Dr. Kai-Friederike Oelbermann
Landesstudienkolleg Hochschule Anhalt

Herbstsemester 2017

Projekte zur linearen Abgebra

Themen

(i) *Blockmatrizen* (leicht)

(ii) *Cramer'sche Regel* (mittel)

(iii) *Gleichung einer Geraden durch zwei Punkte* (leicht)

(iv) *Numerisches Lösen linearer Gleichungssysteme* (mittel)

(v) *Matrizeneigenwertprobleme* (schwer)

(vi) *Orthogonale Matrizen und Eigenwertaufgaben* (sehr schwer)

Aufgabe

(i) Die Studierenden bilden Gruppen mit je 2-4 Personen. Sie wählen ein beliebiges Thema aus der obigen Liste.

(ii) Sie studieren die bereitgestellte Literatur. Rechnen Sie dabei möglichste viele Beispiele. Falls Sie Begriffe nicht kennen, recherchieren Sie diese!

(iii) Bis Freitag den 26. Januar 2019 erstellt jede Gruppe ein Poster im Format A0. Auf dem Poster sollen Definitionen, Theorie, Beispiele und Anwendungen beschrieben werden. Überlegen Sie sich eine Struktur für das Poster. Besorgen Sie sich alle Materialien, die Sie für das Poster benötigen. Seien Sie kreativ!

(iv) Am Freitag hat jede Gruppe 10 Minuten Zeit, um ihr Poster vorzustellen. Überlegen Sie vorher, wie Sie das Poster vorstellen.

Literatur

Seiten 2–16.

1 Blockmatrizen

4.3.8 Blockmatrizen

Bisher wurden Matrizen betrachtet, deren Elemente reelle oder komplexe Zahlen sind. Bei manchen Untersuchungen ist es sinnvoll, Matrizen in Blöcke zu unterteilen. Eine Matrix, deren Elemente selbst Matrizen sind, bezeichnet man als *Blockmatrix* oder *Übermatrix*. Jeder Block als Element der Übermatrix ist dann eine sogenannte *Untermatrix*. Dieses Vorgehen ist insbesondere bei größeren und bei solchen Matrizen nützlich, die Einheitsmatrizen oder Nullmatrizen enthalten.

Beispiel 4.31

$$A = \left(\begin{array}{ccc|c|cc} 3 & 2 & -1 & 6 & 0 & 0 \\ 1 & 0 & 5 & 2 & 0 & 0 \\ \hline 1 & 0 & 0 & 1 & 3 & 8 \\ 0 & 1 & 0 & 4 & -2 & 1 \\ 0 & 0 & 1 & 5 & 1 & 3 \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}$$

ist in 6 Blöcke unterteilt. Darin ist $A_{21} = E$ eine Einheitsmatrix, $A_{13} = N$ eine Nullmatrix und A_{12}, A_{22} sind Spaltenmatrizen.

Bei der Einteilung in Blöcke dürfen die Trennlinien keine Stufen bilden. Wird das beachtet, haben alle Blöcke A_{ik} mit gleichem Zeilenindex i die gleiche Anzahl von Zeilen und alle Blöcke mit gleichem Spaltenindex die gleiche Anzahl von Spalten. Damit ist die Voraussetzung dafür erfüllt, daß man mit den Blöcken einer Blockmatrix nach den üblichen Regeln der Matrizenrechnung umgehen kann. Zu beachten ist dabei lediglich, daß die Elemente nicht Zahlen, sondern Matrizen sind. Voraussetzung für die Multiplikation zweier Blockmatrizen ist die Verkettbarkeit der Untermatrizen. Das folgende Beispiel macht dies deutlich.

Beispiel 4.32

Zu berechnen ist $C = A \cdot B$ aus

$$A = \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & -1 & 3 \\ 0 & 0 & 1 & 5 & 2 \\ \hline 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 \end{array}\right) \quad \text{und} \quad B = \left(\begin{array}{cc|cc} 3 & 4 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ \hline 0 & 0 & 4 & 1 \\ 0 & 0 & 1 & -4 \end{array}\right)$$

Lösung:

Die Blockeinteilung berücksichtigt vorhandene Null- bzw. Einheitsmatrizen. Man erhält

$$\begin{pmatrix} E & A_{12} \\ A_{21} & N \end{pmatrix} \begin{pmatrix} B_{11} & N \\ N & B_{22} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

und daraus

$$\begin{aligned} C_{11} &= EB_{11} + A_{12}N &&= B_{11} \\ C_{12} &= EN + A_{12}B_{22} &&= A_{12}B_{22} \\ C_{21} &= A_{21}B_{11} + NN &&= A_{21}B_{11} \\ C_{22} &= A_{21}N + NB_{22} &&= N \end{aligned}$$

Damit lautet die Produktmatrix

$$C = \begin{pmatrix} B_{11} & A_{12}B_{22} \\ A_{21}B_{11} & N \end{pmatrix}$$

Die Matrizen B_{11} und N können ohne Rechnung übernommen werden. Zu berechnen sind

$$A_{12}B_{22} = \begin{pmatrix} 2 & 1 \\ -1 & 3 \\ 5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 4 & 1 \\ 1 & -4 \end{pmatrix} = \begin{pmatrix} 9 & -2 \\ -1 & -13 \\ 22 & -3 \end{pmatrix}$$

$$A_{21}B_{11} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 2 & 0 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 11 & 6 \\ 7 & 0 \end{pmatrix}$$

Man erhält

$$C = \begin{pmatrix} 3 & 4 & 9 & -2 \\ 2 & 0 & -1 & -13 \\ 1 & -2 & 22 & -3 \\ 11 & 6 & 0 & 0 \\ 7 & 0 & 0 & 0 \end{pmatrix}$$

In ähnlicher Weise können sich Vorteile bei der Berechnung der inversen Matrix ergeben, wenn bei der Zerlegung in Blöcke Null- oder Einheitsmatrizen auftreten. Um die Blöcke der inversen Matrix zu bestimmen, muß man Matrizengleichungen lösen.

2 Cramersche Regel

4.4.3.1 Cramersche Regel

Am Beginn des Abschnitts 4.4.1 wurde ein inhomogenes lineares Gleichungssystem mit 2 Unbekannten betrachtet. Die Formeln (4.11) und (4.12) lassen sich mit Hilfe zweireihiger Determinanten darstellen:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

Demnach erhält man die Lösungen des linearen Gleichungssystems als Quotienten zweier Determinanten, wovon die Nennerdeterminante D mit der Koeffizientendeterminante des Gleichungssystems identisch ist. Die Zählerdeterminante erhält man, indem die Koeffizienten der zu berechnenden Unbekannten durch die rechte Seite des Gleichungssystems ersetzt werden. Das gilt allgemein.

Satz (Cramersche Regel)[1]

> Die Lösungen x_i eines inhomogenen linearen Gleichungssystems $\boldsymbol{Ax} = \boldsymbol{b}$ ergeben sich für $D = \det \boldsymbol{A} \neq 0$ eindeutig aus der Formel
>
> $$x_i = \frac{D_i}{D} \qquad (i = 1, 2, \ldots, n) \tag{4.30}$$
>
> worin die Determinanten D_i aus D entstehen, indem die i-te Spalte durch die Spalte $\boldsymbol{b}$ der rechten Seiten ersetzt wird.

[1] Gabriel Cramer (1704 - 1752), Philosoph und Mathematiker in Genf

Die Voraussetzung $\det \boldsymbol{A} \neq 0$ beschränkt die Anwendbarkeit dieser Regel auf Glei chungssysteme mit quadratischer Koeffizientenmatrix. Man kann sich davon über zeugen, daß für $D = 0$ und $D_i = 0\,(i = 1, 2, \ldots, n)$ das System parameter abhängige frei wählbare Lösungen besitzt, wogegen für $D = 0$ und mindestens ei $D_i \neq 0$ wegen eines Widerspruchs keine Lösung existiert.

Beispiel 4.42

Aus

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 15 \\ x_1 + x_2 \phantom{{}+x_3} &= -2 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$$

folgt

$$D = \begin{vmatrix} 2 & -3 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & -1 \end{vmatrix} = -7$$

$$D_1 = \begin{vmatrix} 15 & -3 & 1 \\ -2 & 1 & 0 \\ 0 & -1 & -1 \end{vmatrix} = -7$$

$$D_2 = \begin{vmatrix} 2 & 15 & 1 \\ 1 & -2 & 0 \\ 1 & 0 & -1 \end{vmatrix} = 21$$

$$D_3 = \begin{vmatrix} 2 & -3 & 15 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{vmatrix} = -28$$

und daraus

$$x_1 = \frac{D_1}{D} = 1\,, \quad x_2 = \frac{D_2}{D} = -3\,, \quad x_3 = \frac{D_3}{D} = 4$$

3 Gleichung einer Geraden durch zwei Punkte

4.4.3.4 Gleichung einer Geraden durch zwei Punkte

Die Gleichung einer Geraden durch die beiden vorgegebenen Punkte $P_1(x_1, y_1)$ und $P_2(x_2, y_2)$ läßt sich aus der sogenannten Zweipunktegleichung errechnen:

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

Eine Umformung ergibt

$$(y - y_1)(x_2 - x_1) - (x - x_1)(y_2 - y_1) = 0$$

und

$$x_2 y - x y_2 - x_1 y + x y_1 + x_1 y_2 - x_2 y_1 = 0$$

bzw.

$$\begin{vmatrix} x_2 & x \\ y_2 & y \end{vmatrix} - \begin{vmatrix} x_1 & x \\ y_1 & y \end{vmatrix} + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = 0$$

Schließlich ermöglicht der Entwicklungssatz von LAPLACE die Darstellung

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x \\ y_1 & y_2 & y \end{vmatrix} = 0$$

4 Numerisches Lösen linearer Gleichungssysteme

4.6 Numerische Lösung linearer Gleichungssysteme

4.6.1 Auswahl eines Lösungsverfahrens

Die bisher gezeigten Beispiele erwecken den Eindruck, als sei ein lineares Gleichungssystem völlig problemlos zu lösen. Das ist ein Irrtum, den man spätestens bei der Lösung sehr umfangreicher Gleichungssysteme mit vielen Unbekannten bemerkt. Auch zahlenmäßig sehr große Unterschiede der Koeffizienten können die Rechnung behindern.

Die vorhandenen Lösungsverfahren lassen sich zwei Gruppen zuordnen, den *direkten Verfahren* oder den *iterativen Verfahren*.

Zu der erstgenannten Gruppe gehören das Austauschverfahren und der Gaußsche Algorithmus, die im Abschnitt 4.1.2 behandelt wurden. Beide Verfahren gelten als Standardverfahren, die sich insbesondere für die Handrechnung gut eignen, die aber auch die Grundlage für leistungsfähige Computerprogramme darstellen. Die Handrechnung ist selbst unter Zuhilfenahme eines Taschenrechners begrenzt auf lineare Gleichungssysteme mit wenigen Variablen. Ein weiteres, im Abschnitt 4.4.3.1 behandeltes Verfahren, die Cramersche Regel, ist beschränkt auf inhomogene lineare Gleichungssysteme mit regulärer Koeffizientenmatrix.

Der Gaußsche Algorithmus verfolgt das Prinzip der schrittweisen Elimination der Unbekannten und wandelt die Koeffizientenmatrix um. Der Vorteil dieses Verfahrens liegt im relativ geringen Schreibaufwand. Das Austauschverfahren ist aufwendiger. Ein vorgegebenes Gleichungssystem $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ muß zunächst umgeformt werden in

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}, \tag{4.35}$$

wodurch die rechten Seiten ein entgegengesetztes Vorzeichen erhalten und $\boldsymbol{y} = \boldsymbol{o}$ vorauszusetzen ist. Diese Voraussetzung berechtigt zur Spaltentilgung. Die Bedingung $\boldsymbol{y} = \boldsymbol{o}$ (bzw. $y_i = 0$ für alle i) erfordert, daß nach der Wahl der Probenelemente in y-Zeilen die Zeilensumme null sein muß. Die weitere Rechnung verläuft so, wie im Abschnitt 4.1.2.1 gezeigt.

Die hier genannten *direkten Verfahren* haben die gemeinsame Eigenschaft, nach endlich vielen Schritten zu einem Ergebnis zu gelangen, das entweder die gesuchte Lösung ist oder die Nichtexistenz einer Lösung offenbart.

Die *iterativen Verfahren* verfolgen das Prinzip der schrittweisen Verbesserung einer Näherungslösung. Sie sind ohne Einsatz geeigneter Computerprogramme praktisch nicht durchführbar. Das gegebene Gleichungssystem ist so umzuformen, daß jede Gleichung nach einer Unbekannten umgestellt wird, wodurch die iterationsfähige Form

$$\boldsymbol{x} = \Phi(\boldsymbol{x}) \tag{4.36}$$

entsteht. Setzt man in die rechte Seite von (4.36) eine erste Näherungslösung $\boldsymbol{x}^{(0)}$ ein, so kann man daraus einen Lösungsvektor $\boldsymbol{x}^{(1)}$ errechnen, von dem man hofft, daß er genauer ist als der Startvektor $\boldsymbol{x}^{(0)}$. Die Iterationsvorschrift

$$\boldsymbol{x}^{(k+1)} = \Phi(\boldsymbol{x}^{(k)})$$

erzeugt eine Folge von Lösungsvektoren $\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(k)}, \ldots$, die bei Konvergenz des Verfahrens die Lösung mit jeder gewünschten Genauigkeit liefert. Einzelheiten und Bedingungen derartiger Verfahren sollen im Rahmen dieses Lehrbuches nicht behandelt werden. Im Bedarfsfall ist dazu geeignete Fachliteratur heranzuziehen.[1]

Sowohl bei den direkten als auch bei den iterativen Verfahren können während der Rechnung numerische Probleme auftreten, auf die im folgenden Abschnitt eingegangen wird.

4.6.2 Konditionierungsprobleme

Bei allen behandelten Lösungsverfahren muß im Laufe der Rechnung mindestens einmal, in der Regel aber mehrmals dividiert werden. Beim Austauschverfahren ist in jedem Schritt durch das Pivotelement zu dividieren, beim Gaußschen Algorithmus durch die Diagonalenelemente der oberen Dreiecksmatrix und bei der Cramerschen Regel durch die Koeffizientendeterminante. Damit das möglich ist, dürfen die erwähnten Elemente nicht null sein.

Sind die Pivotelemente kleine Zahlen nahe null oder ist $\det \boldsymbol{A}$ ein Wert in der Nähe von null, so lassen sich zwar alle Rechnungen ausführen, es kann aber zu großen Ungenauigkeiten kommen. Die verheerende Wirkung derartiger Divisionen wird sofort klar, wenn man bedenkt, daß zum Beispiel die Division durch $0,01$ eine Multiplikation mit 100 bedeutet, wodurch alle Rundungsfehler verhundertfacht werden. Die Subtraktion nahezu gleichgroßer Zahlen kann zu Ziffernauslöschungen führen.

Beispiel 4.50

Für das inhomogene lineare Gleichungssystem

$$\begin{aligned} 5,06x_1 + 2x_2 &= 29,30 \\ 10,14x_1 + 4x_2 &= 58,70 \end{aligned}$$

erhält man $\det \boldsymbol{A} = -0,04$ und $x_1 = 5$, $x_2 = 2$.
Durch Runden auf eine Dezimale entsteht das System

$$\begin{aligned} 5,1x_1 + 2x_2 &= 29,3 \\ 10,1x_1 + 4x_2 &= 58,7 \end{aligned}$$

Hierin sind lediglich die Koeffizienten von x_1 geringfügig verändert worden. Die erneute Rechnung liefert $\det \boldsymbol{A} = 0,2$ und

$$x_1 = \frac{1}{0,2} \begin{vmatrix} 29,3 & 2 \\ 58,7 & 4 \end{vmatrix} = -1$$

[1] z.B. Schwetlick/Kretschmar: Numerische Verfahren für Naturwissenschaftler und Ingenieure. Fachbuchverlag Leipzig.

$$x_2 = \frac{1}{0,2} \begin{vmatrix} 5,1 & 29,3 \\ 10,1 & 58,7 \end{vmatrix} = 17,2$$

Ein sorgloses Runden, ohne die übrigen Bedingungen zu beachten, kann zu völlig anderen Ergebnissen führen.
Der im Beipiel gezeigte Sachverhalt wird durch die folgenden Begriffe beschrieben.

Definition

Ist zwar $\det \boldsymbol{A} \neq 0$, aber der Wert nahe null, so heißt die Matrix $\boldsymbol{A}$ **fast singulär**. Ein lineares Gleichungssystem mit dieser Eigenschaft gilt als **schlecht konditioniert**.

Definition

Als **Maximumnorm** der quadratischen Matrix $\boldsymbol{A}$ bezeichnet man die Zahl

$$\|\boldsymbol{A}\| = \max_{i,k} |a_{ik}| \tag{4.37}$$

d.h. das dem Betrag nach größte Matrixelement bestimmt die Maximumnorm.

Offensichtlich gilt für die Einheitsmatrix $\|\boldsymbol{E}\| = 1$. Neben der Maximumnorm sind noch andere Matrixnormen üblich.
Aus der Definition der Maximumnorm und der des Matrixprodukts ist die Eigenschaft

$$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\| \tag{4.38}$$

herleitbar.

Definition

Die Zahl

$$\operatorname{cond}(\boldsymbol{A}) = \|\boldsymbol{A}\| \cdot \|\boldsymbol{A}^{-1}\|$$

heißt **Kondition** (oder *Konditionalzahl*) von $\boldsymbol{A}$. (4.39)

Es ist stets

$$\operatorname{cond}(\boldsymbol{A}) \geq 1,$$

was unmittelbar aus (4.39), (4.38) und $\|\boldsymbol{E}\| = 1$ folgt. Liegt $\operatorname{cond}(\boldsymbol{A})$ nur wenig über 1, so wird das zugehörige Gleichungssystem als gut konditioniert bezeichnet. Ein Gleichungssystem mit der kleineren Kondition gilt als besser konditioniert als eines mit einer größeren Kondition.

Beispiel 4.51

Im Anschluß an Beispiel 4.50 erhält man

$$\boldsymbol{A}_1 = \begin{pmatrix} 5,06 & 2 \\ 10,14 & 4 \end{pmatrix}, \quad \boldsymbol{A}_1^{-1} = \frac{1}{-0,04} \begin{pmatrix} 4 & -2 \\ -10,14 & 5,06 \end{pmatrix} = \begin{pmatrix} -100 & 50 \\ 253,5 & -126,5 \end{pmatrix}$$

und $\qquad \operatorname{cond}(\boldsymbol{A}_1) = 10,14 \cdot 253,5 = 2570,49$

$$\boldsymbol{A}_2 = \begin{pmatrix} 5,1 & 2 \\ 10,1 & 4 \end{pmatrix}, \quad \boldsymbol{A}_2^{-1} = \frac{1}{0,2} \begin{pmatrix} 4 & -2 \\ -10,1 & 5,1 \end{pmatrix} = \begin{pmatrix} 20 & -10 \\ -50,5 & 25,5 \end{pmatrix}$$

und $\qquad \operatorname{cond}(\boldsymbol{A}_2) = 10,1 \cdot 50,5 = 510,05$

Somit ist $\boldsymbol{A}_2$ besser konditioniert als $\boldsymbol{A}_1$, d.h. bei $\boldsymbol{A}_2$ haben Rechenungenauigkeiten weniger schwere Folgen als bei $\boldsymbol{A}_1$.

Bei der rechentechnischen Bearbeitung derartiger Aufgaben ist zu beachten, daß ein Computer stets nur Zahlen mit begrenzter Ziffernanzahl verarbeiten kann. Verwendet man ein Programm, das die Kondition eines linearen Gleichungssystems nicht prüft, kann es zu Unstimmigkeiten kommen.

5 Matrizeneigenwertprobleme

4.7.1 Matrizeneigenwertprobleme

Die lineare Transformation $\boldsymbol{y} = \boldsymbol{Ax}$ mit einer Matrix $\boldsymbol{A}$ des Typs (m, n) überführt den n-dimensionalen Vektor $\boldsymbol{x}$ in einen m-dimensionalen Vektor $\boldsymbol{y}$. Es soll der Frage nachgegangen werden, ob es Vektoren $\boldsymbol{x}$ gibt, die bei einer derartigen Transformation unverändert bleiben oder in ein Vielfaches von sich selbst übergehen, für die also gilt

$$\boldsymbol{Ax} = \boldsymbol{x} \qquad \text{oder} \qquad \boldsymbol{Ax} = \lambda\boldsymbol{x}\,.$$

Diese Frage ist die Grundfrage der sogenannten Eigenwerttheorie, die eine wichtige Rolle beim Vermeiden instabiler Schwingungen spielt, z.B. des Flatterns von Flugzeugtragflächen oder der Resonanz schwingender Brücken.
Für die folgenden Betrachtungen muß die Matrix $\boldsymbol{A}$ als quadratisch vorausgesetzt werden. Die Elemente von $\boldsymbol{A}$ und des Vektors $\boldsymbol{x}$ dürfen reell- oder komplexwertig sein. Zunächst sollen einige Begriffe vereinbart werden.

Definition

$\boldsymbol{A}$ sei eine komplexe Matrix vom Typ (n, n). Eine Zahl $\lambda \in \mathbb{C}\backslash\{0\}$ heißt ein **Eigenwert** von $\boldsymbol{A}$ und $\boldsymbol{x}$ ein zugehöriger **Eigenvektor**, wenn gilt

$$\boldsymbol{Ax} = \lambda\boldsymbol{x} \quad \text{und} \quad \boldsymbol{x} \neq \boldsymbol{o}\,. \tag{4.40}$$

Die Menge aller Eigenwerte von $\boldsymbol{A}$ heißt **Spektrum** von $\boldsymbol{A}$.

In dieser Definition ist $\boldsymbol{x}$ rechtsseitiger Faktor von $\boldsymbol{A}$, also genauer $\boldsymbol{Ax}_r = \lambda\boldsymbol{x}_r$.
Bei linksseitiger Multiplikation lautet die Aufgabe $\boldsymbol{x}_l^{\mathrm{T}}\boldsymbol{A} = \lambda\boldsymbol{x}_l^{\mathrm{T}}$. Durch Transponieren erhält man $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_l = \lambda\boldsymbol{x}_l$, d.h., die linksseitigen Eigenvektoren einer Matrix stimmen mit den rechtsseitigen Eigenvektoren der transponierten Matrix $\boldsymbol{A}^{\mathrm{T}}$ überein, und demzufolge sind auch die zugehörigen Eigenwerte λ die gleichen. Daher

wird im folgenden das Eigenwertproblem nur für rechtsseitige Eigenvektoren ohne besondere Kennzeichnung betrachtet.
Um eine geometrische Vorstellung von der Gleichung (4.40) zu erhalten, sei eine Darstellung im $\mathbb{R}^2$ gewählt.
$\boldsymbol{x}_0$ ist ein Eigenvektor von $\boldsymbol{A}$ genau dann, wenn $\boldsymbol{Ax}_0 = \lambda_0\boldsymbol{x}_0$, d.h. wenn $\boldsymbol{x}_0$ in ein Vielfaches von sich selbst abgebildet wird. $\lambda_0\boldsymbol{x}_0$ hat die gleiche Richtung wie $\boldsymbol{x}_0$, die Länge ist aber auf das λ_0-fache verändert. Die Vektoren $\lambda_0\boldsymbol{x}_0$ erzeugen damit eine Gerade, die als *Fixgerade* bezeichnet wird, denn durch die Transformation $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ geht jeder beliebige Punkt dieser Geraden wieder in einen Punkt derselben Geraden über. Das Eigenwertproblem $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ besteht darin, alle zu $\boldsymbol{A}$ gehörigen Fixgeraden zu bestimmen.
Diese geometrische Interpretation läßt vermuten, daß es bei der Bestimmung eines Eigenvektors nur auf die Richtung, nicht auf die Länge des Vektors ankommt.

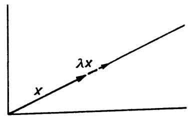

Bild 4.5

Satz

Ist $\boldsymbol{x}$ ein Eigenvektor der Matrix $\boldsymbol{A}$ und $c \neq 0$ eine Zahl mit $c \in \mathbb{C}$, so ist $c\boldsymbol{x}$ ebenfalls ein Eigenvektor zum gleichen Eigenwert, d.h., jeder Eigenvektor ist nur bis auf einen von null verschiedenen Faktor eindeutig bestimmt.

Beweis:
$\boldsymbol{x}$ sei Eigenvektor von $\boldsymbol{A}$, d.h., $\boldsymbol{Ax} = \lambda\boldsymbol{x}$. Ist außerdem $\boldsymbol{y} = c\boldsymbol{x}$, also $\boldsymbol{x} = \frac{1}{c}\boldsymbol{y}$, so erhält man $\boldsymbol{A} \cdot \frac{1}{c}\boldsymbol{y} = \lambda\frac{1}{c}\boldsymbol{y}$ und somit $\boldsymbol{Ay} = \lambda\boldsymbol{y}$. Damit erweist sich auch $\boldsymbol{y}$ als Eigenvektor.

Im folgenden wird gezeigt, wie sich Eigenwerte und Eigenvektoren praktisch berechnen lassen.

Berechnung von Eigenwerten

Die sogenannte *Eigengleichung* $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ läßt sich wegen $\boldsymbol{x} = \boldsymbol{Ex}$ umschreiben in $\boldsymbol{Ax} = \lambda\boldsymbol{Ex}$. Die übliche Darstellung des Eigenwertproblems lautet somit

$$(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{x} = \boldsymbol{o} \tag{4.41}$$

Das ist ein homogenes lineares Gleichungssystem mit der Koeffizientenmatrix $\boldsymbol{A} - \lambda\boldsymbol{E}$ und dem Vektor $\boldsymbol{x}^{\mathrm{T}} = (x_1, x_2, \ldots, x_n)$ der Unbekannten. Anknüpfend an den Abschnitt 4.5.2.3 sind zwei Fälle zu unterscheiden:

1. $\boldsymbol{A} - \lambda\boldsymbol{E}$ ist regulär, d.h. $\det(\boldsymbol{A} - \lambda\boldsymbol{E}) \neq 0$.
   Dann existiert nur die triviale Lösung $\boldsymbol{x} = \boldsymbol{o}$.

2. $\boldsymbol{A} - \lambda\boldsymbol{E}$ ist singulär, d.h. $\det(\boldsymbol{A} - \lambda\boldsymbol{E}) = 0$.
   Dann gibt es auch Lösungen $\boldsymbol{x} \neq \boldsymbol{o}$.

Somit gilt der

Satz

Die Eigengleichung $(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{x} = \boldsymbol{o}$ hat genau dann nichttriviale Lösungen $\boldsymbol{x} \neq \boldsymbol{o}$, wenn

$$\det(\boldsymbol{A} - \lambda\boldsymbol{E}) = 0\,. \tag{4.42}$$

Diese Gleichung heißt *charakteristische Gleichung* der Matrix $\boldsymbol{A}$.

Wird (4.42) ausgeschrieben, so erhält man

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0$$

Nach den Rechenregeln für Determinanten entsteht daraus, geordnet nach Potenzen von λ, das sogenannte *charakteristische Polynom*

$$p(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \ldots a_{nn})\lambda^{n-1} + \ldots + \det\boldsymbol{A} = 0$$

bzw. nach Division durch $(-1)^n$ und unter Verwendung der Spur der Matrix $\boldsymbol{A}$ (s. Abschnitt 4.3.1)

$$\lambda^n - sp(\boldsymbol{A})\lambda^{n-1} + \ldots + (-1)^n\det\boldsymbol{A} = 0\,.$$

Das ist eine algebraische Gleichung n-ten Grades. Nach einer Folgerung aus dem Fundamentalsatz der Algebra besitzt sie genau n reelle oder komplexe Lösungen $\lambda_1, \lambda_2, \ldots, \lambda_n$. Diese sind die gesuchten Eigenwerte der Matrix $\boldsymbol{A}$.

Beispiel 4.52

Zu bestimmen sind die Eigenwerte der Matrix

$$\boldsymbol{A} = \begin{pmatrix} 3 & \frac{5}{9} & -1 \\ 18 & 2 & 0 \\ 2 & 1 & -1 \end{pmatrix}$$

Lösung:
Aus der Bedingung (4.42) erhält man die charakteristische Gleichung

$$\begin{vmatrix} 3-\lambda & \frac{5}{9} & -1 \\ 18 & 2-\lambda & 0 \\ 2 & 1 & -1-\lambda \end{vmatrix} = -\lambda^3 + 4\lambda^2 + 7\lambda - 10 = 0$$

Die Lösungen dieser Gleichung lauten $\lambda_1 = 1$, $\lambda_2 = -2$, $\lambda_3 = 5$.

Beispiel 4.53

Gesucht sind die Eigenwerte zu $\boldsymbol{A} = \begin{pmatrix} 8 & -1 \\ 17 & 6 \end{pmatrix}$.

Lösung:
Die charakteristische Gleichung lautet

$$\begin{vmatrix} 8-\lambda & -1 \\ 17 & 6-\lambda \end{vmatrix} = \lambda^2 - 14\lambda + 65 = 0$$

Daraus erhält man $\lambda_1 = 7 + 4\mathrm{j}$, $\lambda_2 = 7 - 4\mathrm{j}$.

Besonders einfach lassen sich die Eigenwerte berechnen, wenn die Matrix $\boldsymbol{A}$ eine Dreiecksmatrix oder eine Diagonalmatrix ist. Dann gilt der

Satz

Ist $\boldsymbol{A}$ eine Dreiecks- oder eine Diagonalmatrix, so sind die Hauptdiagonalenelemente die Eigenwerte der Matrix.

Der *Beweis* dieses Satzes stützt sich darauf, daß die Determinante einer Dreiecksmatrix gleich dem Produkt der Hauptdiagonalenelemente ist, d.h.

$$\det(\boldsymbol{A} - \lambda\boldsymbol{E}) = (a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda) = 0\,,$$

woraus $\lambda_i = a_{ii}$ für $i = 1, 2, \ldots, n$ folgt.

Berechnung von Eigenvektoren

Zu jedem Eigenwert λ_i gehört ein Eigenvektor $\boldsymbol{x}_i$, der aus dem homogenen linearen Gleichungssystem

$$(\boldsymbol{A} - \lambda_i\boldsymbol{E})\boldsymbol{x}_i = \boldsymbol{o}$$

berechnet werden kann. Wegen $\det(\boldsymbol{A} - \lambda_i\boldsymbol{E}) = 0$ ist $r(\boldsymbol{A} - \lambda_i\boldsymbol{E}) < n$, d.h. mindestens eine der Gleichungen des homogenen Systems ist linear abhängig von den übrigen und mindestens eine der Unbekannten $x_{i1}, x_{i2}, \ldots, x_{in}$ ist frei wählbar. Im Abschnitt 4.5.3 wurde gezeigt, daß die Anzahl der frei wählbaren Unbekannten und damit auch die Anzahl der linear unabhängigen Lösungsvektoren des homogenen Gleichungssystems durch den sogenannten *Rangabfall* $n - r$ festgelegt ist.
Es gilt der

Satz

Die Anzahl der zu einem Eigenwert λ_i gehörigen linear unabhängigen Eigenvektoren ist $n - r(\boldsymbol{A} - \lambda_i\boldsymbol{E})$.

Beispiel 4.54

Laut Beispiel 4.52 besitzt die Matrix

$$A = \begin{pmatrix} 3 & \frac{5}{9} & -1 \\ 18 & 2 & 0 \\ 2 & 1 & -1 \end{pmatrix}$$

die Eigenwerte $\lambda_1 = 1$, $\lambda_2 = -2$, $\lambda_3 = 5$. Zu berechnen sind die zugehörigen Eigenvektoren.

Lösung:

Bezeichnet man den zu $\lambda_1 = 1$ gehörenden Eigenvektor mit $\boldsymbol{x}_1$ und seine Komponenten mit x_{11}, x_{12}, x_{13}, so erhält man aus (4.41) das homogene lineare Gleichungssystem

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{E})\boldsymbol{x}_1 = \begin{pmatrix} 2 & \frac{5}{9} & -1 \\ 18 & 1 & 0 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix} = \boldsymbol{o}$$

bzw.
$$\begin{aligned} 2x_{11} + \tfrac{5}{9}x_{12} - x_{13} &= 0 \\ 18x_{11} + x_{12} \phantom{- 2x_{13}} &= 0 \\ 2x_{11} + x_{12} - 2x_{13} &= 0 \end{aligned}$$

Das Austauschverfahren oder der Gaußsche Algorithmus brechen nach 2 Schritten ab, d.h., es ist $r(\boldsymbol{A} - \lambda_1 \boldsymbol{E}) = 2$ und damit ist die Zahl der frei wählbaren Variablen $n - r = 3 - 2 = 1$.

Man erhält

$$x_{12} = -18x_{11}$$
$$x_{13} = -8x_{11}$$

Setzt man $x_{11} = c_1$, so lautet der gesuchte Eigenvektor

$$\boldsymbol{x}_1 = c_1 \begin{pmatrix} 1 \\ -18 \\ -8 \end{pmatrix},$$

der nur bis auf einen Faktor c_1 bestimmbar ist.

Entsprechende Rechnungen liefern für $\lambda_2 = -2$ den Eigenvektor

$$\boldsymbol{x}_2 = c_2 \begin{pmatrix} 1 \\ -\frac{9}{2} \\ \frac{5}{2} \end{pmatrix}$$

und für $\lambda_3 = 5$ den Eigenvektor

$$\boldsymbol{x}_3 = c_3 \begin{pmatrix} 1 \\ 6 \\ \frac{4}{3} \end{pmatrix}.$$

Wegen $n - r = 1$ existiert zu jedem Eigenwert λ_i nur ein Eigenvektor $\boldsymbol{x}_i$, der bis auf einen Faktor c_i festgelegt ist.

6 Orthogonale Matrizen und Eigenwertaufgaben

5.7 Orthogonale Matrizen und Eigenwertaufgaben

Wir legen im euklidischen Vektorraum $\mathbb{R}^n$ ein kartesisches Koordinaten-System zugrunde und betrachten die durch eine quadratische Matrix A vermittelte Abbildung $\vec{x} \longmapsto A\vec{x}$ des $\mathbb{R}^n$ in sich.

> Eine $n * n$-Matrix A heißt *orthogonal* $\Longleftrightarrow$
> Die zugehörige lineare Abbildung läßt das Skalarprodukt (und damit Abstände und Winkel) unverändert:
> $$A\vec{x} \cdot A\vec{y} = \vec{x} \cdot \vec{y} \,.$$

In Matrizenschreibweise bedeutet dies:

$$A\vec{x} \cdot A\vec{y} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= (x_1, \dots, x_n) \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}}_{A^T} \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}}_{A} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} .$$

Setzt man $\vec{x} = \vec{e}_i$, $\vec{y} = \vec{e}_j$, so folgt

$$(a_{1i}, \dots, a_{ni}) \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \,. \end{cases}$$

D.h., die Spalten von A bilden ein Orthonormal-System, in Matrizenschreibweise:

$$A^T A = E \,.$$

Da wegen $(\det A)^2 = (\det A^T)(\det A) = \det(A^T A) = \det E = 1$ die orthogonale Matrix A regulär ist, folgt durch Multiplikation mit A^{-1} von rechts:

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

Hieraus folgt weiter $A\,A^T = E$, d.h., auch die Zeilen von A bilden ein Orthonormal-System:

> A orthogonal $\Longleftrightarrow$ $A^{-1} = A^T$ $\Longleftrightarrow$ $A^T A = A\,A^T = E$ $\Longleftrightarrow$
> Die Spalten bzw. Zeilen von A bilden ein Orthonormal-System.

Beispiel 5.7.1

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

ist orthogonal, weil die Zeilen (und ebenso die Spalten) Orthonormalsysteme bilden. Man rechnet leicht $A^T A = A\,A^T = E$ nach.
Mit Hilfe des Gauß-Algorithmus ermittelt man die inverse Matrix

$$A^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} ;$$

es ist $A^{-1} = A^T$.

Das Skalarprodukt der Vektoren $A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $A\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ ist

$$(x_1, x_2, x_3) \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} .$$

> Sind im euklidischen $\mathbb{R}^n$ zwei Orthonormal-Basen $\{\vec{e}_1, \dots, \vec{e}_n\}$ (alt) bzw. $\{\vec{f}_1, \dots, \vec{f}_n\}$ (neu) gegeben, so ist die Matrix, welche die Koordinaten eines Vektors bzgl. alter und neuer Basis verknüpft, ebenfalls orthogonal.

Aus $\vec{x} = \sum_{i=1}^n x_i \vec{e}_i = \sum_{j=1}^n \overline{x}_j \vec{f}_j$ erhält man durch skalare Multiplikation mit $\vec{f}_j$ bzw. $\vec{e}_i$:

$$\sum_{i=1}^n x_i (\vec{e}_i \cdot \vec{f}_j) = \overline{x}_j \quad \text{bzw.} \quad \sum_{j=1}^n \overline{x}_j (\vec{f}_j \cdot \vec{e}_i) = x_i \,.$$

Mit $B = (\vec{e}_i \cdot \vec{f}_j)_{i,j=1,\dots,n} = (b_{ij})_{i,j=1,\dots,n}$ lauten die Transformationen:

$$\overline{x}_j = \sum_{i=1}^n b_{ij} x_i \quad \text{bzw.} \quad x_i = \sum_{j=1}^n b_{ij} \overline{x}_j \,,$$

in Matrizenschreibweise:

$$\overline{\vec{x}} = B^T \vec{x} \quad \text{bzw.} \quad \vec{x} = B \overline{\vec{x}} \,,$$

wobei $\vec{x}$ die Koordinatenspalte bzgl. der Basis $\{\vec{e}_1, \dots, \vec{e}_n\}$, $\overline{\vec{x}}$ die Koordinatenspalte bzgl. der Basis $\{\vec{f}_1, \dots, \vec{f}_n\}$ ist. Es folgt

$$B^T = B^{-1} \,.$$

Beispiel 5.7.2 Die Matrizen der Rotations- und Spiegelungs-Transformationen des Koordinaten-Systems aus 5.5 sind orthogonal: Z.B.

$$\begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(Drehung um die } z\text{-Achse)}$$

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(Spiegelung an der } y, z\text{-Ebene).}$$

Wird eine Rotation des Koordinaten-Systems um einen Winkel γ bei festgehaltenen geometrischen Objekten durch

$$\begin{pmatrix} \overline{x} \\ \overline{y} \\ \overline{z} \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

beschrieben, – so beschreibt die Abbildung

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{R^{-1}}_{R^T} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

eine Drehung der geometrischen Objekte um den Drehwinkel $-\gamma$. In jedem Fall sorgt die Orthogonalität der Matrizen R bzw. R^T dafür, daß Längen und Winkel erhalten bleiben.

Als eine Anwendung orthogonaler Matrizen behandeln wir das folgende *Hauptachsen-Problem*:

Gesucht ist eine orthogonale Koordinaten-Transformation $\vec{x} = B\overline{\vec{x}}$ bzw. $\overline{\vec{x}} = \underbrace{B^{-1}}_{B^T}\vec{x}$, so daß die *quadratische Form*

$$Q(\vec{x}) = Q(x_1, \dots, x_n) = \sum_{i,k}^n a_{ik} x_i x_k$$

$$= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} \sum_{k=1}^n a_{1k} x_k \\ \vdots \\ \sum_{k=1}^n a_{nk} x_k \end{pmatrix} = \vec{x} \cdot A\vec{x} = \vec{x}^T A \vec{x}$$

die einfache Gestalt

$$Q(\overline{\vec{x}}) = Q(\overline{x}_1, \dots, \overline{x}_n) = \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix} \cdot \begin{pmatrix} s_1 \overline{x}_1 \\ \vdots \\ s_n \overline{x}_n \end{pmatrix} = \sum_{i=1}^n s_i \overline{x}_i^2$$

bekommt.

Beispiel 5.7.3 Im $\mathbb{R}^2$ haben wir die quadratische Form $5x_1^2 + 6x_1x_2 + 5x_2^2$ durch die Rotation

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix}$$

in die Form $8\overline{x}_1^2 + 2\overline{x}_2^2$ gebracht und so den Kegelschnitt $5x_1^2 + 6x_1x_2 + 5x_2^2 = 2$ im neuen Koordinatensystem als Ellipse $8\overline{x}_1^2 + 2\overline{x}_2^2 = 2$ erkannt.

Zunächst überzeugen wir uns, daß wir die Matrix $A = (a_{ik})_{i,k=1,\dots,n}$ stets *symmetrisch*, d.h. von der Form

$$A = A^T$$

wählen können.
Andernfalls ersetzt man nämlich A durch $A^* = (a^*_{ik}) = \frac{1}{2}(a_{ik} + a_{ki})$:

$$\sum_{i,k=1}^n a^*_{ik} x_i x_k = \frac{1}{2} \sum_{i,k=1}^n a_{ik} x_i x_k + \frac{1}{2} \underbrace{\sum_{i,k=1}^n a_{ki} x_i x_k}_{\sum_{i,k=1}^n a_{ik} x_k x_i} = \sum_{i,k=1}^n a_{ik} x_i x_k \,.$$

Beispiel 5.7.4 Die quadratische Form $Q(x_1, x_2) = 5x_1^2 + 6x_1x_2 + 5x_2^2$ läßt sich durch die Matrix $A = \begin{pmatrix} 5 & 6 \\ 0 & 5 \end{pmatrix}$ beschreiben:

$$Q(x_1, x_2) = (x_1, x_2) \begin{pmatrix} 5 & 6 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2) \begin{pmatrix} 5x_1 + 6x_2 \\ 5x_2 \end{pmatrix} = 5x_1^2 + 6x_1x_2 + 5x_2^2 \,.$$

Man kann zur Beschreibung von $Q(x_1, x_2)$ aber auch von vorneherein die symmetrische Matrix $A^* = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ mit

$$a_{11}^* = a_{11} = 5\,, \quad a_{12}^* = \frac{1}{2}(a_{12} + a_{21}) = 3 = a_{21}^*\,, \quad a_{22}^* = a_{22} = 5$$

verwenden:

$$Q(x_1, x_2) = (x_1, x_2)\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1, x_2)\begin{pmatrix} 5x_1 + 3x_2 \\ 3x_1 + 5x_2 \end{pmatrix} = 5x_1^2 + 6x_1x_2 + 5x_2^2\,.$$

Zur Lösung des Hauptachsenproblems sucht man eine Basis aus Vektoren $\vec{y}$, für die der Bildvektor $A\vec{y}$ jeweils ein skalares Vielfaches $s\vec{y}$ von $\vec{y}$ ist:

> Ein Vektor $\vec{y} \neq \vec{0}$, der bei Anwendung der Matrix A in sein s-faches übergeht, heißt *Eigenvektor* von A *zum Eigenwert* s.

Da $s\vec{y} = sE\vec{y}$ (E Einheitsmatrix) geschrieben werden kann, gilt:

> Eigenvektoren einer Matrix A sind gerade die nicht-trivialen Lösungsvektoren $\vec{y}$ des homogenen linearen Gleichungs-Systems
> $$(A - sE)\vec{y} = \vec{0}\,.$$

Wir suchen zunächst nach Skalaren s, für die dieses Gleichungs-System überhaupt nicht-triviale Lösungen besitzt; dies ist genau dann der Fall, wenn die Determinante dieses Systems verschwindet.

> Die Eigenwerte s_i sind daher die Lösungen der sogenannten *charakteristischen Gleichung*
> $$\det(A - sE) = 0\,.$$

Nach dem Hauptsatz der Algebra hat diese Gleichung n-ten Grades n (im allgemeinen komplexe, eventuell zusammenfallende) Lösungen. Hat man eine von ihnen bestimmt, etwa s_1, so erhält man einen zugehörigen Eigenvektor $\vec{y}_1$ durch Lösung des Gleichungssystems $(A - s_1E)\vec{y} = \vec{0}$. Als Lösung eines homogenen Systems ist $\vec{y}_1$ nur bis auf einen Faktor $\neq 0$ bestimmt, den man bequem wählen kann, etwa so, daß $|\vec{y}_1| = 1$ ist; in diesem Fall nennt man den Eigenvektor *normiert* (d.h., vom Betrag 1).

Beispiel 5.7.5 Für $A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ erhält man aus der charakteristischen Gleichung

$$\det(A - sE) = \begin{vmatrix} 5-s & 3 \\ 3 & 5-s \end{vmatrix} = (5-s)^2 - 9 = 16 - 10s + s^2 = 0$$

die beiden Eigenwerte

$$s_{1/2} = \frac{1}{2}(10 \pm \sqrt{100 - 64}) = \begin{cases} 8 \\ 2 \end{cases}.$$

Die allgemeine Lösung des Systems

$$\begin{pmatrix} 5-8 & 3 \\ 3 & 5-8 \end{pmatrix}\begin{pmatrix} y_{11} \\ y_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ergibt sich nach dem Gauß-Jordan-Algorithmus

$$\begin{array}{cc} 5-8 & 3 \\ 3 & 5-8 \\ \hline -3 & 3 \\ 3 & -3 \\ \hline 1 & -1 \\ 0 & 0 \end{array}$$

zu $\vec{y}_1 = \begin{pmatrix} y_{21} \\ y_{21} \end{pmatrix}$ mit $y_{21} \neq 0$, normiert: $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

Probe: $\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} y_{21} \\ y_{21} \end{pmatrix} = 8 \cdot \begin{pmatrix} y_{21} \\ y_{21} \end{pmatrix}$.

Einen Eigenvektor $\vec{y}_2$ zu $s_2 = 2$ erhält man durch Lösen des Systems

$$\begin{pmatrix} 5-2 & 3 \\ 3 & 5-2 \end{pmatrix}\begin{pmatrix} y_{12} \\ y_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}:$$

$\vec{y}_2 = \begin{pmatrix} -y_{22} \\ y_{22} \end{pmatrix}$ mit $y_{22} \neq 0$, normiert: $\begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

Probe: $\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} -y_{22} \\ y_{22} \end{pmatrix} = 2 \cdot \begin{pmatrix} -y_{22} \\ y_{22} \end{pmatrix}$.

Beispiel 5.7.6 Im allgemeinen kann eine reelle Matrix A sehr wohl komplexe Eigenwerte haben. Z.B. hat $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ zwei konjugiert-komplexe Eigenwerte: Aus der charakteristischen Gleichung

$$\begin{pmatrix} 1-s & 1 \\ -1 & 1-s \end{pmatrix} = (1-s)^2 + 1 = s^2 - 2s + 2 = 0$$

folgt $s_{1/2} = 1 \pm j$. Wegen

$$\begin{array}{cc} -j & 1 \\ -1 & -j \\ \hline 1 & j \\ 0 & 0 \end{array} \qquad \begin{array}{cc} j & 1 \\ -1 & j \\ \hline 1 & -j \\ 0 & 0 \end{array}$$

sind $\begin{pmatrix} -j \\ 1 \end{pmatrix}$ und $\begin{pmatrix} j \\ 1 \end{pmatrix}$ die beiden linear unabhängigen Eigenvektoren:

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -j \\ 1 \end{pmatrix} = (1+j) \cdot \begin{pmatrix} -j \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} j \\ 1 \end{pmatrix} = (1-j) \cdot \begin{pmatrix} j \\ 1 \end{pmatrix}.$$

Es stellt sich heraus, daß Matrizen, deren Eigenwerte paarweise verschieden sind, n linear unabhängige Eigenvektoren haben.

In gewisser Hinsicht erweisen sich nun Matrizen A, zu denen es n linear unabhängige Eigenvektoren gibt, als besonders einfach:

> Gibt es n linear unabhängige Eigenvektoren $\vec{y}_1, \ldots, \vec{y}_n$ zu den Eigenwerten $s_1, \ldots, s_n$ der Matrix A, so wird durch die Matrix B, deren Spalten $\vec{y}_1, \ldots, \vec{y}_n$ sind, das Koordinaten-System so transformiert, daß
> $$B^{-1}AB = \begin{pmatrix} s_1 & 0 & \ldots & \ldots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \ldots & s_i & \ldots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & s_n \end{pmatrix}$$
> eine *Diagonalmatrix* mit den Eigenwerten in der Hauptdiagonale ist.

Beweis: $\vec{e}_j$ sei die Einheitsspalte, die an der j-ten Stelle 1, sonst lauter Nullen enthält. Daher ist $B\vec{e}_j$ die j-te Spalte von B, also gleich $\vec{y}_j$. Da wegen der linearen Unabhängigkeit der $\vec{y}_j$ $(j = 1, \ldots, n)$ die Matrix B invertierbar ist, gilt umgekehrt $\vec{e}_j = B^{-1}\vec{y}_j$. Somit ist

$$(B^{-1}AB)\vec{e}_j = B^{-1}A\vec{y}_j = B^{-1}s_j\vec{y}_j = s_jB^{-1}\vec{y}_j = s_j\vec{e}_j$$

und $B^{-1}AB$ enthält in der j-ten Spalte s_j an der j-ten Stelle, sonst Nullen, ist also die angegebene Diagonalmatrix.
Leider muß es, wenn die Eigenwerte nicht sämtlich verschieden sind, zu einer Matrix A nun nicht immer n linear unabhängige Eigenvektoren geben. Jedoch:

> Für Matrizen A mit der Eigenschaft
> $$A \cdot A^T = A^T \cdot A \quad (\textit{normale}\text{ Matrizen})$$
> existieren n linear unabhängige Eigenvektoren.
> Im Fall reeller Eigenwerte sind sie sogar paarweise orthogonal und können normiert werden, so daß die Matrix B orthogonal gewählt werden kann.

Beispiel 5.7.7 Die Matrix $\begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 1 \end{pmatrix}$ ist normal, weil symmetrisch. Sie hat die charakteristische Gleichung

$$\begin{vmatrix} 1-s & 2 & 0 \\ 2 & -1-s & 2 \\ 0 & 2 & 1-s \end{vmatrix} = (1-s)^2(-1-s) - 4(1-s) - 4(1-s)$$
$$= (1-s)(s^2 - 9) = (1-s)(s-3)(s+3) = 0$$

und daher die (reellen) Eigenwerte $s_1 = 1,\ s_2 = 3,\ s_3 = -3$.

Eigenvektor zu $s_1 = 0$ ist (bis auf einen Faktor $\neq 0$) $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, normiert: $\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

Eigenvektor zu $s_2 = 3$ ist (bis auf einen Faktor $\neq 0$) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, normiert: $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Eigenvektor zu $s_3 = -3$ ist (bis auf einen Faktor $\neq 0$) $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, normiert: $\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.

Diese Eigenvektoren sind paarweise orthogonal:

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0, \quad \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0.$$

Beispiel 5.7.8 Die Matrix $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ ist normal:

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

Laut Beispiel 5.7.6 hat sie zu den komplexen Eigenwerten $s_1 = 1 + j$, $s_2 = 1 - j$ die Eigenvektoren $\begin{pmatrix} -j \\ 1 \end{pmatrix}$ und $\begin{pmatrix} j \\ 1 \end{pmatrix}$. Die Eigenvektoren sind linear unabhängige Vektoren des $\mathbb{C}^2$, nicht mehr des $\mathbb{R}^2$.

Beispiel 5.7.9 Die Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ ist nicht mehr normal, wie man leicht nachrechnet. Sie hat wegen $\begin{vmatrix} 1-s & 1 \\ 0 & 2-s \end{vmatrix} = (s-1)(s-2) = 0$ die beiden Eigenwerte 1 und 2. Eigenvektor zu 1 ist $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, Eigenvektor zu 2 ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
Diese Vektoren sind im $\mathbb{R}^2$ linear unabhängig, aber nicht mehr orthogonal.

Beispiel 5.7.10 Die Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist nicht mehr normal und hat wegen $\begin{vmatrix} 1-s & 1 \\ 0 & 1-s \end{vmatrix} = (s-1)^2 = 0$ den einzigen (zweifachen) Eigenwert 1. Dazu gibt es (bis auf Vielfache) nur einen Eigenvektor, nämlich $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Es gibt also zu dieser Matrix keine 2 linear unabhängigen Eigenvektoren mehr.

> Ein wichtiger Spezialfall normaler Matrizen sind die *symmetrischen* Matrizen mit der Eigenschaft $A = A^T$.
> Für sie sind sämtliche Eigenwerte reell und es gibt stets n orthogonale und damit linear unabhängige Eigenvektoren, die zudem normiert werden können.

Wir haben oben gesehen, daß wir die Matrizen A quadratischer Formen stets symmetrisch wählen können. Deshalb gilt:

Zu jeder quadratischen Form $Q(\vec{x}) = \vec{x}^T A\, \vec{x}$, für die $A = A^T$ gewählt ist, gibt es eine Koordinaten-Transformation

$$\vec{x} = B\overline{\vec{x}}\,,$$

so daß

$$Q(\vec{x}) = \overline{\vec{x}}^T (B^T A\, B)\overline{\vec{x}} = \sum_{i=1}^{n} s_i \overline{x}_i^2$$

ist.
$s_1, \ldots, s_n$ sind dabei die (reellen) Eigenwerte der Matrix A; B hat die orthogonalen und normierten Eigenvektoren der Matrix A als Spalten.

Beispiel 5.7.11 Der Kegelschnitt $3x^2 + 5y^2 - 2\sqrt{3}xy + (1+\sqrt{3})x + (1-\sqrt{3})y - \frac{16}{3} = 0$ soll auf Hauptachsen tranformiert werden.
Die symmetrische Matrix der quadratischen Form $Q = 3x^2 + 5y^2 - 2\sqrt{3}xy$ ist

$$\begin{pmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 5 \end{pmatrix}.$$

Sie hat die charakteristische Gleichung

$$\begin{vmatrix} 3-s & -\sqrt{3} \\ -\sqrt{3} & 5-s \end{vmatrix} = s^2 - 8s + 12 = 0$$

und daher die Eigenwerte $s_1 = 6$, $s_2 = 2$. Normierte Eigenvektoren sind $\frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix}$ und $\frac{1}{2}\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}$.
Die orthogonale Koordinaten-Transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix} \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \overline{x} + \sqrt{3}\overline{y} \\ -\sqrt{3}\overline{x} + \overline{y} \end{pmatrix}$$

ist eine Drehung um den Winkel $\delta = -\frac{\pi}{3}$, denn $\cos(-\frac{\pi}{3}) = \frac{1}{2}$, $\sin(-\frac{\pi}{3}) = -\frac{\sqrt{3}}{2}$.
Man erhält $Q = 6\overline{x}^2 + 2\overline{y}^2$ und als Kegelschnittsgleichung im neuen System:

$$6\overline{x}^2 + 2\overline{y}^2 + 2\overline{x} + 2\overline{y} - \frac{16}{3} = 0 \quad \text{oder} \quad \left(\overline{x} + \frac{1}{6}\right)^2 + \frac{1}{3}\left(\overline{y} + \frac{1}{2}\right)^2 = 1\,.$$

Es handelt sich also dort um eine vertikale Ellipse mit den Achsen $\sqrt{3}$ und 1 und dem Mittelpunkt $(-\frac{1}{6}; -\frac{1}{2})$.

Aufgaben

1 (a) Man ermittle einen zu den Vektoren $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ orthogonalen Vektor, normiere die drei Vektoren und gewinne so eine orthogonale Matrix A!

(b) Wie lautet die zu A inverse Matrix?

(c) Man bilde mittels A den Vektor $\vec{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ ab und zeige, daß die Vektoren $\vec{x}$ und $A\vec{x}$ denselben Betrag haben!