagonalmatrix“ $D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, obwohl A nicht diagonalisierbar ist.

Die Matrix S bringt A nur dann auf Diagonalgestalt, wenn S invertierbar ist. Daher muss geprüft werden, ob $\det S \neq 0$.

Teil III

14.2 Mathematischer Hintergrund: Eigenwerte und Page Ranking

Bei großen Matrizen kann die Bestimmung von Eigenwerten und Eigenvektoren sehr rechenintensiv werden. Es gibt jedoch einen Spezialfall, der sich vergleichsweise einfach algorithmisch behandeln lässt.

Wir betrachten eine diagonalisierbare Matrix A und nehmen an, dass für ihre Eigenwerte $\lambda_1, \ldots, \lambda_n$

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$$

gilt, dass es also einen eindeutigen bestimmten betragsmäßig größten Eigenwert gibt. Ist dann $\boldsymbol{v}$ ein „generischer" Vektor und setzen wir

$$\boldsymbol{v}^{(0)} = \boldsymbol{v}, \qquad \boldsymbol{v}^{(k+1)} = \frac{1}{\lambda_1} \cdot A \cdot \boldsymbol{v}^{(k)} \quad (k \geq 0),$$

so konvergiert die Folge $\boldsymbol{v}^{(k)}$ gegen einen Eigenvektor von A zum Eigenwert λ_1. Dazu schreiben wir

$$\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_n\,,$$

wobei $\boldsymbol{w}_i$ ein Eigenvektor zum Eigenwert λ_i ist. Das ist möglich, da es aufgrund der Diagonalisierbarkeit von A eine Basis, bestehend aus Eigenvektoren von A, gibt. „Generisch" bedeutet in diesem Zusammenhang, dass $\boldsymbol{w}_1 \neq \boldsymbol{0}$. Dann gilt

$$\boldsymbol{v}^{(1)} = A \cdot \boldsymbol{v} = \boldsymbol{w}_1 + \frac{\lambda_2}{\lambda_1} \cdot \boldsymbol{w}_2 + \cdots + \frac{\lambda_n}{\lambda_1} \cdot \boldsymbol{w}_n$$

und allgemein

$$\boldsymbol{v}^{(k)} = \boldsymbol{w}_1 + \frac{\lambda_2^k}{\lambda_1^k} \cdot \boldsymbol{w}_2 + \cdots + \frac{\lambda_n^k}{\lambda_1^k} \cdot \boldsymbol{w}_n\,.$$

Wegen $|\lambda_m| < |\lambda_1|$ für $m > 1$ ist

$$\lim_{k\to\infty} \frac{\lambda_m^k}{\lambda_1^k} = 0 \qquad \text{für } m > 1\,.$$

Hieraus folgt, dass

$$\lim_{k\to\infty} \boldsymbol{v}^{(k)} = \boldsymbol{w}_1$$

existiert und ein Eigenvektor von A zum Eigenwert λ_1 ist. Ist die Matrix A zusammenhängend, d. h. gibt es für je zwei Indizes i, j immer Indizes $i_0, i_1, \ldots, i_t$ mit $i = i_0$ und $j = i_t$, sodass $a_{i_k, i_{k+1}} \neq 0$ für $k = 0, \ldots, t-1$, so ist jeder Einheitsvektor generisch, kann also als Startvektor benutzt werden.

Eine der bekanntesten Anwendungen der Eigenwerttheorie und dieser numerischen Methode zur Bestimmung eines Eigenvektors ist das **Page Ranking**, wie es etwa von Google oder anderen Suchmaschinen verwendet wird. Das Prinzip kann dabei wie folgt beschrieben werden: Die Grundidee hinter dem Page Ranking ist, dass eine Seite im Internet umso wichtiger ist, je mehr andere wichtige Seiten einen Link auf diese Seite enthalten. Diese Idee ist natürlich zunächst mathematisch ungenau, da hier die Wichtigkeit einer Seite mit der Wichtigkeit anderer Seiten erklärt wird; sie soll aber nun präzise gemacht werden.

Für eine Website W_j bezeichnen wir mit l_j die Anzahl aller Links von W_j auf eine andere Website, mit $l_{i,j}$ die Anzahl aller Links von Website W_j zur Website W_i und mit V_i die Menge aller Websites, die einen Link auf W_i haben. Wir nehmen an, dass es zu jeder Seite mindestens eine andere gibt, die einen Link auf diese Seite hat. Damit gilt in jedem Fall

$$\sum_{i: j \in V_i} l_{i,j} = l_j\,,$$

und wir definieren die **Hyperlinkmatrix** $P = \big(p_{i,j}\big)$ durch

$$p_{i,j} = \begin{cases} \frac{l_{i,j}}{l_j}\,, & \text{falls } W_j \in V_i\,, \\ 0\,, & \text{sonst.} \end{cases}$$

In der i-ten Zeile und der j-ten Spalte steht also der Quotient aus der Anzahl der Links, die von Website W_j auf Website W_i zeigen, und der Anzahl der Links, die von Website W_j ausgehen. Damit ist $p_{i,j}$ der Anteil der von W_j ausgehenden Links, der auf W_i zeigt und deshalb gilt für jede Spalte j

$$\sum_i p_{i,j} = \sum_{i: j \in V_i} \frac{l_{i,j}}{l_j} = \frac{l_j}{l_j} = 1\,,$$

weshalb der Vektor

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

ein Eigenvektor von $P^\top$ zum Eigenwert 1 ist, wie wir sofort nachrechnen und bereits in der Anwendung „Eigenwerte und stabile Systeme" aus Abschn. 14.1 gesehen haben. Damit ist aber 1 auch ein Eigenwert von P, da P und $P^\top$ dieselben Eigenwerte haben. Wir behaupten nun, dass alle anderen Eigenwerte λ von P die Ungleichung $|\lambda| \leq 1$ erfüllen.

Dazu reicht es natürlich wieder, $P^\top$ zu betrachten. Ist λ ein Eigenwert von $P^\top$ und ist $\boldsymbol{v}$ ein zugehöriger Eigenvektor, so betrachten wir dessen betragsmäßig größte Komponente v_j (d. h. $|v_i| \leq |v_j|$ für alle i). Dann gilt für die j-te Komponente w_j von $P^\top \cdot \boldsymbol{v}$ einerseits $w_j = \lambda \cdot v_j$ und andererseits

$$|w_j| = |\sum_{i: j \in V_i} p_{i,j} v_i| \leq \sum_{i: j \in V_i} p_{i,j} |v_i| \leq \sum_{i \in L_j} p_{i,j} |v_j| = |v_j|\,,$$

sodass notwendig $|\lambda| \leq 1$. Nehmen wir zusätzlich an, dass P diagonalisierbar ist und dass alle Eigenwerte λ von P außer der 1 die Ungleichung $|\lambda| < 1$ erfüllen, so können wir mit oben beschriebenem Algorithmus einen Eigenvektor $\boldsymbol{v}$ von P zum Eigenwert 1 finden. Bezeichnen wir mit v_i dessen i-te Komponente, so gilt

$$v_i = (P \cdot \boldsymbol{v})_i = \sum_{j:W_j \in V_i} \frac{l_{i,j}}{l_j} \cdot v_j \,,$$

und damit berechnet sich v_i aus den relativen Häufigkeiten der Links auf Website W_i, die von anderen Sites W_j ausgehen, gewichtet mit einem Wichtigkeitsfaktor v_j der verlinkenden Website W_j. Ist daher i der Index für den v_i maximal ist, so können wir W_i als die bedeutendste Website interpretieren.

Wir wollen dazu ein einfaches Beispiel mit fünf Internetseiten betrachten, die, wie in Abb. 14.1 gezeigt, aufeinander verweisen.

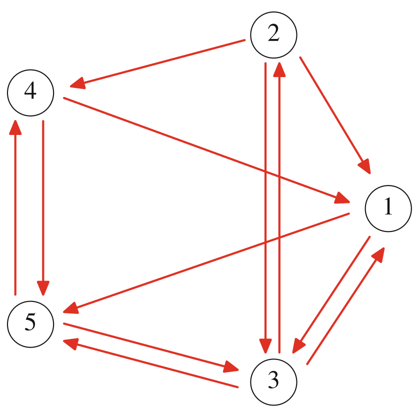


Abb. 14.1 Verlinkte Seiten im Internet

Von Seite 1 gehen also zwei Links aus, wovon einer auf Seite 3 und einer auf Seite 5 zeigt (womit $1 \in V_3$, $1 \in V_5$ und $1 \notin V_j$ für $j = 1, 2, 4$) , und es verweisen die Seiten 2, 3 und 4 auf Seite 1, also

$$\begin{aligned} l_1 &= 2,\ V_1 = \{2,\ 3,\ 5\}\,, \\ l_{3,1} &= l_{5,1} = 1\,, \\ l_{1,2} &= l_{1,3} = l_{1,4} = 1\,. \end{aligned}$$

Führen wir diese Betrachtung für alle Internetseiten dieses Netzes durch, so erhalten wir die Hyperlinkmatrix

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{pmatrix}.$$

Damit ist die Matrix P zusammenhängend. Ferner ist sie auch diagonalisierbar, wie man z. B. mit MATLAB nachrechnet. Einen Eigenvektor zum Eigenwert $\lambda = 1$ erhalten wir durch das skizzierte approximative Verfahren

$$\boldsymbol{v}^{(0)} = \mathbf{e}_1,\ \boldsymbol{v}^{(1)} = P \cdot \boldsymbol{v}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix}, \ldots, \boldsymbol{v}^{(50)} = \begin{pmatrix} 0.2024 \\ 0.0893 \\ 0.2679 \\ 0.1667 \\ 0.2738 \end{pmatrix},$$

und hierfür gilt bereits im Rahmen der Rechengenauigkeit von vier Dezimalstellen: $A \cdot \boldsymbol{v}^{(50)} = \boldsymbol{v}^{(50)}$.

In diesem Internet ist also W_5 die wichtigste Seite. Diese Seite ist sogar wichtiger als die Seiten W_1 und W_3, obwohl es drei Links auf W_1 und W_3 gibt, genauso wie auf W_5. Bei diesen Berechnungen wird nämlich auch berücksichtigt, ob die Seiten, die auf ein W_i verweisen, selbst wieder von vielen anderen wichtigen Seiten aus verlinkt sind. Dadurch wird sichergestellt, dass die Verweise von Seiten, auf die selbst (direkt oder indirekt) häufig verwiesen wird, stärker in die Gewichtung eingehen als Seiten, auf die es keinen Link von einer anderen Webseite gibt.

Wir haben hier einige vereinfachende Annahmen gemacht, etwa dass die Hyperlinkmatrix diagonalisierbar ist, dass sie zusammenhängend ist (dass das Internet also nicht in mehrere, voneinander unabhängige und nicht durch Links verbundene Teilnetze zerfällt), dass es auf alle Seiten des Netzes mindestens einen Verweis gibt und dass 1 der betragsmäßig eindeutig größte Eigenwert ist (dass es also keinen weiteren Eigenwert mit Betrag 1 gibt), damit unser approximatives Verfahren auch zum Ziel führt. Das wird bei der **Googlematrix** G, also bei der Matrix, die Google (oder andere Suchmaschinen) dann tatsächlich benutzen, um die Wichtigkeit von Internetseiten zu bestimmen, sichergestellt, indem die eigentliche Hyperlinkmatrix A durch geeignete Parameter und Matrizen etwas modifiziert und gestört wird. Die Matrix G hat mehr als 10^9 Zeilen und Spalten, allerdings ist sie sehr dünn besetzt, da Internetseiten im Durchschnitt nur zehn Links haben. Aus diesem Grund dauert ein Approximationsverfahren nur wenige Tage. Google führt diese approximativen Eigenvektorberechnungen etwa einmal pro Monat durch.

Eine genauere Abhandlung des Page Rankings finden Sie in Bryan und Leise (2006).

Literatur

K. Bryan, T. Leise: *The \$ 25,000,000,000 Eigenvector: The Linear Algebra behind Google*. SIAM Review 48(3), 569–581. 2006

Teil III

Mit diesem O_1 als Transformationsmatrix berechnen wir

$$B = O_1^\top \cdot A \cdot O_1 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{3}{2} \end{pmatrix}.$$

Damit ist B noch keine Diagonalmatrix, aber erste Zeile und erste Spalte von B haben die gewünschte Gestalt, und die verbleibende Matrix aus zweiter und dritter Zeile und Spalte,

$$B_1 = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{3}{2} \end{pmatrix},$$

ist wieder ein symmetrische Matrix. Diese Matrix hat 0 als Eigenwert, und ein normierter Eigenvektor zum Eigenwert 0 ist $\boldsymbol{u}_1 = \frac{1}{2} \cdot \begin{pmatrix} -\sqrt{3} \\ 1 \end{pmatrix}$. Ergänzen wir $\boldsymbol{u}_1$ zu einer Orthonormalbasis von $\mathbb{R}^2$, etwa durch $\boldsymbol{u}_2 = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$ und bilden die zugehörige orthogonale Matrix

$$U = \frac{1}{2} \cdot \begin{pmatrix} -\sqrt{3} & 1 \\ 1 & \sqrt{3} \end{pmatrix},$$

so erhalten wir durch Transformation von B_1 mit U die Matrix

$$B_2 = U^\top \cdot B_1 \cdot U = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}.$$

Wir bekommen also eine Diagonalmatrix, obwohl wir nur $\boldsymbol{u}_1$ bewusst als Eigenvektor von B_1 gewählt haben. Wir ergänzen nun U zu einer orthogonalen 3×3-Matrix O_2,

$$O_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix},$$

und bilden

$$O = O_1 \cdot O_2 = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \end{pmatrix}.$$

Dann gilt hierfür

$$O^\top \cdot A \cdot O = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

und die Spalten von O bilden eine Orthonormalbasis von $\mathbb{R}^3$, bestehend aus Eigenvektoren von A.

Diese Idee, zunächst nur einen normierten Eigenvektor zu betrachten, diesen in beliebiger Weise zu einer Orthonormalbasis zu ergänzen und dann mit einer kleineren Matrix weiterzumachen, funktioniert ganz allgemein für beliebige symmetrische Matrizen. Wir benutzen dabei das Prinzip der vollständigen Induktion nach n, der Anzahl der Zeilen bzw. Spalten von A (wie in Mathematischer Hintergrund 4.1 beschrieben).

Für $n = 1$ ist $A = (a)$, und wir haben nichts zu zeigen.

Wir nehmen also an, dass $n > 1$ und dass wir die Behauptung für $n - 1$ schon gezeigt haben.

Wir behaupten zunächst, dass A nur reelle Eigenwerte besitzt. Da das charakteristische Polynom $P_A(\lambda)$ in den komplexen Zahlen Nullstellen hat, gibt es sicherlich einen komplexen Eigenwert λ_0 von A und daher auch einen komplexen Vektor $\boldsymbol{v} \in \mathbb{C}^n$, verschieden vom Nullvektor, mit $A \cdot \boldsymbol{v} = \lambda_0 \cdot \boldsymbol{v}$. Bezeichnen wir mit $\langle\ ,\ \rangle$ das komplexe Skalarprodukt, so gilt nach der Wälzformel (12.17)

$$\begin{aligned} \lambda_0 \cdot |\boldsymbol{v}|^2 &= \lambda_0 \cdot \langle \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \langle \lambda_0 \cdot \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \langle A \cdot \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \langle \boldsymbol{v}, A \cdot \boldsymbol{v} \rangle \\ &= \langle \boldsymbol{v}, \lambda_0 \cdot \boldsymbol{v} \rangle \\ &= \overline{\lambda_0} \cdot \langle \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \overline{\lambda_0} \cdot |\boldsymbol{v}|^2 \end{aligned}$$

und daher, da $|\boldsymbol{v}|^2 \neq 0$, auch $\lambda_0 = \overline{\lambda_0}$. Eine komplexe Zahl und ihre komplex-konjugierte Zahl stimmen aber nur dann überein, wenn die Zahl schon reell ist (also ihr Imaginärteil verschwindet). Damit haben wir gezeigt, dass jeder Eigenwert von A schon reell ist

Nun geben wir uns einen Eigenwert λ_1 von A vor und bestimmen einen Eigenvektor $\boldsymbol{u}_1$ dazu, wobei wir $\boldsymbol{u}_1$ auf die Länge 1 normieren. Wir ergänzen (z. B. mit dem Gram-Schmidt-Orthonormalisierungsverfahren) $\boldsymbol{u}_1$ zu einer Orthonormalbasis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ von $\mathbb{R}^n$.

Der entscheidende Punkt für unsere Überlegungen ist nun, dass jeder Vektor $A \cdot \boldsymbol{u}_i$ ($i = 2, \ldots, n$) schon in dem von $\boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ aufgespannten Untervektorraum enthalten ist. Dazu geben wir uns ein $i \in \{2, \ldots, n\}$ vor, schreiben

$$A \cdot \boldsymbol{u}_i = r_1 \boldsymbol{u}_1 + r_2 \boldsymbol{u}_2 + \cdots + r_n \boldsymbol{u}_n$$

und haben zu zeigen, dass $r_1 = 0$. Hierfür rechnen wir

$$\begin{aligned} r_1 &= r_1 \cdot \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle \\ &= \sum_{i=1}^{n} r_i \cdot \langle \boldsymbol{u}_i, \boldsymbol{u}_1 \rangle \\ &= \langle r_1 \boldsymbol{u}_1 + r_2 \boldsymbol{u}_2 + \cdots + r_n \boldsymbol{u}_n,\ \boldsymbol{u}_1 \rangle \\ &= \langle A \cdot \boldsymbol{u}_i, \boldsymbol{u}_1 \rangle \\ &= \langle \boldsymbol{u}_i,\ A \cdot \boldsymbol{u}_1 \rangle \\ &= \langle \boldsymbol{u}_i,\ \lambda_1 \cdot \boldsymbol{u}_1 \rangle \\ &= \lambda_1 \cdot \langle \boldsymbol{u}_i, \boldsymbol{u}_1 \rangle \\ &= 0\,. \end{aligned}$$

Teil III

Dabei haben wir, neben den üblichen Eigenschaften des Skalarprodukts, in der zweiten und in der letzten Zeile ausgenutzt, dass $\langle \boldsymbol{u}_i, \boldsymbol{u}_1\rangle = 0 = \langle \boldsymbol{u}_1, \boldsymbol{u}_i\rangle$ für $i \neq 1$ und $\langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle = 1$ in der ersten Zeile.

Wir betrachten nun die Matrix $U = (\boldsymbol{u}_1 \ \ldots \ \boldsymbol{u}_n)$, die die Vektoren $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ als Spaltenvektoren hat. Da diese Spalten eine Orthonormalbasis von $\mathbb{R}^n$ bilden, ist U eine orthogonale Matrix, sodass also U nach Abschn. 13.4 invertierbar mit $U^{-1} = U^T$ ist. Für die Matrix

$$B = U^{-1} \cdot A \cdot U = U^T \cdot A \cdot U$$

gilt dann

$$\begin{aligned} B \cdot \mathbf{e}_1 &= U^{-1} \cdot A \cdot U \cdot \mathbf{e}_1 \\ &= U^{-1} \cdot A \cdot \boldsymbol{u}_1 \\ &= U^{-1} \cdot \lambda_1 \cdot \boldsymbol{u}_1 \\ &= \lambda_1 \cdot U^{-1} \cdot \boldsymbol{u}_1 \\ &= \lambda_1 \cdot \mathbf{e}_1 , \end{aligned}$$

und für $i \geq 2$ gilt $A \cdot \boldsymbol{u}_i = r_{2,i}\boldsymbol{u}_2 + \cdots + r_{n,i}\boldsymbol{u}_n$, wie wir oben gesehen haben, und damit

$$\begin{aligned} B \cdot \mathbf{e}_i &= U^{-1} \cdot A \cdot U \cdot \mathbf{e}_i \\ &= U^{-1} \cdot A \cdot \boldsymbol{u}_i \\ &= U^{-1} \cdot (r_{2,i}\boldsymbol{u}_2 + \cdots + r_{n,i}\boldsymbol{u}_n) \\ &= r_{2,i} \cdot U^{-1} \cdot \boldsymbol{u}_2 + \cdots + r_{n,i} \cdot U^{-1} \cdot \boldsymbol{u}_n \\ &= r_{2,i}\mathbf{e}_2 + \cdots + r_{n,i}\mathbf{e}_n . \end{aligned}$$

Also hat $U^{-1} \cdot A \cdot U$ die Gestalt

$$U^{-1} \cdot A \cdot U = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}$$

mit einer $(n-1) \times (n-1)$-Matrix C. Genauer ist $C = \big(r_{i,j}\big)$ mit den oben bestimmten $r_{i,j}$. Nun gilt für beliebige Vektoren $\boldsymbol{v}$ und $\boldsymbol{w}$ aufgrund der Formeln (12.16) und (12.17) sowie der Orthogonalität von U:

$$\begin{aligned} \langle U^{-1} \cdot A \cdot U \cdot \boldsymbol{v},\, \boldsymbol{w}\rangle &= \langle U^T \cdot A \cdot U \cdot \boldsymbol{v},\, \boldsymbol{w}\rangle \\ &= \langle A \cdot U \cdot \boldsymbol{v},\, \big(U^T\big)^T \cdot \boldsymbol{w}\rangle \\ &= \langle A \cdot U \cdot \boldsymbol{v},\, U \cdot \boldsymbol{w}\rangle \\ &= \langle U \cdot \boldsymbol{v},\, A \cdot U \cdot \boldsymbol{w}\rangle \\ &= \langle \boldsymbol{v},\, U^T \cdot A \cdot U \cdot \boldsymbol{w}\rangle \\ &= \langle \boldsymbol{v},\, U^{-1} \cdot A \cdot U \cdot \boldsymbol{w}\rangle , \end{aligned}$$

sodass die Matrix $U^{-1} \cdot A \cdot U$ nach Formel (12.18) symmetrisch ist. Damit muss aber auch die Matrix C symmetrisch sein, und wir können unsere Induktionsvoraussetzung anwenden und erhalten, dass C diagonalisierbar ist. Wir finden also eine orthogonale $(n-1) \times (n-1)$-Matrix V, sodass $V^{-1} \cdot C \cdot V$ eine Diagonalmatrix D ist. Setzen wir

$$W := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & V & \\ 0 & & & \end{pmatrix} ,$$

so ist auch diese Matrix orthogonal, und nach den Regeln für die Matrizenrechnung ist

$$W^{-1} \cdot U^{-1} \cdot A \cdot U \cdot W = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{pmatrix} .$$

Setzen wir also $S = U \cdot W$, so hat $S^{-1} \cdot A \cdot S$ Diagonalgestalt, wie gewünscht.

Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix} .$$

Dann ist A symmetrisch und damit diagonalisierbar. Hier sehen wir aber auch sofort, dass

$$P_A(\lambda) = \lambda^2 - 2\lambda - 24 = (\lambda - 6) \cdot (\lambda + 4) ,$$

sodass wir also zwei verschiedene Eigenwerte haben. Ein Eigenvektor zu $\lambda_1 = 6$ ist $\boldsymbol{u}_1 = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, und ein Eigenvektor zu $\lambda_2 = -4$ ist $\boldsymbol{u}_2 = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Die Vektoren $\boldsymbol{u}_1, \boldsymbol{u}_2$ bilden in der Tat eine Orthonormalbasis von $\mathbb{R}^2$, und mit $S := \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ gilt dann

$$S^\top \cdot A \cdot S = \begin{pmatrix} -4 & 0 \\ 0 & 6 \end{pmatrix} .$$

◂

Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 2 & 4 \\ 2 & 4 & 2 \end{pmatrix} . \tag{14.9}$$

Auch hier ist A symmetrisch, und daher sehen wir ohne Rechnung, dass A diagonalisierbar ist. Hier gilt

$$P_A(\lambda) = \lambda^3 - 9\lambda^2 + 108 = (\lambda - 6) \cdot (\lambda - 6) \cdot (\lambda + 3)\,.$$

Diese Matrix hat also nur zwei Eigenwerte, und an der Anzahl der Eigenwerte können wir die Diagonalisierbarkeit nicht ablesen. Wenn wir jedoch die Eigenvektoren berechnen, erhalten wir für $\lambda_1 = 6$ die beiden linear unabhängigen und orthogonalen Eigenvektoren

$$\boldsymbol{u}_1 = \frac{1}{3} \cdot \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \frac{1}{3} \cdot \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$$

und für $\lambda_2 = -3$ den Eigenvektor

$$\boldsymbol{u}_3 = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

Die Vektoren $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ bilden eine Orthonormalbasis von $\mathbb{R}^3$, und für die Matrix

$$S = \frac{1}{3} \cdot \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix}$$

gilt

$$S^\top \cdot A \cdot S = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -3 \end{pmatrix}. \qquad \blacktriangleleft$$

Die Symmetrie einer Matrix hat auch Auswirkungen auf die Berechnungsmöglichkeiten der Lösungen von Gleichungssystemen. Ist nämlich A eine symmetrische $n \times n$-Matrix, bei der alle Eigenwerte positiv sind, so ist $\boldsymbol{v} \in \mathbb{R}^n$ genau dann eine Lösung des linearen Gleichungssystems $A \cdot \boldsymbol{x} = \boldsymbol{b}$, wenn $\boldsymbol{v}$ die Funktion $f(\boldsymbol{x}) = \frac{1}{2} \cdot \boldsymbol{x}^\top \cdot A \cdot \boldsymbol{x} - \boldsymbol{x}^\top \cdot \boldsymbol{b}$ minimiert. Dieses Minimierungsproblem kann sehr gut mit Näherungsverfahren gelöst werden, etwa dem mehrdimensionalen Newton-Verfahren (Bd. 2, Abschn. 2.9).

Definition

Eine komplexe $n \times n$-Matrix heißt **hermitesch**, wenn $A^\top = \overline{A}$.

Die Matrix $A = \begin{pmatrix} 1 & 1+\mathrm{i} \\ 1-\mathrm{i} & 2 \end{pmatrix}$ ist hermitesch.

Hermitesche Matrizen sind diagonalisierbar

Eine hermitesche Matrix A hat nur reelle Eigenwerte und ist (über $\mathbb{C}$) diagonalisierbar. Die Transformationsmatrix U kann unitär gewählt werden (also so, dass $U^{-1} = \overline{U}^\top$).

Beispiel

Die hermitesche Matrix $A = \begin{pmatrix} 1 & 1+\mathrm{i} \\ 1-\mathrm{i} & 2 \end{pmatrix}$ hat die Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 3$. Eine unitäre Transformationsmatrix, die A auf Diagonalgestalt bringt, ist

$$U = \begin{pmatrix} \frac{\sqrt{3}}{3} + \frac{\sqrt{3}}{3} \cdot \mathrm{i} & \frac{\sqrt{6}}{6} + \frac{\sqrt{6}}{6} \cdot \mathrm{i} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{pmatrix}. \qquad \blacktriangleleft$$

Orthogonale Matrizen sind im Komplexen diagonalisierbar

Eine wichtige Klasse von Matrizen sind die orthogonalen Matrizen. Ist A eine orthogonale $n \times n$-Matrix und ist λ ein komplexer Eigenwert von A zum Eigenvektor $\boldsymbol{v}$, so muss notwendigerweise $|\lambda| = 1$ gelten, da wir ja in Abschn. 13.4 gesehen haben, dass

$$|\boldsymbol{v}| = |A \cdot \boldsymbol{v}| = |\lambda| \cdot |\boldsymbol{v}|\,.$$

Die einzigen reellen Eigenwerte, die A haben kann, sind daher 1 und -1. Im Regelfall hat A aber noch weitere (komplexe) Eigenwerte und ist daher im Allgemeinen nicht über den reellen Zahlen diagonalisierbar. Es gilt jedoch: Es gibt eine unitäre $n \times n$-Matrix U, sodass

$$D = \overline{U}^\top \cdot A \cdot U$$

eine Diagonalmatrix ist.

Im Reellen können wir für orthogonale Matrizen die folgende Normalform erreichen:

Normalform orthogonaler Matrizen

Es gibt eine (reelle) orthogonale $n \times n$-Matrix O, sodass

$$O^\top \cdot A \cdot O = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & D_t \end{pmatrix},$$

wobei die D_k Matrizen sind, die eine der folgenden Gestalten haben:

1. D_k ist eine 1×1-Matrix, $D_k = (\pm 1)$.
2. D_k ist eine 2×2-Matrix, und es gibt ein $\varphi \in (0, 2\pi)$, $\varphi \neq \pi$ mit

$$D_k = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Zu der Matrix D_k in 1. gehört eine Achse A in $\mathbb{R}^n$ (also ein eindimensionaler Untervektorraum von $\mathbb{R}^n$), die durch A festgehalten oder auf ihr Negatives abgebildet wird.

Die Matrix D_k in 2. definiert eine Drehung der Ebene, die von den beiden zugehörigen Spalten von O aufgespannt wird, um den Winkel φ.

Beispiel

Die orthogonale 3×3-Matrix $A = \frac{1}{3} \cdot \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix}$ ist nicht reell diagonalisierbar. Sie wird durch die orthogonale Matrix $O = \begin{pmatrix} -\frac{1}{\sqrt{11}} & \frac{1}{\sqrt{2}} & \frac{3}{\sqrt{22}} \\ -\frac{1}{\sqrt{11}} & -\frac{1}{\sqrt{2}} & \frac{3}{\sqrt{22}} \\ \frac{3}{\sqrt{11}} & 0 & \frac{2}{\sqrt{22}} \end{pmatrix}$ auf die Gestalt

$$O^\top \cdot A \cdot O = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{5}{6} & -\frac{\sqrt{11}}{6} \\ 0 & \frac{\sqrt{11}}{6} & \frac{5}{6} \end{pmatrix}$$

gebracht. Die Matrix A dreht also die Ebene, die von

$$\boldsymbol{p} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \boldsymbol{q} = \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}$$

aufgespannt wird, um $\arccos\left(\frac{5}{6}\right)$ und spiegelt $\mathbb{R}^3$ an dieser Ebene. ◀

Ganz allgemein hat jede orthogonale 3×3–Matrix A entweder 1 oder -1 als Eigenwert, da in der Normalform nicht nur der zweite Fall auftreten kann. Falls sie noch weitere (komplexe) Eigenwerte hat, gibt es eine Ebene $E \subset \mathbb{R}^3$ und einen Winkel φ, sodass E durch A um φ gedreht wird. Die Normalengerade zu E wird durch A entweder festgehalten (1 ist Eigenwert) oder an E gespiegelt (-1 ist Eigenwert).

Auch lineare Abbildungen können diagonalisierbar sein

Um die Diagonalisierbarkeit von Matrizen besser zu verstehen, ist es hilfreich, sich an die Beziehungen zwischen Matrizen und linearen Abbildungen (Abschn. 12.3) zu erinnern. Zu einer Matrix A gehört eine lineare Abbildung f, und A ist die darstellende Matrix von f bzgl. der Standardbasis. Die Standardbasis ist aber relativ willkürlich dadurch entstanden, dass wir einen Koordinatenursprung gewählt und dann n Koordinatenachsen festgelegt haben. Genauso gut hätten wir auch andere Achsen wählen können. Wählen wir im $\mathbb{R}^n$ andere Koordinatenachsen, so können wir diese (ausgehend vom ursprünglichen Koordinatensystem) durch n Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ beschreiben. Schreiben wir diese Vektoren als Spalten einer Matrix, $S = (\boldsymbol{v}_1 \ \ldots \ \boldsymbol{v}_n)$, so ist diese Matrix invertierbar (das ist sogar äquivalent damit, dass $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ ein neues Koordinatensystem von $\mathbb{R}^n$ definieren), und in diesem neuen Koordinatensystem wird die lineare Abbildung f durch die Matrix

$$B = S^{-1} \cdot A \cdot S$$

beschrieben. Sind also speziell $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ jetzt n linear unabhängige Eigenvektoren von A, so bedeutet das gerade, dass $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ ein neues Koordinatensystem von $\mathbb{R}^n$ definieren, und in diesem Koordinatensystem wird die lineare Abbildung durch die Diagonalmatrix

$$D = S^{-1} \cdot A \cdot S$$

beschrieben. Bezüglich der durch $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ gegebenen Koordinaten gilt also

$$f(\boldsymbol{v}_i) = \lambda_i \cdot \boldsymbol{v}_i \qquad \text{für alle } i \in \{1, \ldots, n\}$$

oder ganz allgemein für einen beliebigen Vektor $\boldsymbol{w} = \sum_{i=1}^{n} r_i \cdot \boldsymbol{v}_i$

$$f\left(\sum_{i=1}^{n} r_i \cdot \boldsymbol{v}_i\right) = \sum_{i=1}^{n} r_i \cdot \lambda_i \cdot \boldsymbol{v}_i\,.$$

14.3 Normalformen

Wie wir bereits gesehen haben, sind nicht alle Matrizen diagonalisierbar, und selbst reelle Matrizen, die diagonalisierbar sind, sind nicht immer über den reellen Zahlen diagonalisierbar. Auch für diese Fälle gibt es verschiedene Möglichkeiten, eine Matrix A in eine Form zu bringen, die die Arbeit mit der Matrix vereinfacht.

Alle Matrizen lassen sich trigonalisieren

Für viele Anwendungen, etwa im Zusammenhang mit Gleichungssystemen, ist es sehr hilfreich, wenn eine Matrix in oberer Dreiecksform vorliegt.

Teil III

14.3 Mathematischer Hintergrund: Symmetrische Bilinearformen

In Mathematischer Hintergrund 11.3 haben wir bereits symmetrische Bilinearformen auf $\mathbb{R}^n$ behandelt, also Abbildungen $\beta : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, die linear in beiden Komponenten sind und für die $\beta(\boldsymbol{x},\boldsymbol{y}) = \beta(\boldsymbol{y},\boldsymbol{x})$ gilt, und in der Fortführung dieses Themas haben wir in Mathematischer Hintergrund 12.2 gesehen, dass symmetrische Bilinearformen durch symmetrische Matrizen beschrieben werden, dass es also eine symmetrische $n \times n$-Matrix A gibt mit $\beta(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^\top \cdot A \cdot \boldsymbol{y}$.

In diesem Abschnitt haben wir festgestellt, dass eine symmetrische Matrix A durch eine orthogonale Matrix O diagonalisierbar ist, d. h., dass es eine orthogonale Matrix O gibt, sodass $D = O^\top \cdot A \cdot O$ eine Diagonalmatrix ist. Mit $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ bezeichnen wir die Spalten von O (die also eine Orthonormalbasis von $\mathbb{R}^n$, bestehend aus Eigenvektoren von A, bilden) und mit $\lambda_1, \ldots, \lambda_n$ die Diagonalelemente von D. Beachten Sie dabei, dass die λ_i nicht paarweise verschieden sein müssen.

Beliebige Vektoren $\boldsymbol{x}$ und $\boldsymbol{y}$ können mithilfe der Basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ beschrieben werden,

$$\boldsymbol{x} = r_1 \cdot \boldsymbol{v}_1 + \cdots + r_n \cdot \boldsymbol{v}_n, \qquad \boldsymbol{y} = s_1 \cdot \boldsymbol{v}_1 + \cdots + s_n \cdot \boldsymbol{v}_n,$$

wobei die Umrechnung in die Standardbasisdarstellung nach der Formel

$$\boldsymbol{x} = O \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}, \quad \boldsymbol{y} = O \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$$

erfolgt. Damit gilt

$$\begin{aligned}
\beta(\boldsymbol{x},\boldsymbol{y}) &= \boldsymbol{x}^\top \cdot A \cdot \boldsymbol{y} \\
&= \left(O \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right)^\top \cdot A \cdot O \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \\
&= (r_1 \ \ldots \ r_n) \cdot O^\top \cdot A \cdot O \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \\
&= (r_1 \ \ldots \ r_n) \cdot D \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \\
&= \lambda_1 \cdot r_1 \cdot s_1 + \lambda_2 \cdot r_2 \cdot s_2 + \cdots + \lambda_n \cdot r_n \cdot s_n .
\end{aligned}$$

Zu jeder symmetrischen Bilinearform β gibt es also eine Orthonormalbasis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ von $\mathbb{R}^n$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, sodass bzgl. dieser Orthonormalbasis β eine besonders einfache Form hat:

Sind $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ zwei Vektoren und schreiben wir

$$\boldsymbol{v} = r_1 \cdot \boldsymbol{v}_1 + \cdots + r_n \cdot \boldsymbol{v}_n, \quad \boldsymbol{w} = s_1 \cdot \boldsymbol{v}_1 + \cdots + s_n \cdot \boldsymbol{v}_n,$$

so gilt

$$\beta(\boldsymbol{v}, \boldsymbol{w}) = \lambda_1 \cdot r_1 \cdot s_1 + \cdots + \lambda_n \cdot r_n \cdot s_n .$$

Speziell gilt also

$$\beta(\boldsymbol{v}, \boldsymbol{v}) = \lambda_1 \cdot r_1^2 + \cdots + \lambda_n \cdot r_n^2 .$$

Wir sagen in diesem Fall, dass die symmetrische Bilinearform durch die orthogonale Matrix O bzw. durch die Orthonormalbasis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ von $\mathbb{R}^n$ diagonalisiert wird.

Eine besondere Rolle spielt dabei die Anzahl der positiven und der negativen Eigenwerte. Sind alle Eigenwerte positiv, so ist die symmetrische Bilinearform positiv definit; sind alle negativ, so ist sie negativ definit. Im allgemeinen Fall erzeugen die Eigenvektoren, die zu den positiven Eigenwerten gehören, einen Untervektorraum V_+, sodass

$$\beta(\boldsymbol{v}, \boldsymbol{v}) > 0 \quad \text{für alle } \boldsymbol{v} \in V_+ \setminus \{\boldsymbol{0}\},$$

und die Eigenvektoren, die zu den negativen Eigenwerten gehören, erzeugen einen Untervektorraum V_-, so dass

$$\beta(\boldsymbol{v}, \boldsymbol{v}) < 0 \quad \text{für alle } \boldsymbol{v} \in V_- \setminus \{\boldsymbol{0}\}.$$

Diese beiden Untervektorräume sind maximal mit dieser Eigenschaft, und das Tupel $(\dim(V_+), \dim(V_-))$ heißt **Signatur** der symmetrischen Bilinearform β.

Nehmen wir nun an, dass die Orthonormalbasis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ von $\mathbb{R}^n$ und die zugehörigen Eigenwerte $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ von A so gewählt sind, dass $\lambda_1, \ldots \lambda_t > 0$, $\lambda_{t+1}, \ldots, \lambda_{t+s} < 0$ und $\lambda_{t+s+1} = \cdots = \lambda_n = 0$, und ersetzen wir die Orthonormalbasis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ durch die (nicht mehr normierte) Basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ mit

$$\boldsymbol{w}_k = \begin{cases} \frac{1}{\sqrt{\lambda_k}} \cdot \boldsymbol{v}_k & \text{für } 1 \le k \le t, \\ \frac{1}{\sqrt{-\lambda_k}} \cdot \boldsymbol{v}_k & \text{für } t \le k \le t+s, \\ \boldsymbol{v}_k & \text{für } t+s+1 \le k \le n, \end{cases}$$

so hat für einen Vektor $\boldsymbol{v}$ mit der Darstellung

$$\boldsymbol{v} = r_1 \cdot \boldsymbol{w}_1 + \cdots + r_n \cdot \boldsymbol{w}_n$$

die Bilinearform β die besonders einfache Beschreibung

$$\beta(\boldsymbol{v}, \boldsymbol{v}) = r_1^2 + \cdots + r_t^2 - r_{t+1}^2 - \cdots - r_{t+s}^2 ,$$

und es gilt der **Trägheitssatz von Sylvester**:

$$t = \dim(V_+), \quad s = \dim(V_-) .$$

Teil III

14.4 Mathematischer Hintergrund: Definite und semidefinite symmetrische Matrizen

Unter den symmetrischen Bilinearformen bzw. symmetrischen Matrizen bilden die (positiv oder negativ) definiten eine besondere Klasse. In Mathematischer Hintergrund 14.3 haben wir schon die besondere Bedeutung der Diagonalform einer symmetrischen Matrix A für die durch diese Matrix definierte symmetrische Bilinearform β gesehen. Speziell die Vorzeichen der Eigenwerte und die Untervektorräume V_+ und V_-, die von den Eigenvektoren zu positiven und zu negativen Eigenwerten erzeugt werden, spielen eine wichtige Rolle. Ist etwa A positiv semidefinit, so bedeutet das, dass es keinen Vektor $\boldsymbol{v}$ gibt mit $\beta(\boldsymbol{v}, \boldsymbol{v}) < 0$, also dass $V_- = \{\mathbf{0}\}$ der Nullraum ist.

Aus dieser Überlegung erhalten wir: Genau dann ist die symmetrische $n \times n$-Matrix A positiv semidefinit, wenn eine (und damit alle) der folgenden Eigenschaften erfüllt ist:

1. Die Matrix A hat keine negativen Eigenwerte.
2. Die zu A gehörige Bilinearform β hat Signatur $(k, 0)$, wobei $0 \le k \le n$.
3. Es ist $V_- = \{\mathbf{0}\}$.

Entsprechend ist die symmetrische $n \times n$-Matrix A genau dann positiv definit, wenn eine der folgenden Eigenschaften erfüllt ist:

1. Die Matrix A hat nur positive Eigenwerte.
2. Die zu A gehörige Bilinearform β hat Signatur $(n, 0)$.
3. Es ist $V_+ = \mathbb{R}^n$.

Die Matrix $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ etwa ist positiv definit, da sie die beiden positiven Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = 2$ hat.

Da A genau dann negativ (semi-)definit ist, wenn $-A$ positiv (semi-)definit ist, ergeben sich daraus sofort charakterisierende Eigenschaften für diesen Fall. So ist etwa eine symmetrische $n \times n$-Matrix A genau dann negativ definit, wenn eine der folgenden Eigenschaften erfüllt ist:

1. Die Matrix A hat nur negative Eigenwerte.
2. Die zu A gehörige Bilinearform β hat Signatur $(0, n)$.
3. Es ist $V_- = \mathbb{R}^n$.

Die Indefinitheit einer symmetrischen Matrix A schließlich lässt sich dadurch beschreiben, dass A sowohl positive als auch negative Eigenwerte hat.

Die Matrix $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$ etwa ist indefinit, da sie die beiden Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = -2$ hat.

Diese Kriterien zur Bestimmung der Definitheit eignen sich jedoch nicht für alle Anwendungen, denn die Emittlung der Eigenwerte ist für große Matrizen ein sehr komplexes Problem. Daher werden einfachere Entscheidungskriterien benötigt.

Ein solches liefert das **Kriterium von Hurwitz**, das wir nun beschreiben wollen: Zu einer symmetrischen $n \times n$-Matrix A betrachten wir

$$A_k = \left(a_{i,j}\right)_{\substack{1 \le i \le k \\ 1 \le j \le k}} \qquad (k = 1, \ldots, n),$$

also die Matrizen, die wir aus A durch Streichung der letzten $n-k$ Zeilen und Spalten erhalten (sodass etwa $A_1 = (a_{1,1})$ und $A_n = A$). Wir setzen

$$\Delta_k = \det A_k \qquad (k = 1, \ldots, n)$$

und nennen die Δ_k die **Hauptminoren** von A. Dann besagt das Kriterium von Hurwitz, dass die Matrix A genau dann positiv definit ist, wenn $\Delta_k > 0$ für alle $k = 1, \ldots, n$ ist. Da A genau dann negativ definit ist, wenn $-A$ positiv definit ist, folgt hieraus unmittelbar, dass A genau dann negativ definit ist, wenn der erste Hauptminor Δ_1 negativ ist und die Hauptminoren immer wechselndes Vorzeichen haben, also wenn $(-1)^k \cdot \Delta_k > 0$.

Ist A positiv definit, so hat A nur positive Eigenwerte $\lambda_1, \ldots, \lambda_n$. Ist daher O eine orthogonale Matrix, die A auf Diagonalform D bringt, $D = O^\top \cdot A \cdot O$, so gilt

$$\det A = \det\left(O \cdot D \cdot O^\top\right) = 1 \cdot \det D \cdot 1 = \lambda_1 \cdots \lambda_n > 0,$$

und damit ist auf jeden Fall $\Delta_n > 0$. Ist nun $k < n$ und ist $\boldsymbol{v} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$ beliebig, so verlängern wir $\boldsymbol{v}$ zu einem Vektor $\boldsymbol{v}' \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, in dem wir an den Stellen $k+1, \ldots, n$ eine 0 einfügen. Wir erhalten $\boldsymbol{v}'^\top \cdot A \cdot \boldsymbol{v}' > 0$, da A positiv definit ist. Nun rechnen wir aber sofort nach, dass

$$\boldsymbol{v}'^\top \cdot A \cdot \boldsymbol{v}' = \boldsymbol{v}^\top \cdot A_k \cdot \boldsymbol{v},$$

sodass $\boldsymbol{v}^\top \cdot A_k \cdot \boldsymbol{v} > 0$ für alle $\boldsymbol{v} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$. Damit ist auch A_k positiv definit, und obiges Argument zeigt, dass

$$\Delta_k = \det A_k > 0.$$

Ist also A positiv definit, so sind alle Hauptminoren positiv. Die umgekehrte Richtung erfordert ein komplizierteres induktives Argument, auf das wir hier verzichten wollen.

Die Semidefinitheit und die Indefinitheit einer Matrix lassen sich leider nicht so allgemein durch die Vorzeichen der Hauptminoren beschreiben. Wir können jedoch noch sagen: Sind alle Hauptminoren einer symmetrischen Matrix A von 0 verschieden, und folgt das Vorzeichenmuster der Hauptminoren nicht dem Muster einer positiv definiten oder einer negativ definiten Matrix, so ist A indefinit.

Teil III

14.5 Mathematischer Hintergrund: Quadratische Formen und Quadriken

Eine **quadratische Form** q auf $\mathbb{R}^n$ ist eine Abbildung

$$q : \mathbb{R}^n \longrightarrow \mathbb{R}\,,$$

die in der Form

$$q(x_1, \dots, x_n) = \sum_{i,j=1, i\leq j}^{n} \alpha_{i,j} x_i \cdot x_j + \sum_{i=1}^{n} \varphi_i x_i + \delta$$

geschrieben werden kann, wobei mindestens ein $\alpha_{i,j} \neq 0$. Betrachten wir die symmetrische $n \times n$-Matrix $A = (a_{i,j})$ mit

$$a_{i,j} = \begin{cases} \alpha_{i,i}\,, & \text{falls } i = j\,, \\ \frac{\alpha_{i,j}}{2}\,, & \text{falls } i < j\,, \\ \frac{\alpha_{j,i}}{2}\,, & \text{falls } i > j\,, \end{cases}$$

mit zugehöriger symmetrischer Bilinearform

$$\beta : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}\,,$$

gegeben durch $\beta(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^\top \cdot A \cdot \boldsymbol{y}$, und die Linearform

$$\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$$

mit $\varphi(\boldsymbol{x}) = \sum_{i=1}^{n} \varphi_i \cdot x_i$, so gilt

$$q(x_1, \dots, x_n) = \beta(\boldsymbol{x}, \boldsymbol{x}) + \varphi(\boldsymbol{x}) + \delta\,.$$

Wie wir in Mathematischer Hintergrund 14.4 gesehen haben, gibt es eine orthogonale Matrix O, die β diagonalisiert. Es gibt also eine orthogonale Koordinatentransformation des $\mathbb{R}^n$, sodass die quadratische Form q in den neuen Koordinaten die Form

$$q(y_1, \dots, y_n) = \sum_{i=1}^{n} \lambda_i \cdot y_i^2 + \sum_{i=1}^{n} \psi_i \cdot y_i + \delta$$

hat. Dabei können wir annehmen, dass $\lambda_1, \dots, \lambda_k > 0$, $\lambda_{k+1}, \cdots, \lambda_t < 0$ und $\lambda_{t+1} = \cdots = \lambda_n = 0$ (für geeignete $0 \leq k \leq t \leq n$). Da A nicht die Nullmatrix ist (mindestens ein $\alpha_{i,j} \neq 0$), gilt auf jeden Fall $t \geq 1$. Damit kann die quadratische Form aber auch wie folgt geschrieben werden:

$$q(y_1, \dots, y_n) = \sum_{i=1}^{k} \frac{1}{a_i^2} \cdot y_i^2 - \sum_{i=k+1}^{t} \frac{1}{a_i^2} \cdot y_i^2 + \sum_{i=1}^{n} \psi_i \cdot y_i + \delta$$

mit geeigneten positiven reellen Zahlen a_i. Für alle $i \leq t$ führen wir quadratische Ergänzung durch und erhalten

$$\frac{\pm 1}{a_i^2} \cdot y_i^2 + \psi_i y_i = \frac{\pm 1}{a_i^2} \cdot (y_i - \eta_i)^2 + \rho_i\,.$$

Eine Verschiebung der i-ten Koordinate um η_i (also ein Koordinatenwechsel der Form $z_i = y_i - \eta_i$ ($i = 1, \dots, t$)) bringt die quadratische Form dann auf die Gestalt

$$q(z_1, \dots, z_n) = \sum_{i=1}^{k} \frac{1}{a_i^2} \cdot z_i^2 - \sum_{i=k+1}^{t} \frac{1}{a_i^2} \cdot z_i^2 + \rho\,,$$

falls $\psi_i = 0$ für alle $i > t$, und

$$q(z_1, \dots, z_n) = \sum_{i=1}^{k} \frac{1}{a_i^2} \cdot z_i^2 - \sum_{i=k+1}^{t} \frac{1}{a_i^2} \cdot z_i^2 - 2 \cdot z_n$$

falls es ein $i > t$ gibt mit $\psi_i \neq 0$ (in diesem Fall führen wir noch eine weitere Verschiebung und eine orthogonale Koordinatentransformation durch, die $\sum_{i=t+1}^{n} \psi_i y_i + \sum \rho_i + \delta$ in $-2 \cdot z_n$ überführt). Diese Gestalt nennen wir eine **Normalform der quadratischen Form**.

Eine **Quadrik** Q ist die Nullstellenmenge einer quadratischen Form q:

$$Q = \{(x_1, \dots, x_n) \in \mathbb{R}^n \,|\, q(x_1, \dots, x_n) = 0\}$$

Zur Untersuchung von Quadriken können wir nach Wahl eines geeigneten Koordinatensystems immer annehmen, dass die quadratische Form in Normalform vorliegt. Außerdem können wir die Quadrikengleichung immer durch ein beliebige von 0 verschiedene Zahl dividieren, ohne dadurch ihre Nullstellenmenge zu ändern. Daher erhalten wir aus den Normalformen drei Typen von Quadriken:

1. **Quadriken vom kegeligen Typ:** Diese sind in geeigneten Koordinaten gegeben durch eine Gleichung der Form

$$\sum_{i=1}^{k} \frac{1}{a_i^2} \cdot z_i^2 - \sum_{i=k+1}^{t} \frac{1}{a_i^2} \cdot z_i^2 = 0\,,$$

   wobei wir auch noch annehmen können, dass $k \geq t - k$.
2. **Mittelpunktsquadriken:** Hier lautet die Gleichung in geeigneten Koordinaten

$$\sum_{i=1}^{k} \frac{1}{a_i^2} \cdot z_i^2 - \sum_{i=k+1}^{t} \frac{1}{a_i^2} \cdot z_i^2 = 1\,.$$

3. **Quadriken vom parabolischen Typ:** Diese sind in geeigneten Koordinaten gegeben durch eine Gleichung

$$\sum_{i=1}^{k} \frac{1}{a_i^2} \cdot z_i^2 - \sum_{i=k+1}^{t} \frac{1}{a_i^2} \cdot z_i^2 - 2z_n = 0\,,$$

   wobei $t < n$.

Beachten Sie dabei, dass wir bei allen Koordinatentransformationen nur Verschiebungen und orthogonale Transformationen durchführen müssen, also nur Koordinatenwechsel, die Längen und Winkel erhalten.

Anwendung: Kegelschnitte

Im Spezialfall $n = 2$ nennen wir Quadriken **Kegelschnitte**. Mithilfe der Normalformen aus Mathematischer Hintergrund 14.5 lassen sich diese sehr gut klassifizieren.

Kreise

Ein Kreis K ist ein Spezialfall einer Mittelpunktsquadrik, gegeben durch eine Gleichung der Form

$$\frac{1}{r^2}\cdot y^2 + \frac{1}{r^2}\cdot x^2 = 1\,,$$

oder äquivalent, $y^2 + x^2 = r^2$. Er wird beschrieben durch Abb. 14.2.

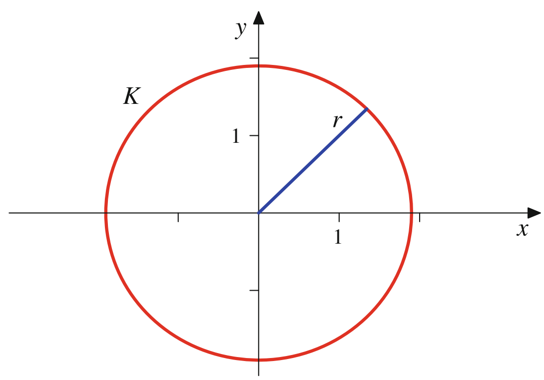


Abb. 14.2 Kegelschnitte: Ein Kreis

Das ist tatsächlich ein Kreis in der Ebene mit Radius r und Mittelpunkt $M = (0,\ 0)$. Lassen wir die Verschiebung $z_i = y_i - \eta_i$ der Koordinaten weg, die wir durchgeführt haben, um die Quadriken auf Normalform zu bringen, so lautet die Kreisgleichung

$$(y - y_0)^2 + (x - x_0)^2 = r^2\,,$$

und wir betrachten einen Kreis mit Radius r und Mittelpunkt $M = (x_0, y_0)$. Diese Form bleibt auch erhalten, wenn wir die orthogonale Koordinatentransformation rückgängig machen, denn beim Kreis ist die Matrix A immer von der Form $A = \frac{1}{r^2}\cdot \mathrm{E}_2$, und diese Gestalt bleibt bei jeder Transformation mit einer (orthogonalen) Matrix erhalten.

Ellipsen

Eine Ellipse E ist eine Mittelpunktsquadrik, gegeben durch eine Gleichung der Form

$$\frac{1}{a^2}\cdot y^2 + \frac{1}{b^2}\cdot x^2 = 1$$

mit $a, b > 0$ und $a \neq b$. Das Bild hiervon ist eine Ellipse mit Mittelpunkt $M = (0,\ 0)$ und zwei Hauptachsen der Länge a und b. Die Ellipse in Abb. 14.3 wird gegeben durch die Gleichung

$$\frac{1}{1.8^2}\cdot y^2 + \frac{1}{3^2}\cdot x^2 = 1\,.$$

Es handelt sich also um eine Ellipse mit langer Halbachse $b = 3$ und kurzer Halbachse $a = 1.8$.

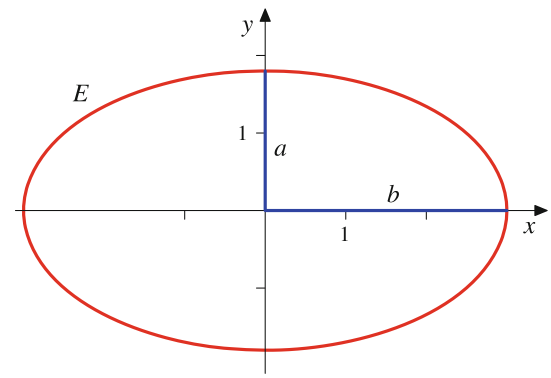


Abb. 14.3 Kegelschnitte: Eine Ellipse

Lassen wir die Verschiebung $z_i = y_i - \eta_i$ der Koordinaten weg, die wir durchgeführt haben, um die Quadriken auf Normalform zu bringen, so lautet die Ellipsengleichung

$$\frac{1}{a^2}\cdot (y - y_0)^2 + \frac{1}{b^2}\cdot (x - x_0)^2 = 1\,,$$

und wir betrachten hier eine Ellipse mit Mittelpunkt $M = (x_0, y_0)$ und Hauptachsen die parallel zu den Koordinatenachsen sind. Lassen wir auch die orthogonale Transformation weg, so sind die Hauptachsen nicht mehr parallel zu den Koordinatenachsen, und deshalb treten in diesem Fall gemischte Terme xy auf, selbst dann, wenn der Mittelpunkt der Ellipse im Koordinatenursprung liegt.

Hyperbeln

Eine Hyperbel H ist ein weiterer Spezialfall einer Mittelpunktsquadrik, gegeben durch eine Gleichung

$$\frac{1}{a^2}\cdot y^2 - \frac{1}{b^2}\cdot x^2 = 1$$

mit $a, b > 0$.

Eine Hyperbel in Normalform hat zwei Ursprungsgeraden als Asymptoten, deren Steigung $\frac{a}{b}$ bzw. $-\frac{a}{b}$ ist. Der Tiefpunkt des oberen Hyperbelastes liegt im Punkt $(0,\ a)$ und der Hochpunkt des unteren Hyperbelastes im Punkt $(0, -a)$. Der Schnittpunkt der beiden Asymptoten, also in Normalform der Punkt $M = (0,\ 0)$, heißt Mittelpunkt der Hyperbel.

Teil III

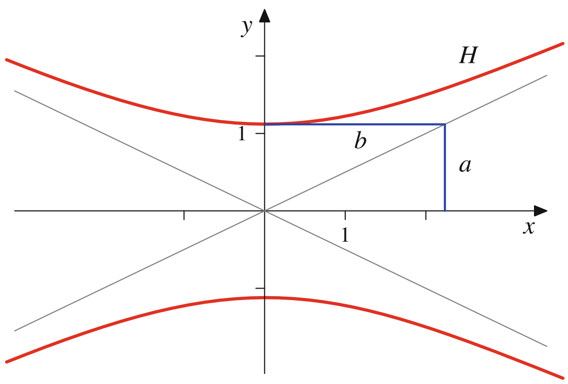


Abb. 14.4 Kegelschnitte: Eine Hyperbel

Die Normalengleichung der Hyperbel aus Abb. 14.4 ist

$$\frac{1}{\left(\frac{\sqrt{5}}{2}\right)^2} \cdot y^2 - \frac{1}{\sqrt{5}^2} \cdot x^2 = 1\,.$$

Wir haben also $a = \frac{\sqrt{5}}{2}$ und $b = \sqrt{5}$.

Bekannt ist die Hyperbel, die explizit durch $y = \frac{1}{x}$ beschrieben wird, also durch die Gleichung $x \cdot y = 1$. Das ist keine Hyperbelgleichung in Normalform. Schreiben wir diese Gleichung jedoch in den durch die Orthonormalbasis

$$\boldsymbol{v}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

gegebenen Koordinaten (drehen wir also die Koordinatenachse um 45° im Uhrzeigersinn), so wird diese Hyperbelgleichung zu

$$2y^2 - 2x^2 = 1\,,$$

liegt also in Normalform mit $a = \frac{\sqrt{2}}{2}$ und $b = \frac{\sqrt{2}}{2}$ vor.

Parabeln

Eine Parabel P ist eine Quadrik vom parabolischen Typ,

$$\frac{1}{a^2} \cdot x^2 - 2y = 0\,.$$

Ihr Bild ist eine Parabel, wie wir sie aus der Analysis kennen. Die Normalengleichung der Parabel in Abb. 14.5 ist $\frac{1}{2^2} \cdot x^2 - 2y = 0$, wir haben also $a = 2$:

Alle Parabeln in Normalform haben ihren **Scheitelpunkt** in $S = (0, 0)$. Der Punkt $B = \left(0, \frac{a^2}{2}\right)$ heißt **Brennpunkt** der Parabel. Senkrecht von oben einfallende Strahlen werden an der Parabel so gespiegelt, dass sie sich alle im Punkt B treffen. Diese Eigenschaft wird ausgenutzt bei der Konstruktion parabolischer Spiegel.

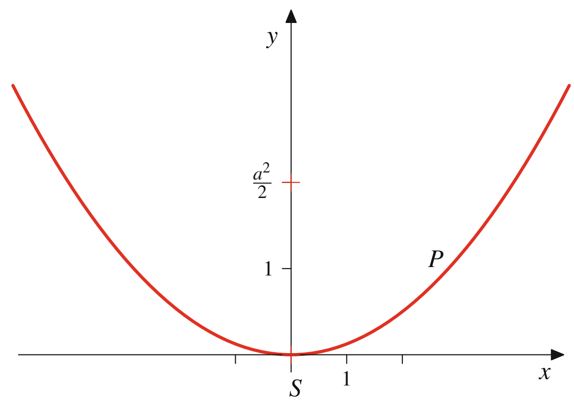


Abb. 14.5 Kegelschnitte: Eine Parabel

Geradenkreuzungen

Eine Geradenkreuzung G ist eine Quadrik vom kegeligen Typ mit der Gleichung

$$\frac{1}{a^2} \cdot y^2 - \frac{1}{b^2} \cdot x^2 = 0\,,$$

die sich nach der dritten binomischen Formel als

$$\left(\frac{1}{a} \cdot y - \frac{1}{b} \cdot x\right) \cdot \left(\frac{1}{a} \cdot y + \frac{1}{b} \cdot x\right) = 0$$

schreiben lässt. Damit ist G die Vereinigung der beiden Geraden $\frac{1}{a} \cdot y - \frac{1}{b} \cdot x = 0$ und $\frac{1}{a} \cdot y + \frac{1}{b} \cdot x = 0$. Es handelt sich um zwei Ursprungsgeraden mit Steigung $\frac{a}{b}$ und $-\frac{a}{b}$. Als Spezialfall einer Geradenkreuzung kann auch eine Doppelgerade, in der Normalform gegeben durch

$$\frac{1}{a^2} \cdot x^2 = 0\,,$$

auftreten. Die Gleichung der Geradenkreuzung in Abb. 14.6 ist $y^2 - \frac{1}{2^2} \cdot x^2 = 0$, wir haben also $a = 1$ und $b = 2$.

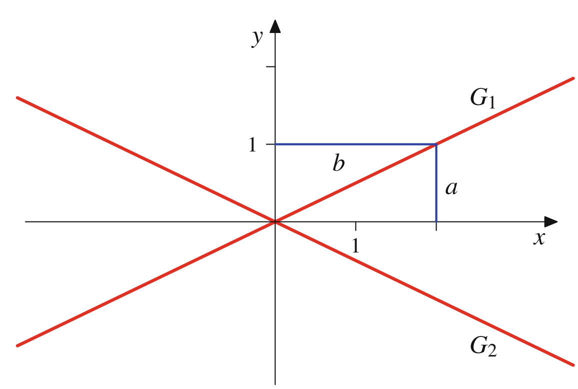


Abb. 14.6 Kegelschnitte: Ein Geradenkreuz

Teil III

Die Quadrik, die wir hier erhalten, besteht also gerade aus den beiden Asymptoten, die wir für die Hyperbel

$$\frac{1}{\left(\frac{\sqrt{5}}{2}\right)^2}\cdot y^2-\frac{1}{\sqrt{5}^2}\cdot x^2=1$$

erhalten haben, da in beiden Fällen $\frac{a}{b}=\frac{1}{2}$.

Nicht mehr zu den Kegelschnitten gezählt werden üblicherweise die Typen

$$\frac{1}{a^2}\cdot y^2+\frac{1}{b^2}\cdot x^2=0 \quad \text{bzw.} \; -\frac{1}{a^2}\cdot y^2-\frac{1}{b^2}\cdot x^2=1$$

(bei denen die Lösungsmenge nur aus dem Punkt $(0,0)$ besteht oder leer ist). Manchmal werden schon Geradenkreuzungen nicht mehr zu den Kegelschnitten gezählt.

Definition

Eine $n\times n$-Matrix A heißt **trigonalisierbar**, wenn es eine invertierbare $n\times n$-Matrix S gibt, sodass

$$T=S^{-1}\cdot A\cdot S$$

eine obere Dreiecksmatrix ist.

Eine obere Dreiecksform kann immer erreicht werden.

Komplexe Matrizen sind trigonalisierbar

Jede $n\times n$-Matrix ist (über den komplexen Zahlen) trigonalisierbar.

Das Vorgehen wird anhand eines Beispiels erläutert.

Wir betrachten die Matrix

$$A=\begin{pmatrix}-7&0&2\\16&3&-3\\-34&-1&9\end{pmatrix}. \tag{14.10}$$

Hierfür gilt

$$P_A(\lambda)=\lambda^3-5\lambda^2+8\lambda-4=(\lambda-1)\cdot(\lambda-2)^2,$$

also hat A die Eigenwerte $\lambda_1=1$ und $\lambda_2=2$.

Zu $\lambda=1$ finden wir als Eigenvektor den Vektor

$$\boldsymbol{v}_1=\begin{pmatrix}1\\-2\\4\end{pmatrix}.$$

Wir ergänzen $\boldsymbol{v}_1$ durch die beiden Standardbasisvektoren $\mathbf{e}_2$ und $\mathbf{e}_3$ zu einer Basis von $\mathbb{R}^3$ und bilden mit diesen Vektoren als Spalten die Umrechnungsmatrix

$$S_1=\begin{pmatrix}1&0&0\\-2&1&0\\4&0&1\end{pmatrix}.$$

Dann ist

$$A_1=S_1^{-1}\cdot A\cdot S_1=\begin{pmatrix}1&0&2\\0&3&1\\0&-1&1\end{pmatrix}.$$

Die Matrix A_1 hat also bereits die erste Spalte einer oberen Dreiecksmatrix. Nach Streichen der ersten Spalte und der ersten Zeile von A_1 erhalten wir die Matrix

$$B=\begin{pmatrix}3&1\\-1&1\end{pmatrix}$$

mit charakteristischem Polynom

$$P_B(\lambda)=\lambda^2-4\lambda+4=(\lambda-2)^2.$$

Also hat B den (einzigen) Eigenwert $\lambda_2=2$, und hierzu erhalten wir als Eigenvektor

$$\boldsymbol{u}_2=\begin{pmatrix}-1\\1\end{pmatrix}.$$

Beachten Sie, dass B keine zwei linear unabhängigen Eigenvektoren besitzt (also nicht diagonalisierbar ist). Wir ergänzen $\boldsymbol{u}_2$ durch den Standardbasisvektor $\mathbf{e}_2$ zu einer Basis von $\mathbb{R}^2$, bilden mit diesen Vektoren die Umrechnungsmatrix

$$S_2=\begin{pmatrix}-1&0\\1&1\end{pmatrix}$$

und erhalten daraus

$$B_1=S_2^{-1}\cdot B\cdot S_2=\begin{pmatrix}2&-1\\0&2\end{pmatrix},$$

also eine obere Dreiecksmatrix. Ergänzen wir S_2 wie folgt durch eine erste Zeile und erste Spalte zu einer invertierbaren 3×3-Matrix,

$$S_3=\begin{pmatrix}1&0&0\\0&-1&0\\0&1&1\end{pmatrix},$$

so erhalten wir hierfür

$$T=S_3^{-1}\cdot A_1\cdot S_3=\begin{pmatrix}1&2&2\\0&2&-1\\0&0&2\end{pmatrix},$$

also eine obere Dreiecksmatrix. Setzen wir $S = S_1 \cdot S_3$, so ist

$$S^{-1} \cdot A \cdot S = S_3^{-1} \cdot S_1^{-1} \cdot A \cdot S_1 \cdot S_3 = S_3^{-1} \cdot A_1 \cdot S_3 = T$$

eine obere Dreiecksmatrix.

Dieses Vorgehen funktioniert bei allen quadratischen Matrizen. Es wird immer erst ein Eigenwert und zu diesem ein Eigenvektor betrachtet. Mit diesem Eigenvektor (und beliebigen $n-1$ Vektoren, die ihn zu einer Basis von $\mathbb{C}^n$ ergänzen) wird eine erste Transformationsmatrix S_1 gebaut und eine erste Transformation

$$A_1 = S_1^{-1} \cdot A \cdot S_1$$

durchgeführt, die immer dazu führt, dass A_1 in der ersten Spalte unterhalb der ersten Zeile nur noch 0 als Einträge hat. Dann werden die erste Zeile und die erste Spalte von A_1 gestrichen, und für die verbleibende $(n-1)\times(n-1)$-Matrix B werden wieder ein Eigenwert und ein zugehöriger Eigenvektor gesucht, und man fährt fort wie mit der Matrix A selbst.

Entlang der Diagonalen treten dabei immer die Eigenwerte von A auf, und zwar jeder Eigenwert genauso oft, wie er als Nullstelle des charakteristischen Polynoms von A auftritt. Sind A und alle ihre Eigenwerte reell, so kann komplett im Reellen gerechnet werden, d. h., in diesem Fall kann auch eine reelle Transformationsmatrix S gefunden werden. Falls aber komplexe Eigenwerte auftreten, so muss auch die Transformationsmatrix im Komplexen konstruiert werden.

Die jordansche Normalform ist eine ausgezeichnete obere Dreiecksform

Bei der oberen Dreiecksmatrix, die wir bei der Trigonalisierung erhalten, ist zu beachten, dass sie nicht eindeutig ist. Betrachten wir für die Matrix (14.10) die Umrechnungsmatrix

$$\widetilde{S} = \begin{pmatrix} 1 & 18 & -4 \\ -2 & -45 & -8 \\ 4 & 81 & -9 \end{pmatrix},$$

so erhalten wir hierfür

$$\widetilde{S}^{-1} \cdot A \cdot \widetilde{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

also wieder eine obere Dreiecksmatrix mit den Eigenwerten $\lambda_1 = 1$ und $\lambda_2 = 2$ entlang der Diagonalen, allerdings mit einer komplett anderen Gestalt oberhalb der Diagonalen.

Unter allen oberen Dreiecksformen, die sich als Transformationen einer gegebenen Matrix finden lassen, gibt es jedoch eine ausgezeichnete.

Jordansche Normalform

Ist A eine $n \times n$-Matrix (über $\mathbb{R}$ oder $\mathbb{C}$), so gibt es eine Transformationsmatrix S, für die die transformierte Matrix $B = S^{-1} \cdot A \cdot S$ die folgende Gestalt hat:

$$J = \begin{pmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ 0 & 0 & J_3 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & J_k \end{pmatrix}$$

mit quadratischen $n_l \times n_l$-Matrizen J_l der Form

$$J_l = \begin{pmatrix} \lambda_l & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_l & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_l & \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_l & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_l \end{pmatrix}$$

Dabei gilt

$$n_1 + n_2 + \dots + n_k = n$$

und die Matrizen J_l sind eindeutig bis auf ihre Reihenfolge entlang der Diagonale.

Die transformierte Matrix $J = S^{-1} \cdot A \cdot S$ heißt **jordansche Normalform** der Matrix A, und die $n_l \times n_l$-Matrizen J_l heißen **Jordan-Kästchen** von A.

Für die Matrix (14.10) ist

$$S := \begin{pmatrix} 1 & 18 & -4 \\ -2 & -45 & -8 \\ 4 & 81 & -9 \end{pmatrix}$$

eine Transformationsmatrix, die A auf die jordansche Normalform

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

bringt. Die Jordan-Kästchen dieser Normalform sind

$$J_1 = \begin{pmatrix} 1 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

Da es hier nur ein Jordan-Kästchen zum (kritischen) Eigenwert 2 gibt, kann die Transformationsmatrix relativ leicht bestimmt werden:

- Die erste Spalte $\boldsymbol{v}_1$ von S ist ein Eigenvektor von A zum Eigenwert 1.

- Die zweite Spalte $\boldsymbol{v}_2$ von S ist ein Eigenvektor von A zum Eigenwert 2.
- Die dritte Spalte $\boldsymbol{v}_3$ von S ist eine Lösung des Gleichungssystems $(A - 2 \cdot \mathrm{E}_3) \cdot \boldsymbol{x} = \boldsymbol{v}_2$, sodass für $\boldsymbol{v}_3$

$$A \cdot \boldsymbol{v}_3 = 2 \cdot \boldsymbol{v}_3 + \boldsymbol{v}_2$$

gilt. Einen solchen Vektor nennen wir einen **Hauptvektor** von A zum Eigenwert λ (der Stufe 2).

Diese Technik funktioniert immer, wenn es zu jedem kritischen Eigenwert nur ein Jordan-Kästchen gibt. (Hätte das Kästchen mehr als zwei Zeilen und Spalten, so würden wir mit einer Lösung $\boldsymbol{v}_4$ des Gleichungssystems $(A - 2 \cdot \mathrm{E}_3) \cdot \boldsymbol{x} = \boldsymbol{v}_3$ fortfahren, usw.).

Die Normalform ist eindeutig bis auf die Anordnung dieser Kästchen, d. h., die einzige andere jordansche Normalform, die noch möglich ist, wäre

$$\widetilde{J} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

die etwa mit der Transformationsmatrix

$$\widetilde{S} = \begin{pmatrix} 18 & -4 & 1 \\ -45 & -8 & -2 \\ 81 & -9 & 4 \end{pmatrix}$$

erreicht werden kann.

Beispiel

Die Matrix

$$A = \begin{pmatrix} 3 & 7 & -13 \\ -2 & -9 & 20 \\ -1 & -4 & 9 \end{pmatrix} \tag{14.11}$$

hat das charakteristische Polynom

$$P_A(\lambda) = \lambda^3 - 3\lambda^2 + 4 = (\lambda + 1) \cdot (\lambda - 2)^2$$

und damit die Eigenwerte $\lambda_1 = -1$ und $\lambda_2 = 2$. Alle Eigenvektoren zum doppelten Eigenwert λ_2 sind Vielfache des Vektors $\boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$, sodass die Matrix A nicht diagonalisierbar ist. Eine Lösung des Gleichungssystems $(A - 2 \cdot \mathrm{E}_3) \cdot \boldsymbol{x} = \boldsymbol{v}_2$ ist der Vektor $\boldsymbol{v}_3 = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$. Daher wird A durch die Transformationsmatrix

$$S := \begin{pmatrix} -2 & -1 & -2 \\ 3 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

auf jordansche Normalform

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

gebracht. Die Jordan-Kästchen sind

$$J_1 = \begin{pmatrix} -1 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$ ◀

Achtung Die Anzahl der Jordan-Kästchen von A zu einem Eigenwert λ_0 ist bestimmt durch die geometrische Vielfachheit von λ_0, also durch $g = \mathrm{Nul}\,(\lambda_0 \cdot \mathrm{E}_n - A)$. Hat λ_0 die algebraische Vielfachheit a, tritt der Linearfaktor $\lambda - \lambda_0$ in der Linearfaktorzerlegung des charakteristischen Polynoms also a-mal auf, so gibt es genau dann Jordan-Kästchen mit mehr als einer Zeile und einer Spalte, wenn $g < a$.

Falls $g \geq 2$ und $g < a$, so kann eine Transformationsmatrix jedoch nicht mehr dadurch bestimmt werden, dass zu einem beliebigen Eigenvektor $\boldsymbol{v}$ von A zu λ_0 das Gleichungssystem $(A - \lambda_0 \cdot \mathrm{E}_n) \cdot \boldsymbol{x} = \boldsymbol{v}$ gelöst wird. In diesem Fall müssen erst die Hauptvektoren (höchster Stufe) gefunden werden. ◀

Die Diagonalisierbarkeit von Matrizen ist in der Theorie der jordanschen Normalformen schon mit eingeschlossen.

Jordansche Normalform und Diagonalisierbarkeit

Genau dann ist eine $n \times n$-Matrix A diagonalisierbar, wenn alle Jordan-Kästchen die Größe 1×1 haben, wenn also die jordansche Normalform von A schon eine Diagonalmatrix ist.

In der Praxis ist die jordansche Normalform schwer zu bestimmen, speziell bei Matrizen, die viele Jordan-Kästchen zu einem Eigenwert haben. Außerdem ist die jordansche Normalform numerisch instabil. Bereits eine minimale Variation der Matrizeneinträge kann zu einer ganz anderen Gestalt der Normalform führen. Sie ist allerdings wichtig für die korrekte Behandlung von Differenzialgleichungssystemen.

Jede Matrix kann in eine obere Dreiecksmatrix und eine orthogonale Matrix zerlegt werden

Aus numerischer Sicht ist es günstiger, mit orthogonalen Matrizen zu arbeiten (wie wir schon beim Orthonormalisierungsverfahren nach Gram-Schmidt diskutiert haben). Eine Transformation mit orthogonalen Matrizen führt jedoch selten zu einer befriedigenden Normalform. Allerdings können damit für die

14.6 Mathematischer Hintergrund: Die Matrixexponentialfunktion

Die Exponentialfunktion $f(x) = \mathrm{e}^x$ spielt eine zentrale Rolle in der Analysis, speziell im Zusammenhang mit der Differenzialrechnung, da sie sich durch die Eigenschaft $f'(x) = f(x)$ auszeichnet. Bei der Untersuchung linearer Differenzialgleichungssysteme ist es nützlich, diese Konstruktion auch für Matrizen zu haben. Dabei gehen wir von der Exponentialreihe aus, also von der Beschreibung

$$\mathrm{e}^a = 1 + \frac{a}{1!} + \frac{a^2}{2!} + \cdots = \sum_{k=0}^{\infty} \frac{a^k}{k!}\,.$$

Dieser Ansatz kann auf quadratische Matrizen übertragen werden und führt für eine $n \times n$-Matrix A zu der Definition

$$\mathrm{e}^A := \mathrm{E}_n + \frac{A}{1!} + \frac{A^2}{2!} + \cdots = \sum_{k=0}^{\infty} \frac{A^k}{k!}\,,$$

wobei A^k hier die k-te Potenz der Matrix A bezeichnet, oder noch etwas allgemeiner zu

$$\mathrm{e}^{t\cdot A} := \sum_{k=0}^{\infty} \frac{(t\cdot A)^k}{k!}\,.$$

Diese Reihe konvergiert. Wir notieren hier einige grundlegende Eigenschaften dieser Konstruktion:

1. Ist $A = 0$, so ist $\mathrm{e}^0 = \mathrm{E}_n$ die Einheitsmatrix.
2. Sind A und B zwei Matrizen mit $A \cdot B = B \cdot A$, so ist $\mathrm{e}^{A+B} = \mathrm{e}^A \cdot \mathrm{e}^B$.
3. Für jede Matrix A ist e^A invertierbar mit $\left(\mathrm{e}^A\right)^{-1} = \mathrm{e}^{-A}$.
4. Ist $A = S \cdot B \cdot S^{-1}$ die Transformation einer Matrix B mit einer invertierbaren Matrix S, so ist $\mathrm{e}^{t\cdot A} = S \cdot \mathrm{e}^{t\cdot B} \cdot S^{-1}$.

Die ersten drei Aussagen beweist man dabei wie im eindimensionalen Fall. Was die vierte Eigenschaft betrifft, so ist zunächst klar, dass $t \cdot A = S \cdot t \cdot B \cdot S^{-1}$. Nun haben wir schon gesehen, dass $(t \cdot A)^k = S \cdot (t \cdot B)^k \cdot S^{-1}$, und hieraus folgt leicht, dass

$$\begin{aligned}
\mathrm{e}^{t\cdot A} &= \sum_{k=0}^{\infty} \frac{(t\cdot A)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{S\cdot (t\cdot B)^k \cdot S^{-1}}{k!} \\
&= S \cdot \sum_{k=0}^{\infty} \frac{(t\cdot B)^k}{k!} \cdot S^{-1} \\
&= S \cdot \mathrm{e}^{t\cdot B} \cdot S^{-1}\,.
\end{aligned}$$

Für eine beliebige Matrix A ist $\mathrm{e}^{t\cdot A}$ direkt nur sehr schwer zu ermitteln; diese vierte Eigenschaft ermöglicht jedoch eine explizite Bestimmung.

Ist dazu zunächst D eine Diagonalmatrix mit Diagonalelementen $d_1, \ldots, d_n$, so ist $(t \cdot D)^k$ ebenfalls eine Diagonalmatrix mit den Diagonalelementen $(td_1)^k, \ldots, (td_n)^k$. Daraus erhalten wir sofort, dass auch $\mathrm{e}^{t\cdot D}$ eine Diagonalmatrix ist, und zwar

$$\mathrm{e}^{t\cdot D} = \begin{pmatrix} \mathrm{e}^{td_1} & 0 & \ldots & 0 \\ 0 & \mathrm{e}^{td_2} & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & e^{td_n} \end{pmatrix}.$$

Ist A diagonalsisierbar und ist $D = S^{-1} \cdot A \cdot S$ eine Transformation von A auf Diagonalgestalt, so ist $A = S \cdot D \cdot S^{-1}$ und damit

$$\mathrm{e}^{t\cdot A} = S \cdot \mathrm{e}^{t\cdot D} \cdot S^{-1}\,.$$

Betrachten wir etwa die Matrix (14.9), also

$$A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 2 & 4 \\ 2 & 4 & 2 \end{pmatrix}.$$

so hat diese den doppelten Eigenwert $\lambda_1 = 6$ und den Eigenwert $\lambda_2 = -3$ und wird durch die orthogonale Matrix

$$S = \frac{1}{3} \cdot \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix}$$

diagonalisiert. Damit erhalten wir

$$\begin{aligned}
\mathrm{e}^{t\cdot A} &= S \cdot \begin{pmatrix} \mathrm{e}^{6t} & 0 & 0 \\ 0 & \mathrm{e}^{6t} & 0 \\ 0 & 0 & e^{-3t} \end{pmatrix} \cdot S^{\top} \\
&= \frac{1}{9} \cdot \begin{pmatrix} 8\mathrm{e}^{6t} - \mathrm{e}^{-3t} & -2\mathrm{e}^{6t} + 2\mathrm{e}^{-3t} & 2\mathrm{e}^{6t} - 2\mathrm{e}^{-3t} \\ -2\mathrm{e}^{6t} + 2\mathrm{e}^{-3t} & 5\mathrm{e}^{6t} + 4\mathrm{e}^{-3t} & 4\mathrm{e}^{6t} - 4\mathrm{e}^{-3t} \\ 2\mathrm{e}^{6t} - 2\mathrm{e}^{-3t} & 4\mathrm{e}^{6t} - 4\mathrm{e}^{-3t} & 5\mathrm{e}^{6t} + 4\mathrm{e}^{-3t} \end{pmatrix}.
\end{aligned}$$

Beachten Sie dabei, dass $\mathrm{e}^{t\cdot A}$ wieder symmetrisch ist.

Es bleibt der Fall einer nicht diagonalisierbaren Matrix A zu betrachten. In diesem Fall können wir zumindest noch eine jordansche Normalform erreichen:

$$S^{-1} \cdot A \cdot S = J = \begin{pmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & J_m \end{pmatrix}$$

Dann folgt aus den Regeln zur Matrizenmultiplikation

$$(t \cdot J)^k = \begin{pmatrix} (t\cdot J_1)^k & 0 & \ldots & 0 \\ 0 & (t\cdot J_2)^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & (t\cdot J_m)^k \end{pmatrix},$$

und wir können uns daher auf ein Jordan-Kästchen beschränken, also eine Matrix der Form

$$J = J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Diese Matrix schreiben wir nun als $J = D + N$, wobei

$$D = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

eine Diagonalmatrix ist und

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

eine Matrix, die in der Nebendiagonalen an allen Stellen den Wert 1 hat und sonst nur die 0 enthält. Hierfür gilt aber nun $D \cdot N = N \cdot D$, woraus $\mathrm{e}^{tJ} = \mathrm{e}^{tD} \cdot \mathrm{e}^{tN}$ folgt. Die Matrix N ist nilpotent, genauer gilt: Ist r die Größe von N, so ist $N^r = 0$ und damit

$$\begin{aligned} \mathrm{e}^{t\cdot N} &= \mathrm{E}_r + t \cdot N + \frac{(t\cdot N)^2}{2!} + \dots + \frac{(t \cdot N)^{r-1}}{(r-1)!} \\ &= \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{r-1}}{(r-1)!} \\ 0 & 1 & t & \dots & \frac{t^{r-2}}{(r-2)!} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \end{aligned}$$

Da $\mathrm{e}^{t\cdot D} = \mathrm{e}^{t\cdot\lambda} \cdot \mathrm{E}_r$, wie wir oben schon gesehen haben, ist zunächst $\mathrm{e}^{t\cdot J_\lambda} = \mathrm{e}^{t\cdot\lambda} \cdot \mathrm{e}^{t\cdot N}$, woraus wir kästchenweise die Matrixexponentialfunktion für J und daraus via $\mathrm{e}^{t\cdot A} = S \cdot \mathrm{e}^{t\cdot J} \cdot S^{-1}$ die Matrixexponentialfunktion ganz allgemein bekommen.

Betrachten wir hierzu die Matrix (14.11), also

$$A = \begin{pmatrix} 3 & 7 & -13 \\ -2 & -9 & 20 \\ -1 & -4 & 9 \end{pmatrix},$$

so haben wir schon die jordansche Normalform J und die Transformationsmatrix S bestimmt. Es ist daher

$$\mathrm{e}^{t\cdot J} = \begin{pmatrix} \mathrm{e}^{-t} & 0 & 0 \\ 0 & \mathrm{e}^{2t} & t\cdot \mathrm{e}^{2t} \\ 0 & 0 & \mathrm{e}^{2t} \end{pmatrix} = \mathrm{e}^{-t} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{3t} & t \cdot \mathrm{e}^{3t} \\ 0 & 0 & \mathrm{e}^{3t} \end{pmatrix}$$

und daher $\mathrm{e}^{t\cdot A} = S \cdot \mathrm{e}^{t\cdot J} \cdot S^{-1} = e^{-t} \cdot B$, wobei B folgende Gestalt hat:

$$\begin{pmatrix} -1 + (2+t)\cdot \mathrm{e}^{3t} & -2 + (2+t) \cdot \mathrm{e}^{3t} & 4 - (4+t) \cdot \mathrm{e}^{3t} \\ -2t \cdot \mathrm{e}^{3t} & 3 - (2+2t)\cdot \mathrm{e}^{3t} & -6 + (6+2t) \cdot \mathrm{e}^{3t} \\ -t \cdot \mathrm{e}^{3t} & 1 - (1+t) \cdot \mathrm{e}^{3t} & -2 + (3+t) \cdot \mathrm{e}^{3t} \end{pmatrix}$$

Wir können die Matrixexponentialfunktion als Matrix von $n \times n$ differenzierbaren Funktionen auffassen. Leiten wir jede dieser Funktionen ab, schreiben wir die Ableitungen wieder als Matrix und bezeichnen diese mit $\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{t\cdot A}\right)$, so gilt

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{t\cdot A}\right) = A \cdot \mathrm{e}^{t\cdot A}.$$

Das kann direkt für Matrizen in jordanscher Normalform nachgerechnet werden und folgt daraus allgemein durch Transformation mit einer Umrechnungsmatrix. Sind daher $y_1, \dots, y_n \in \mathbb{R}$ und setzen wir

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \mathrm{e}^{t\cdot A} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

so erfüllen diese Funktionen die Beziehung

$$\begin{pmatrix} y_1'(t) \\ \vdots \\ y_n'(t) \end{pmatrix} = A \cdot \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

mit den Anfangswerten $y_1(0) = y_1, \dots, y_n(0) = y_n$.

Dieser Zusammenhang wird eine zentrale Rolle in der Behandlung linearer Differenzialgleichungssysteme spielen. Ist die Matrix A nämlich diagonalisierbar, ist $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ eine Basis von $\mathbb{R}^n$, bestehend aus Eigenvektoren von A, und sind $\lambda_1, \dots, \lambda_n$ die zugehörigen Eigenwerte, so erhalten wir aus der Matrixexponentialfunktion, dass sich die allgemeine Lösung obiger Beziehung in folgender Form schreibt:

$$r_1 \cdot \boldsymbol{v}_1 \cdot e^{\lambda_1 \cdot t} + r_2 \cdot \boldsymbol{v}_2 \cdot e^{\lambda_2 \cdot t} + \dots + r_n \cdot \boldsymbol{v}_n \cdot e^{\lambda_n \cdot t}$$

Dieser Ansatz wird in Bd. 2, Abschn. 10.5 sorgfältig ausgearbeitet.

Betrachten wir wieder die Matrix (14.11) und $y_1 = 2$, $y_2 = 1$ und $y_3 = 2$, so erhalten wir

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \mathrm{e}^{t\cdot A} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} (2+2t)\cdot \mathrm{e}^{-t} \\ -3\mathrm{e}^{-t} + (4-4t)\cdot \mathrm{e}^{2t} \\ -\mathrm{e}^{-t} + (2-2t) \cdot \mathrm{e}^{2t} \end{pmatrix},$$

und diese Funktionen erfüllen

$$\begin{aligned} y_1' &= 3y_1 + 7y_2 - 13y_3 \\ y_2' &= -2y_1 - 9y_2 + 20y_3 \\ y_3' &= -y_1 - 4y_2 + 9y_3 \end{aligned}$$

und $y_1(0) = 2$, $y_2(0) = 1$, $y_3(0) = 1$, wie direkt nachgerechnet werden kann.

Praxis sehr hilfreiche Zerlegungen von Matrizen (sogar von beliebigen, also auch von nichtquadratischen) erreicht werden. Zunächst verallgemeinern wir dazu die Begriffe der Dreiecks- und Diagonalmatrizen.

Definition

Eine $m\times n$-Matrix $A = (a_{i,j})$ heißt **obere Dreiecksmatrix**, wenn $a_{i,j} = 0$ für alle $i > j$ und **Diagonalmatrix**, wenn $a_{i,j} = 0$ für alle $i \neq j$.

Die Matrizen

$$R_1 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 4 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}, \quad R_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 4 & 3 \\ 0 & 0 & -3 & 2 \end{pmatrix}$$

sind also obere Dreiecksmatrizen, die Matrizen

$$D_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}$$

sind Diagonalmatrizen.

QR-Zerlegung einer Matrix

Zu einer beliebigen reellen $n \times m$-Matrix A existiert eine orthogonale $n\times n$-Matrix Q und eine obere Dreiecksmatrix R mit

$$A = Q \cdot R\,.$$

Es gibt also stets eine orthogonale Matrix Q, so dass

$$R = Q^\top \cdot A$$

eine obere Dreiecksmatrix ist.

Eine Zerlegung $A = Q\cdot R$ mit einer orthogonalen Matrix Q und einer oberen Dreiecksmatrix R heißt **QR-Zerlegung** der Matrix A.

Wir stellen im Folgenden nun einen Algorithmus vor, die **Householder-Transformation**, mit der die QR-Zerlegung gefunden werden kann, und wir erläutern dieses Prinzip anhand eines Beispiels.

Wir betrachten dazu eine allgemeine reelle $n \times m$-Matrix A und als konkretes Beispiel die 3×4-Matrix

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

Konstruiert werden soll eine orthogonale $n\times n$-Matrix Q, sodass $R := Q^\top \cdot A$ eine obere Dreiecksmatrix wird.

Mit $\boldsymbol{a}^{(j)}$ bezeichnen wir die Spalten dieser Matrix, sodass

$$A = \begin{pmatrix} \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(2)} & \cdots & \boldsymbol{a}^{(m)} \end{pmatrix}.$$

Dann gilt

$$Q^\top \cdot A = \begin{pmatrix} Q^\top \cdot \boldsymbol{a}^{(1)} & Q^\top \cdot \boldsymbol{a}^{(2)} & \cdots & Q^\top \cdot \boldsymbol{a}^{(m)} \end{pmatrix},$$

und daher ist ein Q zu finden mit

$$Q^\top \cdot \boldsymbol{a}^{(1)} = \begin{pmatrix} r_{1,1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad Q^\top \cdot \boldsymbol{a}^{(2)} = \begin{pmatrix} r_{1,2} \\ r_{2,2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots$$

Wie bei der Trigonalisierung gehen wir schrittweise vor und konstruieren zunächst ein orthogonales Q_1 mit

$$Q_1^\top \cdot \boldsymbol{a}^{(1)} = \begin{pmatrix} r_{1,1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Da Q_1 orthogonal ist, muss $|Q_1^\top \cdot a^{(1)}| = |a^{(1)}|$ gelten, also

$$Q_1^\top \cdot a^{(1)} = \begin{pmatrix} \pm|a^{(1)}| \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Falls bereits $a_{2,1} = a_{3,1} = \cdots = a_{n,1} = 0$ gilt, so wählen wir $Q_1 = \mathrm{E}_n$. Anderenfalls setzen wir zunächst

$$\boldsymbol{v} = \begin{cases} \boldsymbol{a}^{(1)} - |\boldsymbol{a}^{(1)}| \cdot \mathbf{e}_1\,, & \text{falls } a_{1,1} > 0\,, \\ \boldsymbol{a}^{(1)} + |\boldsymbol{a}^{(1)}| \cdot \mathbf{e}_1\,, & \text{falls } a_{1,1} < 0\,. \end{cases}$$

Im konkreten Beispiel bedeutet das

$$\boldsymbol{v} = \begin{pmatrix} 1-\sqrt{2} \\ 0 \\ 1 \end{pmatrix}.$$

Ferner definieren wir $\boldsymbol{w} = \frac{1}{|\boldsymbol{v}|} \cdot \boldsymbol{v}$, im konkreten Beispiel also

$$\boldsymbol{w} = \frac{1}{\sqrt{4-2\sqrt{2}}} \cdot \begin{pmatrix} 1-\sqrt{2} \\ 0 \\ 1 \end{pmatrix},$$

und setzen

$$Q_1 = \left(\mathrm{E}_n - 2 \cdot \boldsymbol{w} \cdot \boldsymbol{w}^\top\right)^\top.$$

Im Beispiel ergibt das

$$Q_1^\top = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{2}{4-2\sqrt{2}} \cdot \begin{pmatrix} 1-\sqrt{2} \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1-\sqrt{2} & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Man sieht sofort, dass diese Matrix orthogonal ist, und auch im allgemeinen Fall rechnet man leicht nach, dass $Q_1^\top \cdot Q_1 = \mathrm{E}_n = Q_1 \cdot Q_1^\top$, dass also die Matrix orthogonal ist. In der konkreten Situation erhalten wir

$$A_2 := Q_1^\top \cdot A = \begin{pmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} & \sqrt{2} & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 1 & 1 \\ 0 & -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix},$$

und im Allgemeinen ergibt sich

$$Q_1^\top \cdot \boldsymbol{a}^{(1)} = \boldsymbol{a}^{(1)} - 2 \cdot \boldsymbol{w} \cdot \boldsymbol{w}^\top \cdot \boldsymbol{a}^{(1)} = |\boldsymbol{a}^{(1)}| \cdot \mathbf{e}_1,$$

sodass auch hier die erste Spalte von $A_2 = Q_1^\top \cdot A$ nur in der ersten Zeile einen von 0 verschiedenen Eintrag hat.

Im nächsten Schritt streichen wir aus A_2 die erste Zeile und die erste Spalte und erhalten dadurch eine $(n-1) \times (m-1)$-Matrix $\widetilde{A_2}$, mit der wir in analoger Weise fortfahren. Im konkreten Beispiel ist das

$$\widetilde{A_2} = \begin{pmatrix} 1 & 1 & 1 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Die erste Spalte dieser Matrix ist $\widetilde{\boldsymbol{a}}^{(1)} = \begin{pmatrix} 1 \\ -\frac{\sqrt{2}}{2} \end{pmatrix}$, mit der wir wie oben die Vektoren

$$\boldsymbol{v} = \widetilde{\boldsymbol{a}}^{(1)} - |\widetilde{\boldsymbol{a}}^{(1)}| \cdot \mathbf{e}_1 = \begin{pmatrix} 1-\frac{\sqrt{6}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix}$$

und

$$\boldsymbol{w} = \frac{1}{|\boldsymbol{v}|} \cdot \boldsymbol{v} = \frac{1}{\sqrt{3-\sqrt{6}}} \cdot \begin{pmatrix} 1-\frac{\sqrt{6}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix}$$

bilden. Damit konstruieren wir die orthogonale Matrix

$$\widetilde{Q_2}^\top = \mathrm{E}_2 - 2 \cdot \boldsymbol{w} \cdot \boldsymbol{w}^\top = \begin{pmatrix} \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{3} \end{pmatrix}.$$

Hierfür ist

$$\widetilde{A_3} = \widetilde{Q_2}^\top \cdot \widetilde{A_2} = \begin{pmatrix} \frac{\sqrt{6}}{2} & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & -\frac{2\cdot\sqrt{3}}{3} \end{pmatrix}$$

eine obere Dreiecksmatrix. Ergänzen wir $\widetilde{Q_2}^\top$ wie bei der Trigonalisierung zu einer (orthogonalen) 3×3-Matrix

$$Q_2^\top = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{3} \end{pmatrix},$$

so gilt hierfür

$$A_3 = Q_2^\top \cdot A_2 = \begin{pmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} & \sqrt{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{6}}{2} & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} \\ 0 & 0 & -\frac{\sqrt{3}}{3} & -\frac{2\cdot\sqrt{3}}{3} \end{pmatrix}.$$

Setzen wir daher

$$Q = Q_1 \cdot Q_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \end{pmatrix},$$

so erhalten wir, dass

$$\begin{aligned} Q^\top \cdot A &= Q_2^\top \cdot Q_1^\top \cdot A \\ &= Q_2^\top \cdot A_2 \\ &= A_3 \\ &= \begin{pmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} & \sqrt{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{6}}{2} & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} \\ 0 & 0 & -\frac{\sqrt{3}}{3} & -\frac{2\cdot\sqrt{3}}{3} \end{pmatrix} \end{aligned}$$

eine obere Dreiecksmatrix ist.

QR-Zerlegung einer komplexen Matrix

Für eine beliebige komplexe $n \times m$-Matrix A existiert eine Zerlegung

$$A = U \cdot R$$

mit einer unitären $n \times n$-Matrix U und einer oberne Dreiecksmatrix R.

Kommentar Für eine reguläre $n \times n$-Matrix ist die QR-Zerlegung $A = Q \cdot R$ eindeutig, wenn wir fordern, dass die Diagonalelemente $r_{i,i}$ der oberen Dreiecksmatrix R alle positiv sind. ◀

Auch nichtinvertierbare Matrizen besitzen eine Pseudoinverse

Die QR-Zerlegung basiert darauf, eine Orthonormalbasis des Spaltenraumes $\mathrm{Bild}(A) \subset \mathbb{R}^m$ von A zu finden. Bei zusätzlicher geschickter Wahl einer Orthonormalbasis des Definitionsbereichs $\mathbb{R}^n$ von A kann noch etwas mehr erreicht werden.

Singulärwertzerlegung einer Matrix

Zu einer beliebigen reellen $m \times n$-Matrix A existieren eine orthogonale $m \times m$-Matrix U, eine orthogonale $n \times n$-Matrix V und eine $m \times n$-Diagonalmatrix Σ mit nichtnegativen Einträgen in der Diagonale, so dass

$$A = U \cdot \Sigma \cdot V^\top .$$

Eine Zerlegung einer $m \times n$-Matrix A in $A = U \cdot \Sigma \cdot V^\top$ mit orthogonalen Matrizen U und V und einer Diagonalmatrix Σ heißt **Singulärwertzerlegung** von A. Die Diagonalelemente von Σ heißen die **Singulärwerte** von A.

Eine mögliche Vorgehensweise zur Bestimmung der Singulärwertzerlegung erläutern wir anhand des Beispiels der 3×4-Matrix

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

Zunächst bestimmen wir die 4×4-Matrix

$$B = A^\top \cdot A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} .$$

Diese Matrix ist symmetrisch (jede Matrix der Form $A^\top \cdot A$ ist symmetrisch) und hat das charakteristische Polynom

$$P_B(\lambda) = \lambda^4 - 6\lambda^3 + 10\lambda^2 - 4\lambda$$

und damit die Eigenwerte $\lambda_1 = 0$ (offensichtlich) und $\lambda_2 = 2$ (durch Raten). Nach Division durch $\lambda^2 - 2\lambda$ erhalten wir

$$(\lambda^4 - 6\lambda^3 + 10\lambda^2 - 4\lambda) \div (\lambda^2 - 2\lambda) = \lambda^2 - 4\lambda + 2 .$$

Die zugehörige quadratische Gleichung $\lambda^2 - 4\lambda + 2 = 0$ liefert die beiden weiteren Eigenwerte $\lambda_3 = 2 - \sqrt{2}$ und $\lambda_4 = 2 + \sqrt{2}$. Wir erhalten also vier Eigenwerte, von denen keiner negativ ist (auch das ist immer so bei Matrizen der Form $A^\top \cdot A$). Diese werden nun der Größe nach absteigend angeordnet:

$$\lambda_1 = 2 + \sqrt{2}, \quad \lambda_2 = 2, \quad \lambda_3 = 2 - \sqrt{2}, \quad \lambda_4 = 0 .$$

Da B symmetrisch ist, ist B orthogonal diagonalisierbar. Es gibt also eine Orthonormalbasis von $\mathbb{R}^4$, bestehend aus Eigenvektoren von B, und zwar

$$\boldsymbol{v}_1 = \frac{1}{2 \cdot \sqrt{2 + \sqrt{2}}} \cdot \begin{pmatrix} 1 + \sqrt{2} \\ 1 + \sqrt{2} \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \frac{1}{2} \cdot \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix},$$

$$\boldsymbol{v}_3 = \frac{1}{2 \cdot \sqrt{2 + \sqrt{2}}} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 - \sqrt{2} \\ -1 - \sqrt{2} \end{pmatrix}, \qquad \boldsymbol{v}_4 = \frac{1}{2} \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} .$$

Wir betrachten die von 0 verschiedenen Eigenwerte $\lambda_1, \lambda_2, \lambda_3$ und setzen $\sigma_i = \sqrt{\lambda_i}$, also

$$\sigma_1 = \sqrt{2 + \sqrt{2}}, \quad \sigma_2 = \sqrt{2}, \quad \sigma_3 = \sqrt{2 - \sqrt{2}},$$

und wir bilden

$$\boldsymbol{u}_i = \frac{1}{\sigma_i} \cdot A \cdot \boldsymbol{v}_i ,$$

also hier

$$\boldsymbol{u}_1 = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \end{pmatrix}, \quad \boldsymbol{u}_2 = \frac{1}{2} \cdot \begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \\ 0 \end{pmatrix}, \quad \boldsymbol{u}_3 = \frac{1}{2} \cdot \begin{pmatrix} -1 \\ -1 \\ \sqrt{2} \end{pmatrix} .$$

Dann bilden $\boldsymbol{u}_1, \boldsymbol{u}_2$ und $\boldsymbol{u}_3$ ein orthonormales System von Vektoren (auch diese Konstruktion funktioniert für eine beliebige Matrix der Form $A^\top \cdot A$), in diesem Fall also sogar eine Orthonormalbasis von $\mathbb{R}^3$, da $\operatorname{Rang}(A) = 3$. Gegebenenfalls muss das System der $\boldsymbol{u}_i$ noch durch geeignete Vektoren zu einer Orthonormalbasis von $\mathbb{R}^n$ ergänzt werden.

Mit diesen Daten kann nun die Singulärwertzerlegung konstruiert werden: Wir setzen

$$U = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{2}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{2}}{2} & -\frac{1}{2} \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sqrt{2 + \sqrt{2}} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2 - \sqrt{2}} & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} \frac{1+\sqrt{2}}{2\cdot\sqrt{2+\sqrt{2}}} & -\frac{1}{2} & \frac{1}{2\cdot\sqrt{2+\sqrt{2}}} & -\frac{1}{2} \\ \frac{1+\sqrt{2}}{2\cdot\sqrt{2+\sqrt{2}}} & \frac{1}{2} & \frac{1}{2\cdot\sqrt{2+\sqrt{2}}} & \frac{1}{2} \\ \frac{1}{2\cdot\sqrt{2+\sqrt{2}}} & -\frac{1}{2} & \frac{-1-\sqrt{2}}{2\cdot\sqrt{2+\sqrt{2}}} & \frac{1}{2} \\ \frac{1}{2\cdot\sqrt{2+\sqrt{2}}} & \frac{1}{2} & \frac{-1-\sqrt{2}}{2\cdot\sqrt{2+\sqrt{2}}} & -\frac{1}{2} \end{pmatrix} .$$

Die Matrix U hat also die Vektoren $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ und $\boldsymbol{u}_3$ als Spalten, die Matrix V hat die Vektoren $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ und $\boldsymbol{v}_4$ als Spalten, und die Matrix Σ ist eine Diagonalmatrix mit den Werten σ_1, σ_2 und σ_3 in der Diagonalen. Beachten Sie, dass die Matrizen U und V nach Konstruktion orthogonal sind.

Es kann nun unmittelbar nachgerechnet werden, dass

$$A = U \cdot \Sigma \cdot V^\top .$$

Die Singulärwertzerlegung einer Matrix A wird benutzt, um die **Pseudoinverse** A^+ von A zu bestimmen. Die Pseudoinverse A^+ einer reellen $n \times m$-Matrix A ist definiert als die $m \times n$-Matrix A^+ mit

1. $A \cdot A^+ \cdot A = A$,
2. $A^+ \cdot A \cdot A^+ = A^+$,
3. $(A \cdot A^+)^\top = A \cdot A^+$,
4. $(A^+ \cdot A)^\top = A^+ \cdot A$.

Ist nun $A = U \cdot \Sigma \cdot V^\top$ und sind $\sigma_1, \ldots, \sigma_l$ die von 0 verschiedenen Singulärwerte von A (also die von 0 verschiedenen Diagonalelemente von Σ), so definieren wir zunächst Σ^+ als die $n \times m$-Diagonalmatrix mit

$$\left(\Sigma^+\right)_{i,i} = \frac{1}{\sigma_i} \qquad \text{für } i = 1, \ldots, l .$$

Dann gilt

$$A^+ = V \cdot \Sigma^+ \cdot U^\top .$$

Teil III

Beispiel

Für die Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

haben wir die Singulärwertzerlegung bereits ermittelt. Die Singulärwerte sind hier

$$\sigma_1 = \sqrt{2 + \sqrt{2}}, \quad \sigma_2 = \sqrt{2}, \quad \sigma_3 = \sqrt{2 - \sqrt{2}},$$

und damit ist

$$\Sigma^+ = \begin{pmatrix} \frac{1}{\sqrt{2+\sqrt{2}}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2-\sqrt{2}}} \\ 0 & 0 & 0 \end{pmatrix},$$

und wir erhalten

$$A^+ = V \cdot \Sigma^+ \cdot U^\top = \frac{1}{4} \cdot \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 3 & 1 & -2 \\ 1 & 3 & -2 \end{pmatrix}.$$

Beachten Sie, dass hierfür

$$A \cdot A^+ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

gilt. Natürlich ist nicht $A^+ \cdot A = E_4$, denn $A^+ \cdot A$ hat nur Rang 3. Vielmehr ist

$$A^+ \cdot A = \frac{1}{4} \cdot \begin{pmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}.$$

Dass die definierenden Eigenschaften einer Pseudoinversen hierfür tatsächlich erfüllt sind, rechnet man leicht nach. ◀

Anwendung: Singulärwerte und Optimierung

Matrizen werden häufig bei Optimierungsaufgaben eingesetzt. Dabei können Inputfaktoren $\boldsymbol{x}$ auf verschiedene Weisen eingesetzt werden, um eine Palette von Produkten zu erzeugen, wodurch wiederum bestimmte Erlöse erzielt oder Kosten verursacht werden, die wir mit $\boldsymbol{y}$ bezeichnen. In vielen Fällen wird der Zusammenhang zwischen $\boldsymbol{x}$ und $\boldsymbol{y}$ durch eine lineares Gleichungssystem $A \cdot \boldsymbol{x} = \boldsymbol{y}$ gegeben.

Das Optimierungsproblem reduziert sich dann häufig darauf, einen Vektor $\boldsymbol{x}$ zu finden, der eine bestimmte Größe hat (d. h., der Gesamtwert der einsetzbaren Ressourcen ist fest gegeben) und für den $|A \cdot x|$ minimal oder maximal ist. Auch das lässt sich mithilfe der Singulärwertzerlegung erreichen. Betrachten wir etwa das Problem

Minimiere $|A \cdot x|$ *unter der Bedingung, dass* $|x| = 1$,

so ergibt sich die Lösung unmittelbar aus der Singulärwertzerlegung $A = U \cdot \Sigma \cdot V^\top$ wie folgt: Falls $\mathrm{Kern}(A) \neq \{\mathbf{0}\}$, so erfüllt ein beliebiges $\boldsymbol{x} \in \mathrm{Kern}(A)$ mit $|x| = 1$ diese Forderung. Anderenfalls muss notwendig $m \leq n$ und $\mathrm{Rang}(A) = m$ gelten. Ordnen wir die Singulärwerte der Größe nach absteigend,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m ,$$

und ist $\boldsymbol{v}_m$ die m-te Spalte von V, so löst $\boldsymbol{v}_m$ dieses Minimierungsproblem.

Da V eine orthogonale Matrix ist, ist nämlich $V^\top \cdot \boldsymbol{v}_m = \mathbf{e}_m$, der m-te Standardbasisvektor des $\mathbb{R}^n$. Da Σ eine Diagonalmatrix ist, erhalten wir $\Sigma \cdot \mathbf{e}_m = \sigma_m \cdot \mathbf{e}_m$, und da U eine orthogonale Matrix ist, gilt

$$|A \cdot \boldsymbol{v}_m| = |U \cdot \Sigma \cdot V^\top \cdot \boldsymbol{v}_m| = |U \cdot (\sigma_m \cdot \mathbf{e}_m)| = |\sigma_m \cdot \mathbf{e}_m| = \sigma_m .$$

Ist $\boldsymbol{x}$ ein beliebiger anderer Vektor in $\mathbb{R}^m$ mit $|\boldsymbol{x}| = 1$, so schreiben wir

$$\boldsymbol{x} = r_1 \cdot \boldsymbol{v}_1 + r_2 \cdot \boldsymbol{v}_2 + \cdots + r_m \cdot \boldsymbol{v}_m .$$

Das ist möglich, denn $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ ist eine Basis von $\mathbb{R}^m$. Da diese Basis orthonormal ist, muss

$$1 = |\boldsymbol{x}| = \sqrt{r_1^2 + r_2^2 + \cdots + r_m^2}$$

gelten. Nun erhalten wir

$$V^\top \cdot \boldsymbol{x} = V^\top \cdot (r_1 \cdot \boldsymbol{v}_1 + \cdots + r_m \cdot \boldsymbol{v}_m) = \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix},$$

also

$$\Sigma \cdot V^\top \cdot \boldsymbol{x} = \Sigma \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \begin{pmatrix} \sigma_1 \cdot r_1 \\ \vdots \\ \sigma_m \cdot r_m \end{pmatrix},$$

und damit

$$\begin{aligned} |\Sigma \cdot V^\top \cdot \boldsymbol{x}| &= \sqrt{\sigma_1^2 \cdot r_1^2 + \sigma_2^2 \cdot r_2^2 + \cdots + \sigma_m^2 \cdot r_m^2} \\ &\geq \sqrt{\sigma_m^2 \cdot r_1^2 + \sigma_m^2 \cdot r_2^2 + \cdots + \sigma_m^2 \cdot r_m^2} \\ &= \sigma_m \cdot \sqrt{r_1^2 + \cdots + r_m^2} \\ &= \sigma_m , \end{aligned}$$

wobei wir ausgenutzt haben, dass $\sigma_l \geq \sigma_m$ für alle l. Da U orthogonal, also normerhaltend ist, gilt

$$|A\boldsymbol{x}| = |U \cdot \Sigma \cdot V^\top \cdot \boldsymbol{x}| = |\Sigma \cdot V^\top \cdot \boldsymbol{x}| \geq \sigma_m .$$

Es folgt, dass in der Tat $|A \cdot \boldsymbol{x}| \geq |A \cdot \boldsymbol{v}_m|$ für alle $\boldsymbol{x} \in \mathbb{R}^m$.

Entsprechend wird das Optimierungsproblem

Maximiere $|A \cdot x|$ *unter der Bedingung, dass* $|x| = 1$,

durch den Vektor $\boldsymbol{v}_1$ gelöst.

Im Beispiel der oben behandelten Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

löst also der Vektor

$$\boldsymbol{v}_3 = \frac{1}{2 \cdot \sqrt{2 + \sqrt{2}}} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 - \sqrt{2} \\ -1 - \sqrt{2} \end{pmatrix},$$

den wir im Rahmen der Singulärwertzerlegung bestimmt haben, das Minimierungsproblem, und entsprechend löst der Vektor

$$\boldsymbol{v}_1 = \frac{1}{2 \cdot \sqrt{2 + \sqrt{2}}} \cdot \begin{pmatrix} 1 + \sqrt{2} \\ 1 + \sqrt{2} \\ 1 \\ 1 \end{pmatrix}$$

das Maximierungsproblem.

Weitere Anwendungen der Singulärwertzerlegung finden sich in der (deskriptiven) Statistik bei der Hauptkomponentenanalyse. Ziel der Hauptkomponentenanalyse ist es, mehrdimensionale Daten so in niederdimensionale Räume zu projizieren, dass möglichst wenig Information verloren geht. Das wird mit Hilfe der Singulärwertzerlegung der Kovarianzmatrix der Daten realisiert.

Anwendung: Bestimmung von Normalformen in MATLAB

Mit MATLAB kann die jordansche Normalform einer Matrix mit dem Befehl `jordan` gefunden werden. Dabei ermittelt der Befehl `J = jordan(A)` die jordansche Normalform J von A, und der Befehl `[S, J] = jordan(A)` bestimmt die jordansche Normalform J von A und eine Transformationsmatrix S, die A auf jordansche Normalform bringt:

```
A= [-19,32,-22,56;-6,13,-5,16;...
    30,-41,34,-77;8,-11,8,-18];
J = jordan(A)
J =
     1      0      0      0
     0      3      1      0
     0      0      3      1
     0      0      0      3
[V, J] = jordan(A)
V =
    11     36      0    -10
    11      6     16    -11
   -22    -12    -14     22
   -11      6    -14     11
J =
     1      0      0      0
     0      3      1      0
     0      0      3      1
     0      0      0      3
```

Die Matrix

$$A = \begin{pmatrix} -19 & 32 & -22 & 56 \\ -6 & 13 & -5 & 16 \\ 30 & -41 & 34 & -77 \\ 8 & -11 & -8 & -18 \end{pmatrix}$$

hat also die jordansche Normalform

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Eine QR-Zerlegung einer Matrix liefert in MATLAB der Befehl `qr`, wobei `[Q,R] =qr(A)` eine obere Dreiecksmatrix R und eine orthogonale Matrix Q mit $A = Q \cdot R$ bestimmen:

```
A=[4,-8,-3,5;2,-1,-6,10;4,-5,-3,-1];
[Q,R]=qr(A)
Q =
   -0.6667    0.6667   -0.3333
   -0.3333   -0.6667   -0.6667
   -0.6667   -0.3333    0.6667
R =
   -6.0000    9.0000    6.0000   -6.0000
         0   -3.0000    3.0000   -3.0000
         0         0    3.0000   -9.0000
```

Eine QR-Zerlegung der Matrix

$$A = \begin{pmatrix} 4 & -8 & -3 & 5 \\ 2 & -1 & -6 & 10 \\ 4 & -5 & -3 & -1 \end{pmatrix}$$

ist also

$$A = \begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} -6 & 9 & -6 & -6 \\ 0 & -3 & 3 & -3 \\ 0 & 0 & 3 & -9 \end{pmatrix}.$$

Die Singulärwertzerlegung schließlich kann mit MATLAB mit `svd` gefunden werden: Der Befehl `s = svd(A)` ermittelt den Vektor $\boldsymbol{s}$ der Singulärwerte von A, und `[U,S,V] = svd(A)` bestimmt die komplette Singulärwertzerlegung $A = U \cdot S \cdot V^\top$ von A:

```
 A = [1,-1,1,-1;1,1,1,1;1,2,3,4];
s=svd(A)
s =
    5.7899
    2.0000
    0.6909

[U,S,V] = svd(A)
U =
    0.0640   -0.9806    0.1854
   -0.3202   -0.1961   -0.9268
   -0.9452   -0.0000    0.3265
S =
    5.7899         0         0         0
         0    2.0000         0         0
         0         0    0.6909         0
V =
   -0.2075   -0.5883   -0.6007    0.5000
   -0.3929    0.3922   -0.6647   -0.5000
   -0.5340   -0.5883    0.3445   -0.5000
   -0.7194    0.3922    0.2805    0.5000
```

Die Pseudoinverse einer Matrix A erhalten wir mit `pinv(A)`:

```
A=[1,0,1;1,1,1];
Aplus=pinv(A)
Aplus =

    0.5000         0
   -1.0000    1.0000
    0.5000    0.0000
```

Aufgaben

14.1 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{pmatrix} 2 & -4 \\ -3 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}.$$

14.2 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 2 \\ 2 & -2 \end{pmatrix}.$$

14.3 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 4 & 2 & 1 \end{pmatrix}.$$

14.4 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrix

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

14.5 Bestimmen Sie das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Für welche Werte von α hat diese Matrix (reelle) Eigenwerte?

14.6 Bestimmen Sie das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Für welche Werte von α hat diese Matrix (reelle) Eigenwerte?

14.7 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ -5 & 6 \end{pmatrix}.$$

14.8 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{pmatrix} 1 & 2\mathrm{i} \\ -2\mathrm{i} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \mathrm{i} & -1 \\ 4 & \mathrm{i} \end{pmatrix}.$$

14.9 Wir betrachten die Matrizen

$$A_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

mit reellen Zahlen a, b. Bestimmen Sie die (komplexen) Eigenwerte und Eigenvektoren von $A_{a,b}$ (in Abhängigkeit von a, b).

14.10 Wir betrachten eine $n \times n$-Matrix A, (paarweise verschiedene) Eigenwerte $\lambda_1, \ldots, \lambda_m$ von A und zu jedem λ_i einen Eigenvektor $\boldsymbol{v}_i$. Zeigen Sie, dass die Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ linear unabhängig sind.

14.11 Untersuchen Sie die Matrizen

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A \cdot S_i$ Diagonalgestalt hat.

14.12 Untersuchen Sie die Matrizen

$$A_1 = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 6 & -2 \\ 1 & 2 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A \cdot S_i$ Diagonalgestalt hat.

14.13 Untersuchen Sie die Matrizen

$$A_1 = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -3 & -2 \\ 5 & -1 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A \cdot S_i$ Diagonalgestalt hat.

14.14 Untersuchen Sie die Matrizen

$$A_1 = \begin{pmatrix} 3 & 2 \\ 2 & -2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A_i \cdot S_i$ Diagonalgestalt hat ($i = 1, 2$).

14.15 Untersuchen Sie die Matrizen

$$A_1 = \begin{pmatrix} -2 & 3 \\ -3 & -2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 1 \\ -2 & 2 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A_i \cdot S_i$ Diagonalgestalt hat ($i = 1, 2$).

Teil III

14.16 Untersuchen Sie die Matrix

$$A = \frac{1}{15}\begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. eine Transformationsmatrix S, sodass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

14.17 Untersuchen Sie die Matrizen

$$A_1 = \begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$$

auf Diagonalisierbarkeit. Bestimmen Sie ggf. eine Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A_i \cdot S_i$ Diagonalgestalt hat ($i = 1, 2$).

14.18 Bestimmen Sie Eigenwerte und Eigenvektoren der Matrizen

$$A_1 = \begin{pmatrix} -2 & 6 & 0 \\ -2 & 5 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} -1 & -2 & -1 \\ 6 & 5 & 0 \\ 0 & -2 & -2 \end{pmatrix}$$

und untersuchen Sie, ob sie diagonalisierbar sind. Bestimmen Sie ggf. Transformationsmatrizen S_i, sodass $S_i^{-1} \cdot A_i \cdot S_i$ Diagonalgestalt hat ($i = 1, 2$).

14.19 Wir betrachten die symmetrischen Matrizen

$$A_1 = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}.$$

Bestimmen Sie orthogonale Matrizen S_i, sodass $S_i^\top \cdot A_i \cdot S_i$ Diagonalgestalt hat ($i = 1, 2$).

14.20 Bestimmen Sie (in Abhängigkeit von a) orthogonale Matrizen S_i ($i = 1, 2$), die die symmetrischen Matrizen

$$A_1 = \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$$

diagonalisieren.

14.21 Wir betrachten die Matrizen

$$A_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

mit reellen Zahlen a, b (vgl. Aufgabe 14.9). Bestimmen Sie (in Abhängigkeit von a, b) eine Transformationsmatrix $S = S_{a,b}$, sodass $S^{-1} \cdot A \cdot S$ Diagonalgestalt hat.

14.22 Bestimmen Sie eine Transformationsmatrix S, die die Matrix

$$A = \begin{pmatrix} -56 & -102 & -78 \\ 39 & 71 & 54 \\ -7 & -13 & -10 \end{pmatrix}$$

trigonalisiert.

14.23 Bestimmen Sie eine Transformationsmatrix S, die die Matrix

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 4 & 0 & 2 \\ 3 & -1 & 3 \end{pmatrix}$$

trigonalisiert.

14.24 Bestimmen Sie eine Transformationsmatrix S, die die Matrix

$$A = \begin{pmatrix} -6 & 16 & -11 \\ -3 & 9 & -7 \\ -1 & 4 & -4 \end{pmatrix}$$

in jordansche Normalform bringt.

14.25 Bestimmen Sie $n \times n$-Matrizen A, B, für die gilt

$$e^{A+B} \neq e^A \cdot e^B.$$

14.26 Bestimmen Sie eine QR-Zerlegung der Matrix

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

14.27 Bestimmen Sie eine QR-Zerlegung der Matrix

$$A = \begin{pmatrix} 1 & -4 & 2 \\ 1 & 1 & -4 \\ 1 & 1 & 0 \\ 1 & -4 & -2 \end{pmatrix}.$$

14.28 Bestimmen Sie eine Singulärwertzerlegung der Matrix

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

14.29 Bestimmen Sie eine Singulärwertzerlegung der Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

und bestimmen Sie die Pseudoinverse A^+ von A.