Wiederholung mathematischer Grundlagen

Die folgende Zusammenstellung orientiert sich an den in der Vorlesung *Regelungs- & Systemtechnik 2* Anwendung findenden ingenieurmathematischen Grundlagen. Sie ist bewußt beispielhaft gehalten und als solche nicht dazu gedacht, eine eingängigere Wiederholung und Auseinandersetzung mit den Begriffen und Konzepten anhand einschlägiger Lehrbücher oder Vorlesungsmitschriften der Grundlagenmathematik zu ersetzen. Die Zusammenstellung erhebt keinerlei Anspruch auf Vollständigkeit. Vielmehr soll sie anregen, sich mit den bereits erlernten mathematischen Konzepten auf die Vorlesung fokusiert auseinanderzusetzen. Die mitunter angeführten Schritte zur Lösung einfacherer Aufgaben stellen im allgemeinen nur einen möglichen Lösungsweg dar. Eine in gewisser Hinsicht geschicktere Vorgehensweise ist durchaus möglich.

1. Lösung linearer skalarer Differentialgleichungen

Die Lösung der skalaren linearen Differentialgleichung

$$\dot{x}(t) = a(t)\,x(t) + u(t) \tag{1}$$

mit $x(t), u(t), a(t) \in \mathbb{R}$ ist gegeben als Summe der homogenen Lösung $x_h(t)$ und einer partikulären Lösung $x_p(t)$, d.h.

$$x(t) = x_h(t) + x_p(t)\,.$$

Die zugehörige homogene Differentialgleichung

$$\dot{x}_h(t) = a(t)\,x_h(t)\,.$$

löst die Funktion

$$x_h(t) = \exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right) c$$

mit beliebiger Konstanten $c \in \mathbb{R}$. Eine partikuläre Lösung $x_p(t)$ von (1) erhält man dann z.B. über die Variation der Konstanten, d.h. über den speziellen Lösungsansatz

$$x_p(t) = \exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right) c(t).$$

Einsetzen des Ansatzes in (1) führt auf

$$\begin{aligned}
&\exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right) a(t)c(t) + \exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right) \dot{c}(t) = a(t)\exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right) c(t) + u(t)\\
\iff\quad & \dot{c}(t) = \exp\left(\int_{t_0}^{t} -a(\tau)\,d\tau\right) u(t)\\
\iff\quad & c(t) = \int_{t_0}^{t} \exp\left(\int_{t_0}^{\bar{\tau}} -a(\tau)\,d\tau\right) u(\bar{\tau})\,d\bar{\tau} + c_0
\end{aligned}$$

für beliebiges $c_0 \in \mathbb{R}$. Damit ist eine partikuläre Lösung (für die Wahl $c_0 = 0$) nun

$$\begin{aligned}
x_p(t) &= \exp\left(\int_{t_0}^{t} a(\tau)\,d\tau\right) \int_{t_0}^{t} \exp\left(\int_{t_0}^{\bar{\tau}} -a(\tau)\,d\tau\right) u(\bar{\tau})\,d\bar{\tau}\\
&= \int_{t_0}^{t} \exp\left(\int_{\bar{\tau}}^{t} a(\tau)\,d\tau\right) u(\bar{\tau})\,d\bar{\tau}
\end{aligned}$$

Damit ergibt sich die an den Anfangswert $x(t_0) = x_0$ angepaßte allgemeine Lösung von (1) zu

$$x(t) = \exp\left(\int_{t_0}^{t} a(\tau)\, d\tau\right) x_0 + \int_{t_0}^{t} \exp\left(\int_{\bar{\tau}}^{t} a(\tau)\, d\tau\right) u(\bar{\tau})\, d\bar{\tau}\,. \tag{2}$$

Für $a(t) \equiv a =$ konst. erhält die Lösung (2) die etwas einfachere Darstellung

$$x(t) = \exp\left(a(t-t_0)\right) x_0 + \int_{t_0}^{t} \exp\left(a(t-\tau)\right) u(\tau)\, d\tau\,.$$

2. Rang, Bild und Kern einer Matrix

Der Spaltenrang einer Matrix $A \in \mathbb{R}^{m\times n}$ ist die Anzahl von deren linear unabhängigen Spaltenvektoren. Der Zeilenrang einer Matrix $A \in \mathbb{R}^{m\times n}$ ist die Anzahl von deren linear unabhängigen Zeilenvektoren. Man kann zeigen, daß Spaltenrang und Zeilenrang stets identisch sind. Daher spricht man häufig auch nur vom Rang einer Matrix.

Der Rang einer Matrix kann z.B. mittels elementarer Zeilenoperationen festgestelt werden. Dabei versucht man durch elementare Zeilen- bzw. Spaltenoperation, d.h. Tausch von Zeilen (Spalten) und durch Addition von Vielfacher einer anderen Zeile (Spalte) auf eine Zeile (Spalte), die Matrix so umzuformen, daß die Anzahl der linear unabhängigen Zeilen (Spalten) abgelesen werden kann. Diese Operationen verändern den Rang der Matrix nicht.

Beispiel:

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

Damit ist $\mathrm{Rang}(A) = 3$.

Das Bild einer Matrix $A \in \mathbb{R}^{m\times n}$, bezeichnet als $\mathrm{Bild}(A)$, ist der Raum, den die Spaltenvektoren der Matrix aufspannen. Für die Dimension dieses Raumes gilt offenbar

$$\dim(\mathrm{Bild}(A)) = \mathrm{Rang}(A).$$

Da die Spalten- bzw. Zeilenvektoren einer Matrix nicht notwendigerweise linear unabhängig sein müssen, ist die Dimension des Bildraums

$$\dim(\mathrm{Bild}(A)) \leq \min\{m, n\}.$$

Im vorangehenden Beispiel:

$$\mathrm{Bild}(A) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} \right\} = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Der Kern oder Nullraum einer Matrix $A \in \mathbb{R}^{m\times n}$, bezeichnet als $\mathrm{Kern}(A)$, ist der Raum, den alle möglichen Lösungsvektoren $x \in \mathbb{R}^n$ der Gleichung $A\,x = 0$ aufspannen. Für die Dimensionen von Bild und Kern einer Matrix gilt die Beziehung:

$$\dim(\mathrm{Bild}(A)) + \dim(\mathrm{Kern}(A)) = n\,.$$

Im vorangehenden Beispiel:

Zur Bestimmung von $\mathrm{Kern}(A)$ lösen wir das Gleichungssystem $A\,x = 0$. Wir nutzen, daß

$$\begin{pmatrix} 2 & 1 & 1 & -2 \\ 1 & -1 & 2 & 1 \\ -1 & -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0 \quad \Longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0$$

Demnach ist $x_4 = 0$ und x_3 kann frei gewählt werden. Damit ist $x_1 = -x_3$, $x_2 = x_3$, und somit

$$\mathrm{Kern}(A) = \mathrm{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Die Dimension des Kerns ist $\dim(\mathrm{Kern}(A)) = 1$, wie erwartet.

3. Determinante, charakteristisches Polynom und Spur einer Matrix

Die Begriffe Determinante und charakteristisches Polynom einer Matrix beziehen sich nur auf quadratische Matrizen $A \in \mathbb{R}^{n\times n}$. Anstatt die allgemeine Definition zu wiederholen, soll hier nur die Bestimmung der Determinanten einer Matrix mit Hilfe des Laplaceschen Entwicklungssatzes an einem Beispiel verdeutlicht werden.

Beispiel:

Wir bestimmen die Determinante der Matrix

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

Die Einfärbung in den Farben rot und blau soll dabei negative und positive Vorzeichen für das bei der Berechnung zu verwendende Vorzeichenschachbrett kenntlich machen.

Um den Aufwand zu reduzieren, ist es geschickt nach der 4. Zeile zu entwickeln und die Derminante der verbleibenden Matrix dann anschließend nach der 2. Spalte (oder 3. Zeile). Wir erhalten

$$\det(A) = 3 \begin{vmatrix} \color{blue}{2} & \color{red}{-1} & \color{blue}{-2} \\ \color{red}{1} & \color{blue}{2} & \color{red}{1} \\ \color{blue}{-1} & \color{red}{0} & \color{blue}{-1} \end{vmatrix} = 3\left(1 \begin{vmatrix} \color{blue}{1} & \color{red}{1} \\ \color{red}{-1} & \color{blue}{-1} \end{vmatrix} + 2 \begin{vmatrix} \color{blue}{2} & \color{red}{-2} \\ \color{red}{-1} & \color{blue}{-1} \end{vmatrix}\right) = 3\,(0-8) = -24$$

Für $A, \bar{A} \in \mathbb{R}^{n\times n}$ gilt:

$$\det(A\,\bar{A}) = \det(A)\,\det(\bar{A}).$$

Das Polynom $\det(\lambda\, I - A)$ heißt charakteristisches Polynom der Matrix A. Die dabei verwendete Matrix $\lambda\, I - A$ heißt charakteristische Matrix von A.

Im vorangehenden Beispiel:

$$\det(\lambda\, I - A) = \begin{pmatrix} \color{blue}{\lambda-2} & \color{red}{-1} & \color{blue}{1} & \color{red}{2} \\ \color{red}{-1} & \color{blue}{\lambda+1} & \color{red}{-2} & \color{blue}{-1} \\ \color{blue}{1} & \color{red}{1} & \color{blue}{\lambda} & \color{red}{1} \\ \color{red}{0} & \color{blue}{-3} & \color{red}{0} & \color{blue}{\lambda} \end{pmatrix} = \cdots = \lambda^4 - \lambda^3 - 5\lambda^2 + 14\lambda - 24$$

Das charakteristische Polynom läßt sich in der Form

$$\det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

schreiben. Der Vorfaktor des Monoms höchsten Grades, d.h. λ^n, ist dabei stets 1 und das konstante Element a_0 des charakteristischen Polynoms erfüllt $a_0 = (-1)^n \det(A)$. Der Koeffizient a_{n-1} des Monoms zweithöchsten Grades ergibt sich als $a_{n-1} = -\mathrm{spur}(A)$. Dabei meint der Ausdruck $\mathrm{spur}(A)$ die Spur der Matrix A, d.h. die Summe der Diagonaleinträge von A.

Setzt man in das charakteristische Polynom $\det(\lambda I - A)$ für den Parameter λ die Matrix A selbst ein, dann ergibt sich stets die Nullmatrix (Satz von Cayley-Hamilton). Demnach gilt

$$A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I = 0\,.$$

Damit können alle Potenzen A^i mit $i \geq n$ als Linearkombination von Potenzen niedrigeren Grades ausgedrückt werden.

4. Besondere Matrizen

Für symmetrische Matrizen $A \in \mathbb{R}^{n\times n}$ gilt:

$$A^{\mathrm{T}} = A\,.$$

Diagonale Matrizen sind als Sonderfall trivialerweise symmetrisch.

Für schiefsymmetrische, oder auch antisymmetrisch genannte Matrizen $A \in \mathbb{R}^{n\times n}$ gilt:

$$A^{\mathrm{T}} = -A\,.$$

Die Adjungte zu einer Matrix A, bezeichnet als $\mathrm{adj}(A)$, bestimmt sich komponentenweise jeweils durch Bildung der Unterdeterminanten, die man erhält, wenn man Zeile und Spalte des betrachteten Elements in A streicht und das Vorzeichenschachbrett berücksichtigt. Die sich so ergebende transponierte Matrix ist die Adjungte von A.

Beispiel

Die Adjunkte von

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

bestimmt sich gemäß

$$\mathrm{adj}(A) = \begin{pmatrix} \begin{vmatrix} -1 & 2 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ -1 & -1 \end{vmatrix} \\[2ex] -\begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix} & \begin{vmatrix} 2 & -1 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ -1 & -1 \end{vmatrix} \\[2ex] \begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 2 & -2 & -2 \\ 1 & -1 & 1 \\ 1 & -5 & -3 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 2 & 1 & 1 \\ -2 & -1 & -5 \\ -2 & 1 & -3 \end{pmatrix}$$

Die Inverse A^{-1} einer Matrix A erfüllt die Beziehungen

$$A^{-1}A = A\,A^{-1} = I$$

Sie existiert genau dann, wenn die Matrix A nicht-singulär ist, d.h. $\det(A) \neq 0$. Offenbar gilt

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Die Inverse läst sich über die Adjunkte nach der Beziehung

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

bestimmen. Diese Beziehung ist vor allem dann von Relevanz, wenn die Matrix A nicht zahlenmäßig, sondern in Form von symbolischen Ausdrücken gegeben ist.

Zur Bestimmung der Inversen einer zahlenmäßig gegebenen Matrix A, insbesondere wenn sie viele Nullen aufweist, ist es häufig geschickter, die erweiterte Matrix $(A|I)$ mittels elementarer Zeilen- bzw. Spaltenoperation auf die Form $(I|A^{-1})$ zu transformieren.

Im vorangehenden Beispiel:

$$(A|I) = \left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 0 & 3 & -5 & 1 & -2 & 0 \\ 1 & -1 & 2 & 0 & 1 & 0 \\ 0 & -2 & 2 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 0 & 0 & -2 & 1 & -\frac{1}{2} & \frac{3}{2} \\ 1 & 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -1 & 0 & -\frac{1}{2} & -\frac{1}{2} \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{4} & -\frac{3}{4} \\ 1 & 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{4} & -\frac{5}{4} \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{4} & -\frac{3}{4} \end{array}\right) \Rightarrow A^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{4} & -\frac{5}{4} \\ -\frac{1}{2} & \frac{1}{4} & -\frac{3}{4} \end{pmatrix}$$

Für orthogonale Matrizen $A \in \mathbb{R}^{n \times n}$ gilt besonders vereinfachend

$$A^{\mathrm{T}} A = A\, A^{\mathrm{T}},$$

d.h. es ist $A^{-1} = A^{\mathrm{T}}$. Determinanten orthogonaler Matrizen sind entweder 1 oder -1.

Die Matrixexponentialfunktion einer Matrix A ist definiert über die Reihe

$$e^A := \sum_{i=0}^{\infty} \frac{A^i}{i!}\,.$$

Für zwei Matrizen A, $B \in \mathbb{R}^{n \times n}$ gilt[1]

$$e^A\, e^B = e^{A+B} \quad \Longleftrightarrow \quad A\,B = B\,A\,,$$

d.h. die für Skalare übliche Regel für Exponenten gilt im Matrixfall nur, wenn A und B kommutieren.

Für jede quadratische Matrix $A \in \mathbb{R}^{n \times n}$ existiert die Inverse von e^A. Sie lautet:

$$\left(e^A\right)^{-1} = e^{-A}\,.$$

[1] Man kann zeigen, daß die Reihe der Matrixexponentialfunktion absolut konvergiert. Dann kann die Cauchysche Regel für das Produkt absolut konvergenter Reihen $\sum_{i=0}^{\infty} a_i$ und $\sum_{j=0}^{\infty} b_j$ verwendet werden. Das Produkt der Reihen ist dann ebenso absolut konvergent und es gilt:

$$\left(\sum_{i=0}^{\infty} a_i\right)\left(\sum_{j=0}^{\infty} b_j\right) = \sum_{k=0}^{\infty} c_k \quad \text{mit} \quad c_k = \sum_{n=0}^{k} a_n\, b_{k-n}\,.$$

5. Eigenwerte und Eigenvektoren einer Matrix, Ähnlichkeitstransformation

Die Eigenwerte einer Matrix $A \in \mathbb{R}^{n\times n}$ sind gerade die Nullstellen (Wurzeln) des charakteristischen Polynoms $\det(\lambda\, I - A)$. Für reellwertige Matrizen A sind die Koeffizienten des charakteristischen Polynoms reell. Damit sind die Eigenwerte entweder reell oder sie treten als konjugiert komplexe Paare auf. Zu einer Matrix der Dimension $n \times n$ gehören n Eigenwerte $\lambda_1, \dots, \lambda_n$, die aber im allgemeinen nicht paarweise voneinander verschieden sein müssen. Es liegen dann mehrfache Nullstellen des charakteristischen Polynoms vor. Die Vielfachheit der unterschiedlichen Eigenwerte λ_i wird als algebraische Vielfachheit $\mathrm{mult}_{\mathrm{alg}}(\lambda_i)$ bezeichnet.

Die Eigenwerte einer Matrix sind genau diejenigen Zahlen λ, welche die charakteristische Matrix $\lambda\, I - A$ singulär machen. Damit ist für jeden Eigenwert λ_i auch das zugehörige homogene lineare Gleichungssystem

$$(\lambda_i\, I - A)\, v_i = 0,$$

d.h. die sogenannte charakteristische Gleichung lösbar. Lösungsvektoren $v_i \in \mathbb{C}^n$ können mit beliebigen, von null verschiedenen Faktoren skaliert bzw. normiert werden und sind dann wieder Lösungen der charakteristischen Gleichung. Im allgemeinen lösen zu einem Eigenwert λ_i mehrere linear unabhängige Vektoren die charakteristische Gleichung. Diese zu einem Eigenwert λ_i gehörenden linear unabhängigen Vektoren, heißen Eigenvektoren des Eigenwerts λ_i. Alle zusammen spannen den zu λ_i gehörigen Eigenraum auf.

Zu einem Eigenwert λ_i gibt es genau so viele linear unahängige Lösungsvektoren der charakteristischen Gleichung, wie der Rangdefekt der charakteristischen Matrix ist, d.h. deren Anzahl ist genau $\dim(\mathrm{Kern}(\lambda_i\, I - A) = n - \dim(\mathrm{Bild}(\lambda_i\, I - A)$. Man nennt diese Anzahl auch geometrische Vielfachheit $\mathrm{mult}_{\mathrm{geo}}(\lambda_i)$ des Eigenwerts λ_i.

Matrizen, bei denen für jeden Eigenwert λ_i algebraische und geometrische Vielfachheit zusammenfallen, d.h. wenn für alle Eigenwerte λ_i einer Matrix gilt

$$\mathrm{mult}_{\mathrm{geo}}(\lambda_i) = \mathrm{mult}_{\mathrm{alg}}(\lambda_i)$$

heißen diagonalisierbar, andernfalls nicht-diagonalisierbar.

Da stets gilt

$$1 \leq \mathrm{mult}_{\mathrm{geo}}(\lambda_i) \leq \mathrm{mult}_{\mathrm{alg}}(\lambda_i)\,,$$

stellen Matrizen mit paarweise verschiedenen Eigenwerten einen Sonderfall der diagonalisierbaren Matrizen dar. Ist eine Matrix diagonalisierbar, so sagt man auch, sie ist diagonalähnlich.

Jede diagonalisierbare Matrix $A \in \mathbb{R}^{n\times n}$ läßt sich mit einer nicht-singulären, im allgemeinen aber komplexwertigen Matrix $V \in \mathbb{C}^{n\times n}$ in folgender Form schreiben:

$$A = V\, \Lambda\, V^{-1}\,.$$

Dabei ist Λ eine aus den Eigenwerten von A gebildete Diagonalmatrix

$$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

und V bestimmt sich aus den in den zugehörigen Spalten angeordneten Eigenvektoren, d.h.

$$V = \left(v_1^{(1)}, \dots, v_1^{(\kappa_1)}, \dots, v_N^{(1)}, \dots, v_N^{(\kappa_N)}\right)$$

der entsprechenden geometrischen Vielfachheit $\kappa_i = \mathrm{mult}_{\mathrm{geo}}(\lambda_i)$ der N paarweise verschiedenen Eigenwerte (Reihenfolge beachten!).

Zwei Matrizen $A, \bar{A} \in \mathbb{C}^{n\times n}$ heißen zueinander ähnlich, wenn sie über eine sogenannte Ähnlichkeitstransformation in Zusammenhang stehen, d.h. wenn es eine nicht-singuläre Matrix $T \in \mathbb{C}^{n\times n}$ so gibt, daß

$$\bar{A} = T^{-1} A\, T\,.$$

Zu einander ähnliche Matrizen haben u.a. das selbe charakterische Polynom, wie man leicht mit Hilfe der Regeln zur Determinatenbildung von Matrizenprodukten herleiten kann.

Auch an einer Matrix durchgeführte elementare Zeilen- und Spaltenumformungen ändern deren Eigenwerte nicht.

Beispiel 1

Wir untersuchen die Eigenwerte und Eigenvektoren der folgenden Matrix vom Typ Frobenius-Matrix (auch Begleitmatrix genannt)

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

Eigenwerte

Das charakeristische Polynom berechnet sich zu

$$\begin{aligned} \det(\lambda\, I - A) &= \det \begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 1 & 1 & \lambda+1 \end{pmatrix} = 1 \begin{vmatrix} -1 & 0 \\ \lambda & -1 \end{vmatrix} - 1 \begin{vmatrix} \lambda & 0 \\ 0 & -1 \end{vmatrix} + (\lambda+1) \begin{vmatrix} \lambda & -1 \\ 0 & \lambda \end{vmatrix} \\ &= \lambda^3 + \lambda^2 + \lambda + 1 \end{aligned}$$

Die Nullstellen dieses Polynoms, d.h. die Eigenwerte von A, sind $\lambda_1 = -1$, $\lambda_2 = j$ und $\lambda_3 = -j$. Jeder Eigenwert hat die algebraische Vielfachheit $\mathrm{mult}_{\mathrm{alg}}(\lambda_i) = 1$, d.h. die Eigenwerte sind paarweise verschieden und A ist diagonalisierbar.

Eigenvektoren

Eigenvektoren zu $\lambda_1 = -1$: Wir suchen Vektoren $v_1 \in \mathbb{C}^3$ mit

$$(\lambda_1\, I - A)\, v_1 = 0 \quad \Longleftrightarrow \quad \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix} v_1 = 0\,.$$

Elementare Zeilenumformungen liefern:

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

Wählen wir die zweite Komponente von v_1 frei, so sieht man, daß alle Vektoren

$$v_1 = c_1 \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \quad \text{mit } c_1 \neq 0$$

Eigenvektoren zum Eigenwert $\lambda_1 = -1$ sind. Ohne Beschränkung der Allgemeinheit wählen wir die Konstante c_1 zu $c_1 = -1$ und erhalten als einzigen linear unabhängigen Eigenvektor

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Eigenvektoren zu $\lambda_2 = j$: Wir suchen Vektoren $v_2 \in \mathbb{C}^3$ mit

$$(\lambda_2\, I - A)\, v_2 = 0 \quad\Longleftrightarrow\quad \begin{pmatrix} j & -1 & 0 \\ 0 & j & -1 \\ 1 & 1 & j+1 \end{pmatrix} v_2 = 0$$

Elementare Zeilenumformungen liefern:

$$\begin{pmatrix} j & -1 & 0 \\ 0 & j & -1 \\ 1 & 1 & j+1 \end{pmatrix} \rightarrow \begin{pmatrix} j & -1 & 0 \\ 0 & j & -1 \\ 0 & 1-j & j+1 \end{pmatrix} \rightarrow \begin{pmatrix} j & 0 & j \\ 0 & j & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Wählen wir die dritte Komponente von v_3 frei, so sieht man, daß alle Vektoren

$$v_2 = c_2 \begin{pmatrix} -1 \\ -j \\ 1 \end{pmatrix} \quad \text{mit } c_2 \neq 0$$

Eigenvektoren zum Eigenwert $\lambda_2 = j$ sind. Ohne Beschränkung der Allgemeinheit wählen wir die Konstante c_2 zu $c_2 = -1$ und erhalten als einzigen linear unabhängigen Eigenvektor

$$v_2 = \begin{pmatrix} 1 \\ j \\ -1 \end{pmatrix}.$$

Eigenvektoren zu $\lambda_3 = -j$: Da zu konjugiert komplexen Eigenwerten konjugiert komplexe Eigenvektoren gehören, folgt sofort der zum Eigenwert $\lambda_3 = -j$ gehörende linear unabhängige Eigenvektor

$$v_3 = \begin{pmatrix} 1 \\ -j \\ -1 \end{pmatrix}.$$

Bei jedem Eigenwert ist die geometrische Vielfachheit gleich der algebraischen Vielfachheit, wie man anhand der paarweisen Verschiedenheit der Eigenwerte schon erschließen konnte.

Damit folgt, daß die Matrix der spaltenweise eingetragenen Eigenvektoren

$$V = \begin{pmatrix} 1 & 1 & 1 \\ -1 & j & -j \\ 1 & -1 & -1 \end{pmatrix}$$

die zugehörige Matrix A diagonalisiert, denn

$$V^{-1} A\, V = \Lambda = \begin{pmatrix} -1 & 0 & 0 \\ 0 & j & 0 \\ 0 & 0 & -j \end{pmatrix},$$

wie man leicht nachprüft.[2]

[2] Da A hier eine Frobenius-Matrix ist, hätte man die Koeffizienten des charakteristischen Polynoms $\det(\lambda\, I - A) = \lambda^n + a_{n-1}\,\lambda^{n-1} + \cdots + a_1\,\lambda + a_0$ natürlich auch direkt als negative Einträge der letzten Zeile von A ablesen können, d.h. es gilt: $a_0 = -A_{n1}$, $a_1 = -A_{n2}$, …, $a_{n-1} = -A_{n,n-1}$. Zudem gilt speziell für Frobenius-Matrizen, daß Eigenwerte eine geometrische Vielfachheit von 1 aufweisen. Hat man die Eigenwerte λ_i der Frobenius-Matrix ermittelt, so bestimmen sich die zugehörigen linear unabhängigen Eigenvektoren v_i einfach aus der Beziehung $v_i = (1, \lambda_i, \lambda_i^2, \ldots, \lambda_i^{n-1})^{\mathrm{T}}$, was den Berechnungsaufwand deutlich reduzieren hilft.

Beispiel 2

Wir untersuchen die Eigenwerte und Eigenvektoren der Matrix

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

Eigenwerte

Da es sich um einer Dreiecksmatrix handelt, kann man die Eigenwerte der Matrix sofort von der Diagonalen ablesen und das charakteristische Polynom ist dann

$$\det(\lambda\, I - A) = (\lambda - 1)^3,$$

d.h. der einzige Eigenwert $\lambda_1 = 1$ hat die algebraische Vielfachheit $\text{mult}_{\text{alg}}(\lambda_1) = 3$.

Eigenvektoren

Wir bestimmen die zu $\lambda = 1$ gehörenden Eigenvektoren, d.h. suchen Vektoren $v_1 \in \mathbb{C}^3$ mit

$$(\lambda_1\, I - A)\, v_1 = 0 \quad \Longleftrightarrow \quad \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} v_1 = 0$$

und erhalten

$$v_1 = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{mit } c_1, c_2 \neq 0\,.$$

Es gibt also zu $\lambda_1 = 1$ genau 2 linear unabhängige Eigenvektoren, d.h. λ_1 hat die geometrische Vielfachheit $\text{mult}_{\text{geo}}(\lambda_1) = 2$. Da die algebraische Vielfachheit jedoch $\text{mult}_{\text{alg}}(\lambda_1) = 3$ beträgt, unterscheiden sich die Vielfachheiten. Konsequenz: Matrix A ist nicht-diagonalisierbar.

6. Eigenwerte und Eigenvektoren besonderer Matrizen

Transponierte Matrix

Die Eigenwerte der Transponierten A^{T} fallen mit den Eigenwerten von A zusammen. Zu jedem Eigenwert sind die zugehörigen Eigenvektoren von A^{T} orthogonal zu den entsprechenden Eigenvektoren von A.

Symmetrische Matrix

Symmetrische Matrizen weisen ausschließlich reelle Eigenwerte auf. Sind immer diagonalähnlich und Eigenvektoren zu verschiedenen Eigenwerten sind zueinander orthogonal.

Zu jeder symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ kann eine orthogonale Transformationsmatrix $T \in \mathbb{R}^{n \times n}$ gefunden werden, so daß[3]

$$T^{\mathrm{T}} A\, T = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

mit den reellen Eigenwerten $\lambda_1, \ldots, \lambda_n$ von A.

[3]Dies gelingt auf Grundlage der Eigenvektoren von A, indem man je (unterschiedlichen) Eigenwert alle zugehörigen Eigenvektoren des Eigenraums einer Orthonormalisierung (Gram-Schmidt-Verfahren) unterzieht und diese Vektoren dann spaltenweise in der Matrix T entsprechend der Numerierung der Eigenwerte anordnet.

Inverse Matrix

Die Eigenwerte von A^{-1} ergeben sich als Kehrwerte der Eigenwerte von A. Die zu den jeweiligen Eigenwerten gehörenden Eigenvektoren sind genau die entsprechenden von A.

Orthogonale Matrix

Jeder Eigenwert λ einer orthogonalen Matrix hat den Betrag $|\lambda| = 1$, d.h. alle Eigenwerte befinden sich auf dem Einheitskreis der komplexen Zahlenebene.

6. Definitheit

Die Definitheit einer Matrix definiert man üblicherweise anhand der Bilinearform $x^{\mathrm{T}} A\, x$. Da für beliebige $A \in \mathbb{R}^{n \times n}$ stets die Zerlegung

$$A = \frac{1}{2}\left(A + A^{\mathrm{T}}\right) + \frac{1}{2}\left(A - A^{\mathrm{T}}\right)$$

gilt, kann man jede Matrix als Summe einer symmetrischen und einer schiefsymmetrischen Matrix darstellen. Man zeigt leicht, daß wenn A schiefsymmetrisch ist, für belieibge $x \in \mathbb{R}^n$ gilt: $x^{\mathrm{T}} A\, x = 0$. Daher führt man Definitheit von Matrizen häufig nur für symmetrische Matrizen ein.

Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ heißt,

- positiv definit, wenn $x^{\mathrm{T}} A\, x > 0$
- positiv semidefinit, wenn $x^{\mathrm{T}} A\, x \geq 0$
- negativ definit, wenn $x^{\mathrm{T}} A\, x < 0$
- negativ semidefinit, wenn $x^{\mathrm{T}} A\, x \leq 0$

für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt. Sie heißt indefinit, wenn sie weder positiv noch negativ semidefinit ist.

Kriterien I: Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ ist

- positiv definit, wenn alle Eigenwerte $\lambda_i > 0$
- positiv semidefinit, wenn alle Eigenwerte $\lambda_i \geq 0$
- negativ definit, wenn alle Eigenwerte $\lambda_i < 0$
- negativ semidefinit, wenn alle Eigenwerte $\lambda_i \leq 0$
- indefinit, wenn sie positive und negative Eigenwerte hat

Kriterien II: Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ ist

- positiv definit, wenn alle führenden Hauptminoren (Hauptabschnittsdeterminanten) von A positiv sind
- negativ definit, wenn alle führenden Hauptminoren (Hauptabschnittsdeterminanten) von $-A$ positiv sind

Beispiel

Die Matrix

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

ist positiv definit, denn alle führenden Hauptminoren sind positiv:

$$\frac{3}{2} > 0, \qquad \begin{vmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{vmatrix} = 2 > 0, \qquad \begin{vmatrix} \frac{3}{2} & -\frac{1}{2} & -1 \\ -\frac{1}{2} & \frac{3}{2} & 1 \\ -1 & 1 & 3 \end{vmatrix} = 4 > 0\,.$$

7. Jacobi-Matrix

Sei $f = f(x)$ mit $f : \mathrm{R}^n \to \mathrm{R}^m$ eine vektorwertige Funktion der vektorwertigen Größe x. Die Jacobi-Matrix $\frac{\partial f}{\partial x}$ ist gegeben als

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Beispiel 1

Wir bestimmen die Jacobi-Matrix von f mit $f(x) = A\,x$ für $A \in \mathbb{R}^{m\times n}$. Mit $f_i(x) = A_{i1}x_1 + \cdots + A_{in}x_n$, $i = 1, \ldots, m$, zeigt man leicht:

$$\frac{\partial f}{\partial x} = A\,.$$

Beispiel 2

Wir bestimmen die Jacobi-Matrix von f mit $f(x) = x^{\mathrm{T}} A\,x$ für $A \in \mathbb{R}^{n\times n}$, d.h. den Gradienten der skalarwertigen Funktion f. Mit $f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j$ zeigt man:

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \ldots, \frac{\partial f_1}{\partial x_n} \right) = x^{\mathrm{T}} \left(A + A^{\mathrm{T}} \right).$$