Lineare Algebra B

Sommersemester 2002

W. Ebeling

Institut für Mathematik
Universität Hannover
Postfach 6009
30060 Hannover
E-mail: ebeling@math.uni-hannover.de

1 Determinanten

Wir wollen nun den Begriff der Determinante einer Matrix einführen. Dazu betrachten wir zunächst Permutationen.

Definition Eine *Permutation* σ der Menge $\{1, 2, \ldots, n\}$ ist eine Anordnung $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ dieser Zahlen ohne Auslassungen oder Wiederholungen.

Beispiel 1.1 Die Permutationen der Menge $\{1, 2\}$ sind

$$(1, 2) \quad (2, 1).$$

Die Permutationen der Menge $\{1, 2, 3\}$ sind

$$\begin{array}{ccc} (1,2,3) & (2,3,1) & (3,1,2) \\ (1,3,2) & (2,1,3) & (3,2,1). \end{array}$$

Beispiel 1.2 Wir wollen die Permutationen der Menge $\{1, 2, 3, 4\}$ bestimmen. Dazu kann man wie folgt vorgehen.

Zuerst müssen wir die Zahl $\sigma(1)$ wählen, die an erster Stelle steht. Diese können wir beliebig aus den vier Zahlen $1, 2, 3, 4$ wählen (4 Wahlmöglichkeiten).

Nun müssen wir die Zahl $\sigma(2)$ wählen, die an zweiter Stelle steht. Dies darf nicht die schon gewählte Zahl $\sigma(1)$ sein. Also haben wir noch 3 Wahlmöglichkeiten.

Nun müssen wir die Zahl $\sigma(3)$ aus den verbliebenen 2 Zahlen wählen (2 Wahlmöglichkeiten).

Es bleibt nur noch eine Zahl übrig. Die muss an vierter Stelle stehen.

Insgesamt haben wir also $4 \cdot 3 \cdot 2 \cdot 1 = 24$ Wahlmöglichkeiten. Es gibt also 24 Permutationen der Menge $\{1, 2, 3, 4\}$.

Die Wahlmöglichkeiten können wir uns auch an einem Baum veranschaulichen

$$\begin{array}{lcccccccccccc}
\sigma(1) = & & 1 & & & 2 & & & 3 & & & 4 & \\
 & / & | & \backslash & / & | & \backslash & / & | & \backslash & / & | & \backslash \\
\sigma(2) = & 2 & 3 & 4 & 1 & 3 & 4 & 1 & 2 & 4 & 1 & 2 & 3 \\
 & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash & /\backslash \\
\sigma(3) = & 34 & 24 & 23 & 34 & 14 & 13 & 24 & 14 & 12 & 23 & 13 & 12 \\
 & || & || & || & || & || & || & || & || & || & || & || & || \\
\sigma(4) = & 43 & 42 & 32 & 43 & 41 & 31 & 42 & 41 & 21 & 32 & 31 & 21
\end{array}$$

Wie wir in dem Beispiel gesehen haben, gibt es genau $n \cdot (n-1) \cdots 2 \cdot 1 = n!$ Permutationen der Menge $\{1, 2, \ldots, n\}$.

Definition Sind σ und τ Permutationen der Menge $\{1, 2, \dots, n\}$, so ist ihre *Hintereinanderausführung* $\tau \circ \sigma$ definiert durch

$$\tau \circ \sigma := (\tau(\sigma(1)), \tau(\sigma(2)), \dots, \tau(\sigma(n))).$$

Beispiel 1.3 Ist $\sigma = (2, 1, 4, 3)$ und $\tau = (4, 1, 3, 2)$, so gilt $\tau \circ \sigma = (1, 4, 2, 3)$. Dies sieht man so ein: Wir fassen die Permutationen als Vertauschungen von $(1, 2, 3, 4)$ auf und schreiben deshalb

$$\sigma = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{array}, \quad \tau = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{array}.$$

Schreiben wir nun die entsprechenden Zahlen für $\tau \circ \sigma$ untereinander, so erhalten wir

$$\tau \circ \sigma = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 1 & 4 & 2 & 3 \end{array}.$$

Satz 1.1 *Die Menge der Permutationen der Menge $\{1, 2, \dots, n\}$ bildet mit der Hintereinanderschaltung als Verknüpfung eine Gruppe.*

Beweis. Das Assoziativgesetz ist leicht nachzuweisen.

Das neutrale Element ist die Permutation $(1, 2, \dots, n)$.

Das inverse Element zu einer Permuation $\sigma = (\sigma(1), \sigma(2), \dots, \sigma(n))$ ist die Permutation $(j_1, j_2, \dots, j_n)$ mit $\sigma(j_1) = 1$, $\sigma(j_2) = 2$, ..., $\sigma(j_n) = n$. □

Definition Eine *Inversion* der Permutation σ der Menge $\{1, 2, \dots, n\}$ ist ein Zahlenpaar (i, j) mit $1 \leq i < j \leq n$ aber $\sigma(i) > \sigma(j)$, d.h. ein Zahlenpaar, bei der die Reihenfolge bei der Permutation vertauscht wird.

Es sei $m(\sigma)$ die Anzahl der Inversionen der Permutation σ. Diese Anzahl kann man durch den folgenden Algorithmus bestimmen:

(1) $m :=$ Anzahl der Zahlen, die kleiner als $\sigma(1)$ sind.

(2) $i := 2$.

(3) $m := m+$ Anzahl der Zahlen, die kleiner als $\sigma(i)$ sind, aber in der Permutation erst nach $\sigma(i)$ stehen.

(4) $i := i + 1$. Falls $i < n$ gehe nach (3), andernfalls $m(\sigma) := m$.

Beispiel 1.4 Die Anzahl der Inversionen von $(3, 2, 4, 1)$ ist $2 + 1 + 1$.

Definition Eine Permutation heißt *gerade*, wenn sie eine gerade Anzahl von Inversionen enthält, andernfalls heißt sie *ungerade*. Das *Signum* der Permutation σ, in Zeichen $\operatorname{sign}(\sigma)$, ist $+$, falls die Permutation gerade ist, und $-$, falls die Permutation ungerade ist.

Beispiel 1.5 In der folgenden Tabelle werden die Permutationen von $\{1,2,3\}$ einschließlich der Anzahl der Inversionen und des Signums aufgelistet:

Permutationen	Anzahl der Inversionen	Signum
$(1,2,3)$	0	$+$
$(1,3,2)$	1	$-$
$(2,3,1)$	2	$+$
$(2,1,3)$	1	$-$
$(3,1,2)$	2	$+$
$(3,2,1)$	3	$-$

Im Folgenden sei nun A eine $n \times n$-Matrix.

Definition Ein *elementares Produkt* von Einträgen aus A ist ein Produkt von n Einträgen aus A, wobei aus jeder Zeile und jeder Spalte genau ein Eintrag stammt. Ein elementares Produkt von Einträgen aus A ist also ein Ausdruck

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)},$$

wobei σ eine Permutation der Menge $\{1,2,\ldots,n\}$ ist.

Beispiel 1.6 Die elementaren Produkte der Einträge der Matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

sind

$$a_{11}a_{22} \text{ und } a_{12}a_{21}.$$

Die elementaren Produkte der Einträge der Matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

sind

$$\begin{array}{ccc} a_{11}a_{22}a_{33} & a_{12}a_{23}a_{31} & a_{13}a_{21}a_{32} \\ a_{13}a_{22}a_{31} & a_{11}a_{23}a_{32} & a_{12}a_{21}a_{33}. \end{array}$$

Definition Die *Determinante* der Matrix A, in Zeichen $\det A$, ist definiert durch die Formel

$$\det A = \sum_{\sigma} \operatorname{sign}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)},$$

wobei sich die Summe über alle Permutationen der Menge $\{1,2,\ldots,n\}$ erstreckt, d.h.

(1) bilde alle elementaren Produkte der Einträge der Matrix A,

(2) versehe sie mit dem Signum der entsprechenden Permutation,

(3) summiere alles auf.

Warnung Man beachte, dass die Determinante nur für eine *quadratische* Matrix definiert ist!

Beispiel 1.7 Mit Hilfe von Beispiel 1.6 erhalten wir

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{array}$$

Die Berechnung der Determinante einer 3×3-Matrix merkt man sich mit Hilfe der folgenden Regel:

Regel von Sarrus: *Man schreibe die ersten beiden Spalten der Matrix noch einmal hinter die Matrix und multipliziere entlang der angedeuteten Pfeile, wobei die elementaren Produkte längs der nach oben gerichteten Pfeile mit dem Vorzeichen $-$ zu versehen sind:*

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

Beispiel 1.8 Die Determinante der Matrix

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

lautet

$$\det A = 2 + 0 + (-6) - (-4) - 1 - 0 = -1.$$

Warnung Die Regel von Sarrus funktioniert nur für 3×3-Matrizen, *nicht* für größere Matrizen!

Die Berechnung von Determinanten mit Hilfe der Definition erfordert im Allgemeinen einen großen Rechenaufwand. Um eine Determinante einer 4×4-Matrix zu berechnen, braucht man schon $4! = 24$ Terme, bei einer 10×10-Matrix $10! = 3\,628\,800$. Selbst für schnelle Computer wächst die Rechenzeit mit der Größe der Matrix sehr schnell ins Unermessliche.

Notation Statt $\det A$ schreibt man auch $|A|$, also z.B.

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Wir wollen uns nun mit Eigenschaften von Determinanten beschäftigen.

Satz 1.2 *Es sei A eine $n \times n$-Matrix. Dann gilt*

$$\det A^T = \det A.$$

Beweis. Es gilt

$$\begin{aligned} \det A^T &= \sum_\sigma \operatorname{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ &= \sum_\sigma \operatorname{sign}(\sigma) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}, \end{aligned}$$

Hierbei durchläuft σ alle Permutationen der Menge $\{1, 2, \ldots, n\}$. Da es aber zu jeder solchen Permutation σ nach Satz 1.1 genau eine inverse Permutation σ^{-1} gibt, können wir die Summe genauso gut über σ^{-1} laufen lassen und σ^{-1} durchläuft alle Permutationen der Menge $\{1, 2, \ldots, n\}$. Es gilt $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$. Setzen wir also $\tau = \sigma^{-1}$, so folgt

$$\det A^T = \sum_\tau \operatorname{sign}(\tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)},$$

wobei sich die Summe über alle Permutationen τ erstreckt. Daraus folgt die Behauptung. □

Aus diesem Satz folgt, dass wir in allen folgenden Sätzen über Determinanten, die sich auf die Zeilen der Matrix beziehen, das Wort "Zeile" durch "Spalte" ersetzen können, wir also einen entsprechenden Satz für Spalten der Matrix erhalten.

Satz 1.3 *Es sei A eine $n \times n$-Matrix.*

(a) *Ist B die Matrix, die aus A durch Vertauschung zweier Zeilen (oder Spalten) entsteht, so ist* $\det B = -\det A$.

(b) *Kommt in A eine Zeile (oder Spalte) doppelt vor, so gilt* $\det A = 0$.

(c) *Ist B die Matrix, die durch Multiplikation einer Zeile (oder Spalte) von A mit einer Konstanten λ entsteht, so ist* $\det B = \lambda \det A$.

(d) *Es seien* $\mathbf{a}_1, \dots, \mathbf{a_n}$ *die Zeilenvektoren von* A *und* $\mathbf{b} \in \mathbb{R}^n$*. Dann gilt*

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k + \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

(Eine entsprechende Aussage gilt auch für Spaltenvektoren.)

Beweis. Wir beweisen nur (d). Die anderen Gleichungen beweist man analog. Es gilt

$$\begin{aligned} \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k + \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} &= \sum_\sigma \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots (a_{k\sigma(k)} + b_{\sigma(k)}) \cdots a_{n\sigma(n)} \\ &= \sum_\sigma \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{n\sigma(n)} \\ &\quad + \sum_\sigma \operatorname{sign}(\sigma) a_{k\sigma(1)} \cdots b_{\sigma(k)} \cdots a_{n\sigma(n)} \\ &= \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + \det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}. \end{aligned}$$

□

Aus diesem Satz können wir Folgerungen ziehen:

Satz 1.4 *Es sei* A *eine* $n \times n$*-Matrix.*

(a) *Ist* B *die Matrix, die aus* A *durch Addition eines Vielfachen einer Zeile (oder Spalte) zu einer anderen entsteht, so gilt* $\det B = \det A$.

(b) *Enthält* A *eine Nullzeile (oder -spalte), so ist* $\det A = 0$.

(c) *Enthält* A *zwei linear abhängige Zeilen (oder Spalten), so ist* $\det A = 0$.

Beweis. (a) Die Matrix B entstehe aus A durch Addition des λ-fachen der k-ten Zeile zur l-ten Zeile. Nach Satz 1.3 gilt dann

$$\det B = \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a_k} \\ \vdots \\ \mathbf{a}_l + \lambda \mathbf{a}_k \\ \vdots \\ \mathbf{a}_n \end{vmatrix} = \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_l \\ \vdots \\ \mathbf{a}_n \end{vmatrix} + \lambda \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_n \end{vmatrix} = \det A$$

(b) Die l-te Zeile der Matrix A sei eine Nullzeile. Addieren wir eine beliebige andere Zeile, etwa die k-te Zeile, $k \neq l$, zu dieser Zeile, so enthält die neue Matrix B eine Zeile doppelt. Nach Satz 1.3 (b) gilt $\det B = 0$.

(c) Enthält die Matrix A zwei linear abhängige Zeilen (oder Spalten), so kann man durch Addition eines geeigneten Vielfachen der einen Zeile (oder Spalte) zur anderen eine Nullzeile (oder Nullspalte) erzeugen. □

Wir berechnen nun die Determinanten der in LAA § 6 betrachteten Elementarmatrizen.

Satz 1.5 (1) $\det F_{ik} = -1$.

(2) $\det F_{ik}^{\lambda} = 1$.

(3) $\det F_i^{\lambda} = \lambda$.

Beweis. Die Determinante der Einheitsmatrix ist 1, wie man leicht sieht. Man erhält die Elementarmatrizen durch elementare Zeilenumformungen aus der Einheitsmatrix. Damit folgt die Behauptung aus Satz 1.3 und Satz 1.4. □

Wir wollen nun untersuchen, wie sich die Determinante bei Matrixoperationen verhält.

Satz 1.6 *Es sei A eine $n \times n$-Matrix und $\lambda \in \mathbb{R}$. Dann gilt*

$$\det(\lambda A) = \lambda^n \det A.$$

Beweis. Jede der n Zeilen der Matrix λA enthält den gemeinsamen Faktor λ. Nach Satz 1.3 (c) kann er aus jeder Zeile vor die Determinante gezogen werden. Insgesamt ergibt sich also der Faktor λ^n. □

Warnung Es gilt im Allgemeinen *nicht* $\det(A + B) = \det A + \det B$!

Beispiel 1.9 Es sei

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A + B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Dann gilt $\det A = \det B = 1$, $\det A + \det B = 2$, aber $\det(A + B) = 4$.

Anders verhält es sich aber mit dem Produkt von Matrizen:

Satz 1.7 (Determinantenmultiplikationssatz) *Es seien A und B zwei $n \times n$-Matrizen. Dann gilt*

$$\det AB = \det A \det B.$$

Beweis. (a) Wir zeigen den Satz zunächst für den Spezialfall, dass A eine Elementarmatrix ist. Es sei etwa $A = F_{ik}$. Dann entsteht $AB = F_{ik}B$ aus B durch Vertauschung der i-ten mit der k-ten Zeile. Nach Satz 1.3 gilt $\det AB = -\det B$. Nach Satz 1.5 gilt $\det F_{ik} = -1$, also

$$\det F_{ik}B = \det F_{ik} \det B.$$

Analog zeigt man die Behauptung für die anderen Elementarmatrizen.

(b) Wir zeigen den Satz für den Spezialfall, dass A eine Nullzeile enthält. In diesem Fall enthält auch AB eine Nullzeile. Nach Satz 1.4 gilt dann $\det A = \det AB = 0$, also auch

$$\det AB = \det A \det B.$$

(c) Wir zeigen den Satz nun für den Fall, dass A invertierbar ist. Nach LAA, Satz 6.6, lässt sich A dann als Produkt von Elementarmatrizen schreiben. Damit folgt die Behauptung durch wiederholte Anwendung von (a).

(d) Schließlich betrachten wir den Fall, dass A nicht invertierbar ist. Nach LAA, Satz 12.5, hat A dann nicht maximalen Rang. Das bedeutet, dass die Zeilen von A linear abhängig sind. Durch geeignete Zeilenumformungen kann man dann eine Nullzeile erzeugen. Den Zeilenumformungen entspricht aber die Multiplikation der Matrix A mit Elementarmatrizen. Also folgt die Behauptung aus Teil (a) und (b). □

Korollar 1.1 *Eine $n \times n$-Matrix A ist genau dann invertierbar, wenn* $\det A \neq 0$ *gilt. In diesem Fall gilt*

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

Beweis. Ist A invertierbar, so gilt $AA^{-1} = E$. Nach Satz 1.7 folgt

$$\det A \det A^{-1} = \det E = 1.$$

Also folgt $\det A \neq 0$ und wir können die Gleichung durch $\det A$ dividieren.

Ist A nicht invertierbar, so sind die Zeilen von A nach LAA, Satz 12.5, linear abhängig. Durch geeignete Zeilenumformungen kann man dann eine Nullzeile erzeugen. Aus Satz 1.3 und Satz 1.4 folgt dann $\det A = 0$. $\square$

Wir behandeln nun ein wichtiges Verfahren zur Berechnung von Determinanten. Dazu sehen wir uns noch einmal die Determinante einer 3×3-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

an. In Beispiel 1.7 hatten wir gesehen

$$\begin{aligned} \det A &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Dies können wir auch wie folgt schreiben:

$$\det A = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}).$$

Die Ausdrücke in den Klammern sind aber Determinanten von Untermatrizen.

Definition Es sei A eine $n \times n$-Matrix. Die Matrix A_{ij} entstehe aus A durch Streichen der i-ten Zeile und j-ten Spalte von A, d.h.

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i1} & \cdots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \cdots & a_{in} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}$$

Mit dieser Definition können wir also schreiben

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31}.$$

Allgemein gilt der folgende Satz

Satz 1.8 (Laplacescher Entwicklungssatz) *Es sei A eine $n \times n$-Matrix, $1 \leq k \leq n$. Dann gilt*

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} \quad \textit{(Entwicklung nach der k-ten Spalte)}$$

und

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj} \quad \textit{(Entwicklung nach der k-ten Zeile).}$$

Diesen Satz kann man ähnlich beweisen, wie wir es für den Fall $n = 3$ (Entwicklung nach der ersten Spalte) ausgeführt haben. Der Ausdruck, der die Determinante definiert, ist geeignet umzuformen. Da dies aber sehr aufwendig ist, verzichten wir auf einen allgemeinen Beweis.

Bemerkung 1.1 Die Vorzeichen in Satz 1.8 sind gemäß eines Schachbrettmusters verteilt:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Beispiel 1.10 Wir entwickeln die Determinante der Matrix

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

nach der ersten Spalte:

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

Wir betrachten nun Anwendungen von Determinanten.

Satz 1.9 (Cramersche Regel) *Es sei A eine invertierbare $n \times n$-Matrix und $\mathbf{b} \in \mathbb{R}^n$. Mit A_i bezeichnen wir die Matrix, die dadurch entsteht, dass wir die i-te Spalte von A durch den Spaltenvektor $\mathbf{b}$ ersetzen. Dann werden die Komponenten x_i der Lösung $\mathbf{x}$ des linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ durch die Formel*

$$x_i = \frac{\det A_i}{\det A}, \quad 1 \le i \le n,$$

gegeben.

Beweis. Wir bezeichnen die Spaltenvektoren der Matrix A mit $\mathbf{s}_1, \ldots, \mathbf{s}_n$. Dann gilt

$$\begin{aligned} \det A_1 &= \det(\mathbf{b}, \mathbf{s}_2, \ldots, \mathbf{s}_n) \\ &= \det(x_1\mathbf{s}_1 + \cdots + x_n\mathbf{s}_n, \mathbf{s}_2, \ldots, \mathbf{s}_n) \\ &= x_1 \det(\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n) + \cdots + x_n \det(\mathbf{s}_n, \mathbf{s}_2, \ldots, \mathbf{s}_n) \\ &= x_1 \det A. \end{aligned}$$

Der Beweis für die anderen Komponenten geht analog. □

Die Cramersche Regel ist für praktische Berechnungen nicht sehr geeignet, wohl aber für theoretische Überlegungen.

Beispiel 1.11 Wir betrachten das folgende Gleichungssystem

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

Dann ergibt die Cramersche Regel:

$$x_1 = \frac{\begin{vmatrix} 4 & -1 \\ -3 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ -3 & 2 \end{vmatrix}} = 5, \quad x_2 = \frac{\begin{vmatrix} 2 & 4 \\ -3 & -3 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ -3 & 2 \end{vmatrix}} = 6.$$

Wir wenden nun die Cramersche Regel auf die Bestimmung der inversen Matrix einer invertierbaren Matrix an. Es sei A eine invertierbare $n \times n$-Matrix. Die inverse Matrix bestimmt sich durch das lineare Gleichungssystem

$$AX = E$$

oder

$$A\mathbf{x}_j = \mathbf{e}_j, \quad j = 1, \ldots, n,$$

wobei wir mit $\mathbf{x}_1, \ldots, \mathbf{x}_n$ die Spalten der Matrix X bezeichnen. Nach der Cramerschen Regel gilt

$$x_{ij} = \frac{1}{\det A} \det A_i,$$

wobei A_i die Matrix ist, die aus A entsteht, indem man die i-te Spalte durch den Vektor $\mathbf{e}_j$ ersetzt. Nach dem Laplaceschen Entwicklungssatz gilt

$$x_{ij} = \frac{1}{\det A}(-1)^{i+j} \det A_{ji}.$$

Definition Die Matrix $A^* := (a^*_{ij})$ mit $a^*_{ij} = (-1)^{i+j} \det A_{ji}$ heißt die *adjungierte Matrix* von A.

Wir haben damit bewiesen:

Satz 1.10 *Für eine invertierbare Matrix A ist*

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

Bemerkung 1.2 Man beachte die Vertauschung der Indizes i und j bei der Definition der adjungierten Matrix.

Beispiel 1.12 Die adjungierte Matrix zu der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix}$$

lautet

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

Wegen $\det A = 1$ ist $A^{-1} = A^*$.

2 Lineare Abbildungen

Wir wollen nun lineare Abbildungen betrachten. Dazu müssen wir zunächst den Begriff einer Abbildung einführen.

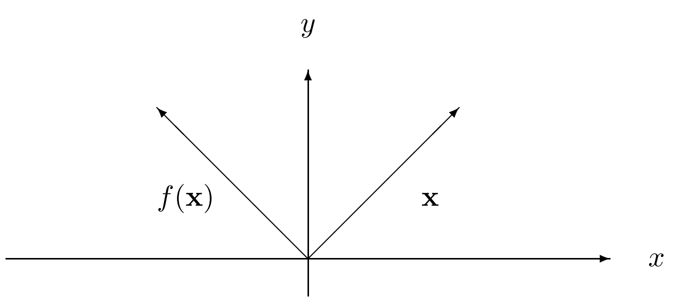
Definition Es seien X und Y Mengen. Unter einer *Abbildung* von X nach Y versteht man eine Vorschrift f, die jedem $x \in X$ genau ein $f(x) \in Y$ zuordnet. Man schreibt dafür

$$\begin{array}{rccc} f: & X & \longrightarrow & Y \\ & x & \longmapsto & f(x) \end{array}.$$

Man nennt X die *Definitionsmenge* und Y die *Wertemenge* der Abbildung f. Das Element $f(x)$ bezeichnen wir auch als das *Bild* von x unter der Abbildung f. Zwei Abbildungen $f : X \to Y$ und $g : X \to Y$ heißen *gleich*, in Zeichen $f = g$, genau dann, wenn $f(x) = g(x)$ für alle $x \in X$ gilt.

Beispiel 2.1 Beispiele für Abbildungen sind Funktionen $f : \mathbb{R} \to \mathbb{R}$, also zum Beispiel die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$.

Beispiel 2.2 Wir betrachten nun Abbildungen der Ebene. Es sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Abbildung, die jedem Vektor $\mathbf{x} \in \mathbb{R}^2$ sein Spiegelbild bezüglich der y-Achse zuordnet.



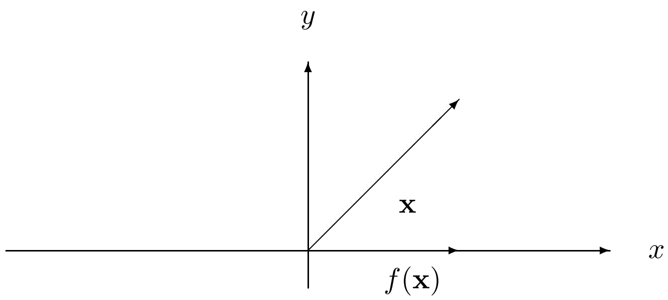
Das Bild $f(\mathbf{x})$ ist wieder ein Vektor. Wir bezeichnen die Komponenten dieses Vektors mit $f_1(\mathbf{x})$ und $f_2(\mathbf{x})$. Es gilt

$$\begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}.$$

Dies können wir auch so ausdrücken:

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Beispiel 2.3 Wir betrachten die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$, die jedem Vektor $\mathbf{x}$ seine Orthogonalprojektion auf die x-Achse zuordnet.



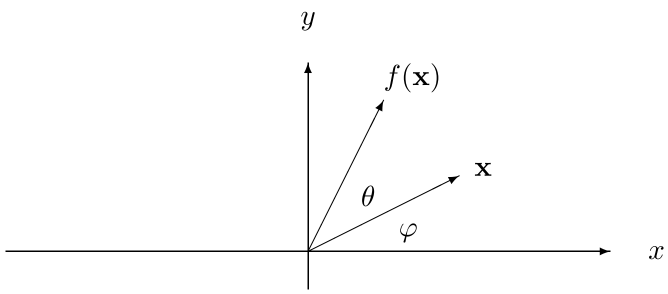
Es gilt

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}.$$

Mit Hilfe einer Matrix können wir dies so schreiben:

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Beispiel 2.4 Wir betrachten nun die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$, die einen Vektor $\mathbf{x}$ um den Winkel θ dreht. Um eine Beschreibung für $f_1(x,y)$ und $f_2(x,y)$ abzuleiten, betrachten wir eine Drehung um einen positiven Winkel θ. Es sei φ der Winkel zwischen dem Vektor $\mathbf{x}$ und der positiven x-Achse und r die Länge von $\mathbf{x}$.



Dann gilt

$$x = r\cos\varphi, \quad y = r\sin\varphi$$

und

$$f_1(x,y) = r\cos(\varphi+\theta), \quad f_2(x,y) = r\sin(\varphi+\theta).$$

Durch Anwendung der Additionstheoreme von sin und cos ergibt sich hieraus

$$\begin{aligned} f_1(x,y) &= r\cos\theta\cos\varphi - r\sin\theta\sin\varphi \\ f_2(x,y) &= r\sin\theta\cos\varphi + r\cos\theta\sin\varphi, \end{aligned}$$

und schließlich

$$\begin{aligned} f_1(x,y) &= (\cos\theta)x - (\sin\theta)y \\ f_2(x,y) &= (\sin\theta)x + (\cos\theta)y. \end{aligned}$$

In Matrizenschreibweise lautet dies

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Beispiel 2.5 Nun betrachten wir auch Abbildungen des Raumes. Es sei $f : \mathbb{R}^3 \to \mathbb{R}^3$ eine Drehung um die x-Achse um den Winkel θ. Wie im vorigen Beispiel leitet man her, dass diese Abbildung durch die folgende Vorschrift gegeben wird:

$$\begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Beispiel 2.6 Die Beispiele von Abbildungen der Ebene und des Raumes sind Spezialfälle der folgenden Konstruktion. Einer $m \times n$-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

kann man wie folgt eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ zuordnen: Wir definieren

$$f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix},$$

oder anders ausgedrückt

$$f(\mathbf{x}) = A\mathbf{x}.$$

Nach den Rechenregeln für Matrizen gilt

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}, \quad A(\lambda\mathbf{x}) = \lambda A\mathbf{x}.$$

Also hat diese Abbildung die Eigenschaften

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}), \quad f(\lambda\mathbf{x}) = \lambda f(\mathbf{x}).$$

Eine Abbildung mit den Eigenschaften des letzten Beispiels heißt *linear*. Diesen Begriff wollen wir nun ganz allgemein definieren.

Definition Es seien V und W K-Vektorräume. Eine Abbildung $f : V \to W$ heißt *linear* genau dann, wenn gilt

(L1) $f(\mathbf{v}+\mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$ für alle $\mathbf{v}, \mathbf{w} \in V$,

(L2) $f(\lambda\mathbf{v}) = \lambda f(\mathbf{v})$ für alle $\mathbf{v} \in V$ und $\lambda \in K$.

Bemerkung 2.1 Die beiden Bedingungen (L1) und (L2) kann man auch zu der Bedingung

(L) $f(\lambda\mathbf{v} + \mu\mathbf{w}) = \lambda f(\mathbf{v}) + \mu f(\mathbf{w})$ für alle $\mathbf{v}, \mathbf{w} \in V$ und $\lambda, \mu \in K$.

zusammenfassen. Man überlegt sich leicht, dass die beiden Bedingungen (L1) und (L2) zusammen äquivalent zu der Bedingung (L) sind.

Beispiel 2.7 Die Abbildungen von Beispiel 2.6 und damit aus allen vorherigen Beispielen mit Ausnahme von Beispiel 2.1 sind linear. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, ist nicht linear, denn es gilt

$$f(\lambda x) = (\lambda x)^2 = \lambda^2 x^2 \text{ für alle } \lambda \in \mathbb{R}$$

und etwa für $\lambda = 2$ ist $\lambda^2 \neq \lambda$.

Wir geben nun einige Eigenschaften von linearen Abbildungen an.

Satz 2.1 *Es sei $f : V \to W$ eine lineare Abbildung. Dann gilt für alle Vektoren $\mathbf{v}, \mathbf{w} \in V$:*

(a) $f(\mathbf{0}) = \mathbf{0}$.

(b) $f(-\mathbf{v}) = -f(\mathbf{v})$.

(c) $f(\mathbf{v}-\mathbf{w}) = f(\mathbf{v}) - f(\mathbf{w})$.

Beweis. (a) $f(\mathbf{0}) = f(0\mathbf{v}) \overset{(L2)}{=} 0 f(\mathbf{v}) = \mathbf{0}$.

(b) $f(-\mathbf{v}) = f((-1)\mathbf{v}) \overset{(L2)}{=} (-1) f(\mathbf{v}) = -f(\mathbf{v})$.

(c) $f(\mathbf{v}-\mathbf{w}) = f(\mathbf{v}+(-1)\mathbf{w}) \overset{(L)}{=} f(\mathbf{v}) + (-1) f(\mathbf{w}) = f(\mathbf{v}) - f(\mathbf{w})$. □

Satz 2.2 *Es sei V ein endlich dimensionaler K-Vektorraum. Dann ist eine lineare Abbildung $f : V \to W$ bereits durch die Bilder der Vektoren einer Basis vollständig festgelegt.*

Beweis. Es sei $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ eine Basis von V. Dann lässt sich jeder Vektor $\mathbf{v} \in V$ als Linearkombination bezüglich dieser Basis schreiben:

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n.$$

Mit (L) erhalten wir

$$\begin{aligned} f(\mathbf{v}) &= f(\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n) \\ &= \lambda_1 f(\mathbf{v}_1) + \cdots + \lambda_n f(\mathbf{v}_n). \end{aligned}$$

□

Beispiel 2.8 Es sei A eine $m \times n$-Matrix und $f : \mathbb{R}^n \to \mathbb{R}^m$ die lineare Abbildung $f(\mathbf{x}) = A\mathbf{x}$. Es sei $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ die Standardbasis von $\mathbb{R}^n$. Dann ist f durch die Bilder $f(\mathbf{e}_i)$ der Standardbasisvektoren festgelegt. Aber

$$A\mathbf{e}_1, \ldots, A\mathbf{e}_n$$

sind gerade die Spalten der Matrix A in dieser Reihenfolge. Man merke sich also: *Die Spalten der Matrix A sind die Bilder der Standardbasisvektoren.*

Man kann nun Abbildungen auch hintereinanderschalten.

Definition Sind X, Y, Z Mengen und $f : X \to Y$ sowie $g : Y \to Z$ Abbildungen, so heißt die Abbildung

$$g \circ f : X \to Z, \quad x \mapsto (g \circ f)(x) := g(f(x)),$$

die *Komposition* (oder *Hintereinanderschaltung*) von f und g. (Man sagt zu $g \circ f$ auch g "Kringel" f.)

Beispiel 2.9 In § 1 hatten wir die Hintereinanderausführung $\tau \circ \sigma$ von zwei Permutationen σ und τ der Menge $\{1, 2, \ldots, n\}$ betrachtet. Eine Permutation σ kann man auch als eine Abbildung $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, $i \mapsto \sigma(i)$, auffassen. In diesem Sinne ist die Hintereinanderausführung $\tau \circ \sigma$ die Komposition der Abbildungen σ und τ.

Satz 2.3 *Sind U, V, W K-Vektorräume und $f : U \to V$, $g : V \to W$ lineare Abbildungen, so ist auch die Abbildung*

$$g \circ f : U \to W$$

linear.

Beweis. Für $\mathbf{u}, \mathbf{v} \in U$ gilt

$$\begin{aligned}
(g \circ f)(\mathbf{u}+\mathbf{v}) &= g(f(\mathbf{u}+\mathbf{v})) \\
&= g(f(\mathbf{u}) + f(\mathbf{v})) \quad \text{(da } f \text{ linear)} \\
&= g(f(\mathbf{u})) + g(f(\mathbf{v})) \quad \text{(da } g \text{ linear)} \\
&= (g \circ f)(\mathbf{u}) + (g \circ f)(\mathbf{v}).
\end{aligned}$$

Analog zeigt man $(g \circ f)(\lambda \mathbf{u}) = \lambda (g \circ f)(\mathbf{u})$ für $\lambda \in K$. □

3 Kern und Bild

Wir versuchen nun, die Geometrie linearer Abbildungen besser zu verstehen. Dazu führen wir die Begriffe Kern und Bild ein.

Definition Für eine lineare Abbildung $f : V \to W$ definieren wir

$$\begin{aligned}
\operatorname{Im} f &:= f(V) := \{f(\mathbf{v}) \mid \mathbf{v} \in V\}, \quad \textit{Bild} \text{ von } f, \\
\operatorname{Ker} f &:= \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}\}, \quad \textit{Kern} \text{ von } f.
\end{aligned}$$

Satz 3.1 *Für eine lineare Abbildung $f : V \to W$ gilt:*

(a) $\operatorname{Im} f$ *ist ein Unterraum von W.*

(b) $\operatorname{Ker} f$ *ist ein Unterraum von V.*

Beweis. (a) Wegen $f(\mathbf{0}) = \mathbf{0}$ enthält $\operatorname{Im} f$ mindestens den Nullvektor. Es seien $\mathbf{w}$, $\mathbf{w}'$ Vektoren aus $\operatorname{Im} f$. Dann gibt es $\mathbf{v}$, $\mathbf{v}' \in V$ mit $f(\mathbf{v}) = \mathbf{w}$, $f(\mathbf{v}') = \mathbf{w}'$. Dann gilt

$$f(\mathbf{v}+\mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}') = \mathbf{w} + \mathbf{w}'.$$

Also ist auch $\mathbf{w} + \mathbf{w}' \in \operatorname{Im} f$.

Analog zeigt man, dass mit $\mathbf{w} \in \operatorname{Im} f$ und $\lambda \in K$ auch $\lambda \mathbf{w} \in \operatorname{Im} f$ ist.

(b) Wegen $f(\mathbf{0}) = \mathbf{0}$ enthält $\operatorname{Ker} f$ mindestens den Nullvektor. Es seien $\mathbf{v}$, $\mathbf{v}' \in \operatorname{Ker} f$. Dann gilt

$$f(\mathbf{v}+\mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}') = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Also ist auch $\mathbf{v} + \mathbf{v}' \in \operatorname{Ker} f$.

Analog zeigt man, dass mit $\mathbf{v} \in \operatorname{Ker} f$ und $\lambda \in K$ auch $\lambda \mathbf{v} \in \operatorname{Ker} f$ ist.
□

Beispiel 3.1 Es sei A eine $m \times n$-Matrix. Wir betrachten die lineare Abbildung

$$f : \mathbb{R}^n \to \mathbb{R}^m, \quad f(\mathbf{x}) = A\mathbf{x}.$$

Dann ist der Kern von f die Lösungsmenge des linearen Gleichungssystems $A\mathbf{x} = \mathbf{0}$. Das Bild von f ist die Menge aller Vektoren $\mathbf{b} \in \mathbb{R}^m$, für die das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ eine Lösung besitzt.

Satz 3.2 (Dimensionsformel) *Es sei $f : V \to W$ eine lineare Abbildung von einem endlich dimensionalen K-Vektorraum V in einen K-Vektorraum W. Dann gilt*

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V.$$

Beweis. Es sei $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ eine Basis von $\operatorname{Ker} f$ und $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$ eine Basis von $\operatorname{Im} f$. Es seien $\mathbf{u}_1, \dots, \mathbf{u}_r$ beliebige Vektoren aus V mit $f(\mathbf{u}_1) = \mathbf{w}_1$, ... , $f(\mathbf{u}_r) = \mathbf{w}_r$. Wir zeigen

Behauptung $S = \{\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_1, \dots, \mathbf{v}_k\}$ ist eine Basis von V.

Es sei $\mathbf{v} \in V$. Dann gilt

$$f(\mathbf{v}) = \mu_1 \mathbf{w}_1 + \cdots + \mu_r \mathbf{w}_r.$$

Es sei

$$\mathbf{v}' = \mu_1 \mathbf{u}_1 + \cdots + \mu_r \mathbf{u}_r.$$

Dann gilt

$$f(\mathbf{v}') = \mu_1 f(\mathbf{u}_1) + \cdots + \mu_r f(\mathbf{u}_r) = f(\mathbf{v}).$$

Also ist $\mathbf{v} - \mathbf{v}' \in \operatorname{Ker} f$. Das bedeutet, dass es $\lambda_1, \dots, \lambda_k \in K$ gibt mit

$$\mathbf{v} - \mathbf{v}' = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k.$$

Dann folgt aber

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k + \mu_1 \mathbf{u}_1 + \cdots + \mu_r \mathbf{u}_r.$$

Also ist S ein Erzeugendensystem von V.

Um zu zeigen, dass S linear unabhängig ist, betrachten wir die Gleichung

$$\mu_1 \mathbf{u}_1 + \cdots + \mu_r \mathbf{u}_r + \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k = \mathbf{0}. \tag{1}$$

Es gilt

$$0 = f(\mu_1 \mathbf{u}_1 + \cdots + \mu_r \mathbf{u}_r + \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k) = \mu_1 \mathbf{w}_1 + \cdots + \mu_r \mathbf{w}_r.$$

Da $\mathbf{w}_1, \dots, \mathbf{w}_r$ linear unabhängig sind, folgt $\mu_1 = \cdots = \mu_r = 0$. In (1) eingesetzt ergibt sich

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k = \mathbf{0}.$$

Da $\mathbf{v}_1, \dots, \mathbf{v}_k$ linear unabhängig sind, folgt daraus $\lambda_1, = \cdots = \lambda_k = 0$. Daraus folgt die Behauptung.

Aus der Behauptung folgt $r + k = \dim V$ und daraus die Behauptung des Satzes. □

Definition Es sei $f : V \to W$ eine lineare Abbildung. Der *Rang* von F, in Zeichen $\operatorname{Rang} f$, ist die Dimension von $\operatorname{Im} f$.

Beispiel 3.2 Es sei A eine $m \times n$-Matrix. Für den Rang der linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ mit $f(\mathbf{x}) = A\mathbf{x}$ gilt $\operatorname{Rang} f = \operatorname{Rang} A$.

Wir betrachten nun wichtige Eigenschaften von Abbildungen.

Definition Es seien X und Y beliebige Mengen. Eine Abbildung $f : X \to Y$ heißt

- *injektiv*, falls aus x, $x' \in X$ und $x \neq x'$ stets $f(x) \neq f(x')$ folgt.
- *surjektiv*, falls es zu jedem $y \in Y$ ein $x \in X$ gibt mit $y = f(x)$, d.h. falls $f(X) = Y$.
- *bijektiv*, falls f injektiv und surjektiv ist.

Bemerkung 3.1 Ist f bijektiv, so gibt es zu jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$. In diesem Fall kann man also eine *Umkehrabbildung*

$$\begin{array}{rccl} f^{-1}: & Y & \longrightarrow & X \\ & y & \longmapsto & x = f^{-1}(y) \text{ mit } y = f(x) \end{array}$$

erklären.

Satz 3.3 *Eine lineare Abbildung $f : V \to W$ ist genau dann injektiv, wenn der Kern von f nur den Nullvektor enthält.*

Beweis.
"$\Rightarrow$": Es sei $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$. Da $f(\mathbf{0}) = \mathbf{0}$ und f injektiv ist, gilt $f(\mathbf{v}) \neq \mathbf{0}$. Also enthält der Kern von f nur den Nullvektor.
"$\Leftarrow$": Es seien $\mathbf{v}, \mathbf{v}' \in V$ mit $\mathbf{v} \neq \mathbf{v}'$. Wir müssen zeigen, dass $f(\mathbf{v}) \neq f(\mathbf{v}')$ gilt. Angenommen, $f(\mathbf{v}) = f(\mathbf{v}')$. Dann gilt

$$\mathbf{0} = f(\mathbf{v}) - f(\mathbf{v}') = f(\mathbf{v} - \mathbf{v}').$$

Also liegt $\mathbf{v} - \mathbf{v}'$ im Kern von f. Da der Kern von f nur den Nullvektor enthält, folgt $\mathbf{v} - \mathbf{v}' = 0$, ein Widerspruch. □

Satz 3.4 *Es seien V, W endlich dimensionale K-Vektorräume. Eine lineare Abbildung $f : V \to W$ ist genau dann surjektiv, wenn* $\operatorname{Rang} f = \dim W$ *gilt.*

Beweis. Die lineare Abbildung f ist genau dann surjektiv, wenn $\operatorname{Im} f = W$ gilt. Da $\operatorname{Im} f$ ein Unterraum von W ist, ist dies nach LAA, Satz 13.4, genau dann der Fall, wenn $\operatorname{Rang} f = \dim \operatorname{Im} f = \dim W$ gilt. □

Satz 3.5 *Für eine lineare Abbildung $f : V \to W$ zwischen endlich dimensionalen K-Vektorräumen* gleicher Dimension *sind die folgenden Bedingungen äquivalent:*

(a) *f ist injektiv.*

(b) *f ist surjektiv.*

(c) *f ist bijektiv.*

Beweis. Dies folgt sofort aus Satz 3.3, Satz 3.4 und der Dimensionsformel (Satz 3.2). □

4 Lineare Abbildungen und Matrizen

Wir wollen nun den Zusammenhang zwischen linearen Abbildungen und Matrizen studieren. Im folgenden seien V und W endlich dimensionale Vektorräume.

Es sei $f : V \to W$ eine lineare Abbildung. Wir wollen dieser Abbildung eine Matrix zuordnen. Dies geschieht wie folgt:

(1) Wir wählen eine Basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ von V und eine Basis $\mathcal{B}' = \{\mathbf{w}_1, \dots \mathbf{w}_m\}$ von W.

(2) Wir stellen die Bilder $f(\mathbf{v}_j)$ der Vektoren der Basis $\mathcal{B}$ von V in der Basis $\mathcal{B}'$ von W dar:

$$f(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i, \quad a_{ij} \in K.$$

Also ist

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

der Koordinatenvektor von $f(\mathbf{v}_j)$ bezüglich der Basis $\mathcal{B}'$ von W.

(3) Wir bilden die Matrix A, die die Koordinatenvektoren von $f(\mathbf{v}_1), \dots,$ $f(\mathbf{v}_n)$ bezüglich der Basis $\mathcal{B}'$ von W als Spaltenvektoren hat:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Definition Die Matrix A heißt *Darstellungsmatrix von f bezüglich der Basen $\mathcal{B}$ und $\mathcal{B}'$*. Wir bezeichnen sie mit $M^{\mathcal{B}}_{\mathcal{B}'}(f)$.

Beispiel 4.1 Wir betrachten die lineare Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^2$, die definiert ist durch

$$\begin{aligned} f_1(x,y,z) &= 2x - 3y + z, \\ f_2(x,y,z) &= -x + 2y - z. \end{aligned}$$

Mit $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ bezeichnen wir die Standardbasisvektoren von $\mathbb{R}^3$. Dann gilt

$$f\,(\mathbf{e}_1) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \; f(\mathbf{e}_2) = \begin{pmatrix} -3 \\ 2 \end{pmatrix}, \; f(\mathbf{e}_3) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Also hat f bezüglich der Standardbasen von $\mathbb{R}^3$ und $\mathbb{R}^2$ die Darstellungsmatrix

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

Bemerkung 4.1 Die lineare Abbildung f ist durch ihre Darstellungsmatrix A bezüglich der Basen $\mathcal{B}$ und $\mathcal{B}'$ eindeutig bestimmt, denn nach Satz 2.2 ist eine lineare Abbildung durch die Bilder der Vektoren einer Basis vollständig festgelegt. Daraus folgt, dass es nach Wahl von Basen $\mathcal{B}$ von V und $\mathcal{B}'$ von W eine eineindeutige Beziehung zwischen linearen Abbildungen und Matrizen gibt: Eine lineare Abbildung bestimmt nach der obigen Konstruktion eine Matrix und eine Matrix A bestimmt umgekehrt eine lineare Abbildung $f : V \to W$ wie folgt: Es sei $\mathbf{v} \in V$ gegeben. Ist $\mathbf{x}$ der Koordinatenvektor des Vektors $\mathbf{v}$ bezüglich der Basis $\mathcal{B}$ von V, so erhält man den Koordinatenvektor $\mathbf{y}$ des Vektors $f(\mathbf{v})$ bezüglich der Basis $\mathcal{B}'$ von W wie folgt:

$$\mathbf{y} = A\mathbf{x}.$$

Satz 4.1 *Es seien U, V, W endlich dimensionale Vektorräume, $f : U \to V$, $g : V \to W$ lineare Abbildungen, $\mathcal{B}$ Basis von U, $\mathcal{B}'$ Basis von V, $\mathcal{B}''$ Basis von W. Dann gilt*

$$M^{\mathcal{B}}_{\mathcal{B}''}(g \circ f) = M^{\mathcal{B}'}_{\mathcal{B}''}(g) M^{\mathcal{B}}_{\mathcal{B}'}(f).$$

Beweis. Es sei

$$\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}, \; \mathcal{B}' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}, \; \mathcal{B}'' = \{\mathbf{w}_1, \dots, \mathbf{w}_m\},$$

$$f(\mathbf{u}_j) = \sum_{k=1}^{n} a_{kj}\mathbf{v}_k, \; g(\mathbf{v}_k) = \sum_{i=1}^{m} b_{ik}\mathbf{w}_i, \; (g \circ f)(\mathbf{u}_j) = \sum_{i=1}^{m} c_{ij}\mathbf{w}_i.$$

Dann gilt

$$\begin{aligned}
(g \circ f)(\mathbf{u}_j) &= g(f(\mathbf{u}_j)) \\
&= g\left(\sum_{k=1}^{n} a_{kj}\mathbf{v}_k\right) \\
&= \sum_{k=1}^{n} a_{kj} g(\mathbf{v}_k) \\
&= \sum_{k=1}^{n} a_{kj}\left(\sum_{i=1}^{m} b_{ik}\mathbf{w}_i\right) \\
&= \sum_{i=1}^{m}\left(\sum_{k=1}^{n} b_{ik}a_{kj}\right)\mathbf{w}_i.
\end{aligned}$$

Daraus folgt

$$c_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj},$$

was zu zeigen war. □

Satz 4.2 *Es sei V ein endlich dimensionaler Vektorraum, $f : V \to V$ eine bijektive lineare Abbildung und $\mathcal{B}$ eine Basis von V. Dann existiert die Umkehrabbildung $f^{-1} : V \to V$ und es gilt*

$$M_{\mathcal{B}}^{\mathcal{B}}(f^{-1}) = M_{\mathcal{B}}^{\mathcal{B}}(f)^{-1}.$$

Beweis. Es gilt $f \circ f^{-1} = \mathrm{id}$, wobei $\mathrm{id} : V \to V$ die identische Abbildung mit $\mathrm{id}(\mathbf{v}) = \mathbf{v}$ ist. Nach Satz 4.1 gilt dann

$$M_{\mathcal{B}}^{\mathcal{B}}(f)M_{\mathcal{B}}^{\mathcal{B}}(f^{-1}) = M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}) = E.$$

Daraus folgt die Behauptung. □

Man beachte, dass die Darstellungsmatrix einer linearen Abbildung von den Basen $\mathcal{B}$ und $\mathcal{B}'$ abhängt. Wir diskutieren nun, wie sich die Darstellungsmatrix einer linearen Abbildung ändert, wenn wir zu anderen Basen übergeben.

Dazu nehmen wir an, dass in dem endlich dimensionalen Vektorraum V zwei Basen $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ und $\mathcal{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ gegeben sind. Dann können wir jeden Vektor $\mathbf{v}_j$ als Linearkombination der Basisvektoren $\mathbf{v}'_1$, $\ldots$, $\mathbf{v}'_n$ darstellen:

$$\mathbf{v}_j = \sum_{i=1}^{n} t_{ij}\mathbf{v}'_i.$$

Definition Die Matrix $T_{\mathcal{B}'}^{\mathcal{B}} := (t_{ij})$, deren Spalten die Koordinatenvektoren der Basisvektoren $\mathbf{v}_j$ bezüglich der Basis $\mathcal{B}'$ sind, heißt die *Transformationsmatrix von $\mathcal{B}$ nach $\mathcal{B}'$*.

Bemerkung 4.2 Es sei $\mathrm{id} : V \to V$ die identische Abbildung, die durch $\mathrm{id}(\mathbf{v}) = \mathbf{v}$ definiert ist. Dann gilt $T_{\mathcal{B}'}^{\mathcal{B}} = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id})$.

Beispiel 4.2 Es sei $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ die Standardbasis des $\mathbb{R}^3$ und $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ die Basis mit

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} \mathbf{e}_1 &= 3\mathbf{v}_1 - \mathbf{v}_2, \\ \mathbf{e}_2 &= -\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3, \\ \mathbf{e}_3 &= -\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3. \end{aligned}$$

Also lautet die Transformationsmatrix von $\mathcal{B}$ nach $\mathcal{B}'$

$$T_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Die Transformationsmatrix von $\mathcal{B}'$ nach $\mathcal{B}$ lautet

$$T_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix}.$$

Nach Beispiel 1.12 folgt, dass

$$T_{\mathcal{B}}^{\mathcal{B}'} = (T_{\mathcal{B}'}^{\mathcal{B}})^{-1}.$$

Ist $\mathbf{x}$ der Koordinatenvektor eines Vektors $\mathbf{v} \in V$ bezüglich der Basis $\mathcal{B}$, so erhält man den Koordinatenvektor $\mathbf{y}$ von $\mathbf{v}$ bezüglich $\mathcal{B}'$ durch

$$\mathbf{y} = T_{\mathcal{B}'}^{\mathcal{B}} \mathbf{x}.$$

Das bedeutet, dass die Transformationsmatrix $T_{\mathcal{B}'}^{\mathcal{B}}$ angibt, wie sich die Koordinaten zur Basis $\mathcal{B}$ in die Koordinaten zur Basis $\mathcal{B}'$ transformieren. Durch Auflösen der obigen Gleichung nach $\mathbf{x}$ erhalten wir

$$\mathbf{x} = (T_{\mathcal{B}'}^{\mathcal{B}})^{-1} \mathbf{y}.$$

Das bedeutet, dass die Matrix $(T_{\mathcal{B}'}^{\mathcal{B}})^{-1}$ die Transformation von den Koordinaten zu $\mathcal{B}'$ in die Koordinaten zu $\mathcal{B}$ angibt. Damit erhalten wir

Satz 4.3 *Es seien $\mathcal{B}$ und $\mathcal{B}'$ Basen des endlich dimensionalen Vektorraums V. Dann ist die Transformationsmatrix $T^{\mathcal{B}}_{\mathcal{B}'}$ invertierbar und es gilt*

$$T^{\mathcal{B}'}_{\mathcal{B}} = (T^{\mathcal{B}}_{\mathcal{B}'})^{-1}.$$

Wir können nun eine Antwort auf die Frage geben, wie sich die Darstellungsmatrix einer linearen Abbildung bei Einführung neuer Basen ändert.

Satz 4.4 (Transformationsformel) *Es sei $f : V \to W$ eine lineare Abbildung, $\mathcal{B}$ und $\mathcal{B}'$ Basen von V und $\mathcal{C}$ und $\mathcal{C}'$ Basen von W. Dann gilt*

$$M^{\mathcal{B}'}_{\mathcal{C}'}(f) = T^{\mathcal{C}}_{\mathcal{C}'} M^{\mathcal{B}}_{\mathcal{C}}(f)(T^{\mathcal{B}}_{\mathcal{B}'})^{-1}.$$

Setzen wir $A := M^{\mathcal{B}}_{\mathcal{C}}(f)$, $B := M^{\mathcal{B}'}_{\mathcal{C}'}(f)$, $T := T^{\mathcal{B}}_{\mathcal{B}'}$ und $S := T^{\mathcal{C}}_{\mathcal{C}'}$, so können wir diese Gleichung auch so ausdrücken

$$B = SAT^{-1}.$$

Beweis. Dies folgt leicht aus Satz 4.1. □

Für den Spezialfall $W = V$ ergibt sich

Korollar 4.1 *Es sei $f : V \to V$ eine lineare Abbildung und es seien $\mathcal{B}$, $\mathcal{B}'$ Basen von V. Wir setzen $A := M^{\mathcal{B}}_{\mathcal{B}}(f)$, $B := M^{\mathcal{B}'}_{\mathcal{B}'}(f)$, $T := T^{\mathcal{B}}_{\mathcal{B}'}$. Dann gilt*

$$B = TAT^{-1}.$$

Beispiel 4.3 Es sei $f : \mathbb{R}^3 \to \mathbb{R}^2$ die lineare Abbildung von Beispiel 4.1. Es sei $\mathcal{B}$ die Standardbasis von $\mathbb{R}^3$ und $\mathcal{C}$ die Standardbasis von $\mathbb{R}^2$. Es sei $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ die Basis von $\mathbb{R}^3$ mit

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

(vgl. Beispiel 4.2). Schließlich sei $\mathcal{C}' = \{\mathbf{w}_1, \mathbf{w}_2\}$ die Basis von $\mathbb{R}^2$ mit

$$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Nach Beispiel 4.2 gilt

$$(T^{\mathcal{B}}_{\mathcal{B}'})^{-1} == \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix}.$$

Für $T^{\mathcal{C}}_{\mathcal{C}'}$ berechnen wir

$$T^{\mathcal{C}}_{\mathcal{C}'} = (T^{\mathcal{C}'}_{\mathcal{C}})^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

Dann gilt für die Darstellungsmatrix von f bezüglich der Basen $\mathcal{B}'$ und $\mathcal{C}'$ nach Satz 4.4

$$\begin{aligned} M_{\mathcal{C}'}^{\mathcal{B}'}(f) &= \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 3 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 0 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -5 & -2 \\ 0 & 8 & 3 \end{pmatrix}. \end{aligned}$$

Definition Zwei $m \times n$-Matrizen A und B heißen *äquivalent*, genau dann, wenn es eine invertierbare $m \times m$-Matrix S und eine invertierbare $n \times n$-Matrix T gibt, so dass

$$B = SAT^{-1}$$

gilt.

Zwei quadratische $n \times n$-Matrizen A und B heißen *ähnlich*, genau dann, wenn es eine invertierbare $n \times n$-Matrix T gibt, so dass

$$B = TAT^{-1}$$

gilt.

Bemerkung 4.3 Damit können wir festhalten: *Zwei Matrizen sind genau dann äquivalent, wenn sie bezüglich verschiedener Paare von Basen die gleiche lineare Abbildung beschreiben. Zwei quadratische Matrizen sind genau dann ähnlich, wenn sie bezüglich verschiedener Basen die gleiche lineare Selbstabbildung beschreiben.*

Satz 4.5 *Es sei $f : V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Vektorräumen vom Rang r. Dann gibt es eine Basis $\mathcal{B}$ von V und eine Basis $\mathcal{C}$ von W, so dass*

$$M_{\mathcal{C}}^{\mathcal{B}}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Beweis. Es sei $\mathcal{B}$ die Basis $\mathcal{B} := \{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots, \mathbf{v}_k\}$ aus dem Beweis von Satz 3.2. Wir ergänzen die Basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ von $\operatorname{Im} f$ aus diesem Beweis zu einer Basis $\mathcal{C}$ von W. Wegen $f(\mathbf{u}_i) = \mathbf{w}_i$, $i = 1, \ldots, r$, hat f dann bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$ die angebene Darstellungsmatrix. □

Korollar 4.2 *Eine $m \times n$-Matrix A vom Rang r ist äquivalent zu der Matrix*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Beweis. Es sei $f : \mathbb{R}^n \to \mathbb{R}^m$ die lineare Abbildung mit $f(\mathbf{x}) = A\mathbf{x}$. Dann hat f den Rang r. Nach Satz 4.5 gibt es eine Basis $\mathcal{B}$ von $\mathbb{R}^n$ und eine Basis $\mathcal{C}$ von $\mathbb{R}^m$, so dass $B := M_{\mathcal{C}}^{\mathcal{B}}(f)$ die angebene Form hat. Es sei T die Transformationsmatrix der Standardbasis von $\mathbb{R}^n$ nach $\mathcal{B}$ und S die Transformationsmatrix der Standardbasis von $\mathbb{R}^m$ nach $\mathcal{C}$. Dann gilt nach Satz 4.4

$$B = SAT^{-1}.$$

Also sind die Matrizen A und B äquivalent. □

5 Eigenwerte und Eigenvektoren

Wir haben gerade gesehen, dass jede Matrix äquivalent zu einer Matrix in besonders einfacher Form ist, nämlich zu der Matrix

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Wir betrachten nun eine quadratische Matrix A und suchen eine Matrix B in möglichst einfacher Form, die ähnlich zu der Matrix A ist, d.h. wir suchen eine invertierbare Matrix T, so dass die Matrix

$$B = TAT^{-1}$$

in möglichst einfacher Form ist. Dieses Problem ist schwieriger zu behandeln, da wir nur *eine* Matrix T zur Verfügung haben. Die Behandlung dieses Problems führt zu der Betrachtung von Eigenwerten und Eigenvektoren.

Ist $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung und $\mathbf{x} \in \mathbb{R}^n$, so besteht im Allgemeinen keine Beziehung zwischen den Vektoren $\mathbf{x}$ und $f(\mathbf{x})$. Häufig existieren jedoch von Null verschiedene Vektoren $\mathbf{x}$, so dass $\mathbf{x}$ und $f(\mathbf{x})$ skalare Vielfache voneinander sind. Solche Vektoren nennt man Eigenvektoren. Sie spielen in Anwendungen eine große Rolle, so zum Beispiel bei der Untersuchung von Schwingungen und elektrischen Systemen.

Im Folgenden sei V ein n-dimensionaler Vektorraum.

Definition Es sei $f : V \to V$ eine lineare Abbildung. Ein Vektor $\mathbf{v} \in V$ heißt *Eigenvektor* von f, wenn $\mathbf{v} \neq \mathbf{0}$ und es ein $\lambda \in \mathbb{R}$ gibt mit

$$f(\mathbf{v}) = \lambda\mathbf{v}.$$

Der Skalar λ heißt dann *Eigenwert* von f.

Wegen des engen Zusammenhangs zwischen linearen Abbildungen und Matrizen können wir auch zu quadratischen Matrizen Eigenvektoren und Eigenwerte definieren.

Definition Es sei A eine $n \times n$-Matrix. Ein Vektor $\mathbf{x} \in \mathbb{R}^n$ heißt *Eigenvektor* von A, wenn $\mathbf{x} \neq \mathbf{0}$ und es eine Zahl $\lambda \in \mathbb{R}$ gibt mit

$$A\mathbf{x} = \lambda \mathbf{x}.$$

Die Zahl λ heißt dann *Eigenwert* von A.

Bemerkung 5.1 Man beachte, dass der Nullvektor ausdrücklich ausgeschlossen ist. Denn für den Nullvektor gilt $f(\mathbf{0}) = \mathbf{0} = \lambda\mathbf{0}$ bzw. $A\mathbf{0} = \mathbf{0} = \lambda\mathbf{0}$ für jeden Skalar λ.

Beispiel 5.1 Es sei $V = \mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der Geraden $x - y = 0$. Die Abbildung f hat bezüglich der Standardbasis die Darstellungsmatrix

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

Es ist anschaulich klar, dass $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 1 ist und $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert -1 ist.

Ist $\mathbf{x}$ Eigenvektor der linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ zum Eigenwert λ, so ist $\mathbf{x}$ auch Eigenvektor der Matrixdarstellung A von f bezüglich der Standardbasis zu demselben Eigenwert. Umgekehrt ist jeder Eigenvektor einer $n \times n$-Matrix A zum Eigenwert λ auch Eigenvektor der linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{x}) = A\mathbf{x}$, zu demselben Eigenwert. Deswegen werden wir im Folgenden nur Eigenvektoren und Eigenwerte einer Matrix betrachten.

Wie bestimmt man nun die Eigenwerte einer $n \times n$-Matrix A? Dazu formen wir die Gleichung $A\mathbf{x} = \lambda\mathbf{x}$ um:

$$\begin{aligned} & A\mathbf{x} = \lambda\mathbf{x} \\ \Leftrightarrow\ & A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \\ \Leftrightarrow\ & A\mathbf{x} - \lambda E\mathbf{x} = \mathbf{0} \\ \Leftrightarrow\ & (A - \lambda E)\mathbf{x} = \mathbf{0}. \end{aligned}$$

Damit λ ein Eigenwert von A ist, muss die Gleichung

$$(A - \lambda E)\mathbf{x} = \mathbf{0}$$

eine nichttriviale Lösung besitzen. Nach LAA, Satz 6.6, ist das genau dann der Fall, wenn die Matrix $A - \lambda E$ nicht invertierbar ist. Nach Korollar 1.1 ist dies äquivalent zu

$$\det(A - \lambda E) = 0.$$

Definition Die Gleichung

$$\det(A - \lambda E) = 0$$

für die Unbekannte λ heißt die *charakteristische Gleichung* von A. Ihre Lösungen sind die Eigenwerte von A. Entwickelt man $\det(A - \lambda E)$, so ergibt sich ein Polynom $P_A(\lambda)$ in λ, das als *charakteristisches Polynom* von A bezeichnet wird.

Es gilt

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda E) \\ &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} \\ &= (-1)^n \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n. \end{aligned}$$

Das charakteristische Polynom einer $n \times n$-Matrix ist also ein Polynom n-ten Grades. Ein solches Polynom hat nach dem Fundamentalsatz der Algebra höchstens n verschiedene Nullstellen. Deswegen besitzt eine $n \times n$-Matrix höchstens n verschiedene Eigenwerte.

Beispiel 5.2 Das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

lautet

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda E) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} \\ &= (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc). \end{aligned}$$

Für die Matrix

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

erhalten wir

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

Die Eigenwerte von A sind also $+1$ und -1 in Übereinstimmung mit Beispiel 5.1.

Satz 5.1 *Es sei*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix. Dann sind die Eigenwerte von A die Hauptdiagonalelemente von A.

Beweis. Es gilt

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda E) \\ &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} - \lambda \end{vmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda). \end{aligned}$$

Also sind die Eigenwerte

$$\lambda = a_{11}, \quad \lambda = a_{22}, \quad \ldots, \quad \lambda = a_{nn}.$$

□

Das charakteristische Polynom einer Matrix mit reellen Einträgen kann auch komplexe Nullstellen haben.

Beispiel 5.3 Die Matrix

$$A = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat das charakteristische Polynom

$$\det(A - \lambda E) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

Dieses Polynom hat die komplexen Nullstellen $\lambda = i$ und $\lambda = -i$.

Wir befassen uns nun mit der Bestimmung der Eigenvektoren zu einem gegebenen Eigenwert.

Definition Zu einem Eigenwert $\lambda \in \mathbb{R}$ nennen wir

$$\mathrm{Eig}(A, \lambda) := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

den *Eigenraum* von A zum Eigenwert λ.

Bemerkung 5.2 $\mathrm{Eig}(A, \lambda)$ ist ein Unterraum des $\mathbb{R}^n$, denn $\mathrm{Eig}(A, \lambda)$ ist der Lösungsraum des linearen Gleichungssystems $(A - \lambda E)\mathbf{x} = \mathbf{0}$.

Zur Bestimmung des Eigenraums zu einem Eigenwert λ muss man also das lineare Gleichungssystem

$$(A - \lambda E)\mathbf{x} = \mathbf{0}$$

lösen.

Beispiel 5.4 Wir bestimmen eine Basis des Eigenraums zum Eigenwert 1 der Matrix

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

Dazu müssen wir eine Basis des Lösungsraum des Gleichungssystems

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

bestimmen. Man sieht leicht, dass dieses lineare Gleichungssystem einen eindimensionalen Lösungsraum mit der Basis $\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$ hat.

6 Diagonalisierung

Wir kommen nun auf das eingangs gestellte Problem der Vereinfachung einer quadratischen Matrix zurück.

Definition Eine *Diagonalmatrix* A ist eine quadratische Matrix, bei der alle Elemente außer eventuell den Diagonalelementen gleich Null sind, d.h. A hat die Form

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Definition Eine $n \times n$-Matrix A heißt *diagonalisierbar*, wenn sie ähnlich zu einer Diagonalmatrix ist, d.h. wenn eine invertierbare Matrix T existiert, so dass TAT^{-1} Diagonalgestalt hat.

Satz 6.1 *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

Beweis. Es sei $B = TAT^{-1}$. Dann gilt

$$\begin{aligned} \det(B - \lambda E) &= \det(TAT^{-1} - \lambda TET^{-1}) = \det(T(A - \lambda E)T^{-1}) \\ &= \det T \det(A - \lambda E) \det T^{-1} = \det(A - \lambda E). \end{aligned}$$

□

Korollar 6.1 *Ist A diagonalisierbar, so ist A ähnlich zu einer Diagonalmatrix, bei der auf der Diagonalen die Eigenwerte von A stehen.*

Beweis. Dies folgt sofort aus Satz 6.1. □

Satz 6.2 *Eine $n \times n$-Matrix A ist genau dann diagonalisierbar, wenn sie n linear unabhängige Eigenvektoren besitzt.*

Beweis. "$\Rightarrow$": Es sei A diagonalisierbar. Dann gibt es eine invertierbare $n \times n$-Matrix T, so dass TAT^{-1} Diagonalgestalt hat. Also ist $TAT^{-1} = D$ mit

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

Aus $TAT^{-1} = D$ folgt $AT^{-1} = T^{-1}D$. Es sei $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ die Standardbasis von $\mathbb{R}^n$. Dann gilt

$$A(T^{-1}(\mathbf{e}_i)) = T^{-1}(\lambda_i \mathbf{e}_i) = \lambda_i T^{-1}(\mathbf{e}_i) \quad \text{für } i = 1, \ldots, n.$$

Also ist $T^{-1}(\mathbf{e}_i)$ ein Eigenvektor von A zum Eigenwert λ_i, $i = 1, \ldots, n$. Da $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ linear unabhängig ist und T^{-1} invertierbar ist, ist auch $\{T^{-1}(\mathbf{e}_1), \ldots, T^{-1}(\mathbf{e}_n)\}$ linear unabhängig.

"$\Leftarrow$": Es seien $\mathbf{v}_1, \ldots, \mathbf{v}_n$ die linear unabhängigen Eigenvektoren von A zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Es sei $\mathcal{B}$ die Standardbasis von $\mathbb{R}^n$ und $\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Es sei D die Matrix

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

Dann ist D die Darstellungsmatrix $M_{\mathcal{B}'}^{\mathcal{B}'}(f)$ der linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ mit $f(\mathbf{x}) = A\mathbf{x}$ bezüglich der Basis $\mathcal{B}'$. Nach Satz 4.4 gilt

$$A = T_{\mathcal{B}}^{\mathcal{B}'} D (T_{\mathcal{B}}^{\mathcal{B}'})^{-1}.$$

Also ist A diagonalisierbar. □

Beispiel 6.1 Wir untersuchen die Matrix

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

auf Diagonalisierbarkeit. Das charakteristische Polynom

$$\det(A - \lambda E) = (1 - \lambda)^2$$

liefert 1 als einzigen Eigenwert, aber der Eigenraum $\mathrm{Eig}(A, 1)$ zum Eigenwert 1 wird von dem Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ erzeugt. Also besitzt die Matrix A nur einen linear unabhängigen Eigenvektor. Nach Satz 6.2 ist A nicht diagonalisierbar.

Satz 6.3 *Es sei A eine $n \times n$-Matrix. Es seien $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von A und $\mathbf{v}_1, \dots, \mathbf{v}_k$ zugehörige Eigenvektoren. Dann ist die Menge $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ linear unabhängig.*

Beweis. Wir führen Induktion über k durch.

Induktionsanfang: Da ein Eigenvektor $\mathbf{v}_1$ nach Definition verschieden vom Nullvektor ist und die Menge $\{\mathbf{v}_1\}$ linear unabhängig ist, ist der Satz für k=1 richtig.

Induktionsschritt: Es sei $k \geq 2$. Wir nehmen an, dass die Aussage bereits für $k-1$ bewiesen ist. Wir müssen zeigen, dass $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ linear unabhängig ist. Dazu betrachten wir die Gleichung

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}. \tag{2}$$

Wir multiplizieren diese Gleichung einerseits mit A und andererseits mit λ_1:

$$\begin{aligned} \alpha_1 \lambda_1 \mathbf{v}_1 + \alpha_2 \lambda_2 \mathbf{v}_2 + \cdots + \alpha_k \lambda_k \mathbf{v}_k &= \mathbf{0} \\ \alpha_1 \lambda_1 \mathbf{v}_1 + \alpha_2 \lambda_1 \mathbf{v}_2 + \cdots + \alpha_k \lambda_1 \mathbf{v}_k &= \mathbf{0} \end{aligned}$$

Subtrahieren wir die zweite Gleichung von der ersten, so ergibt sich

$$\alpha_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \cdots + \alpha_k(\lambda_k - \lambda_1)\mathbf{v}_k = \mathbf{0}.$$

Nach Induktionsannahme sind die Vektoren $\mathbf{v}_2, \dots, \mathbf{v}_k$ linear unabhängig. Daraus folgt

$$\alpha_2(\lambda_2 - \lambda_1) = \cdots = \alpha_k(\lambda_k - \lambda_1) = 0.$$

Da nach Voraussetzung die Eigenwerte paarweise verschieden sind, ergibt sich daraus

$$\alpha_2 = \cdots = \alpha_k = 0.$$

Setzt man dies in Gleichung (2) ein, so folgt

$$\alpha_1 \mathbf{v}_1 = \mathbf{0}$$

und daraus $\alpha_1 = 0$, da $\mathbf{v} \neq \mathbf{0}$. □

Aus Satz 6.3 erhalten wir als Folgerung:

Satz 6.4 *Besitzt eine $n \times n$-Matrix A n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.*

Beweis. Es seien $\mathbf{v}_1, \dots, \mathbf{v}_n$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$. Nach Satz 6.3 sind $\mathbf{v}_1, \dots, \mathbf{v}_n$ linear unabhängig. Nach Satz 6.2 ist A deshalb diagonalisierbar. □

Satz 6.4 liefert nur ein hinreichendes, aber kein notwendiges Kriterium für die Diagonalisierbarkeit. Zum Beispiel ist die $n \times n$-Einheitsmatrix schon eine Diagonalmatrix, also diagonalisierbar, sie hat aber lauter gleiche Eigenwerte, nämlich den Eigenwert 1 mit der Vielfachheit n. In Beispiel 6.1 hatten wir gesehen, dass eine Matrix den Eigenwert 1 mit der Vielfachheit 2 haben kann, aber nicht diagonalisierbar zu sein braucht. Wir haben in diesem Beispiel gesehen, dass die Vielfachheit eines Eigenwerts und die Dimension des zugehörigen Eigenraums eine Rolle spielen.

Definition Es sei A eine $n \times n$-Matrix und λ_0 ein Eigenwert von A. Die Vielfachheit der Nullstelle λ_0, d.h. die Anzahl der Faktoren $(\lambda - \lambda_0)$, im charakteristischen Polynom von A heißt die *algebraische Vielfachheit* des Eigenwerts λ_0.

Die Dimension des Eigenraums zum Eigenwert λ_0 von A heißt die *geometrische Vielfachheit* des Eigenwerts λ_0.

Beispiel 6.2 In Beispiel 6.1 war die algebraische Vielfachheit des Eigenwerts 1 von A gleich 2, aber die geometrische Vielfachheit des Eigenwerts 1 gleich 1.

Satz 6.5 *Für eine $n \times n$-Matrix A gilt: Die geometrische Vielfachheit eines Eigenwertes von A ist nicht größer als seine algebraische Vielfachheit.*

Beweis. Es sei $\mathcal{B}$ die Standardbasis des $\mathbb{R}^n$. Es sei $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ eine Basis des Eigenraums $\mathrm{Eig}(A, \lambda_0)$ von A zu einem Eigenwert λ_0. Wir ergänzen diese Basis zu einer Basis

$$\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$$

von $\mathbb{R}^n$. Dann gilt

$$T_{\mathcal{B}'}^{\mathcal{B}} A (T_{\mathcal{B}'}^{\mathcal{B}})^{-1} = \left(\begin{array}{ccc|c} \lambda_0 & & 0 & \\ & \ddots & & * \\ 0 & & \lambda_0 & \\ \hline & 0 & & B \end{array} \right).$$

Daraus folgt nach Satz 6.1

$$\det(A - \lambda E) = \det(T_{\mathcal{B}'}^{\mathcal{B}} A (T_{\mathcal{B}'}^{\mathcal{B}})^{-1} - \lambda E) = (\lambda - \lambda_0)^m \det(B - \lambda E).$$

□

Ohne Beweis geben wir den folgenden Satz an, der eine Charakterisierung diagonalisierbarer Matrizen liefert.

Satz 6.6 *Eine $n \times n$-Matrix A ist genau dann diagonalisierbar, wenn sie mit Vielfachheit gezählt n reelle Eigenwerte hat und die geometrische Vielfachheit jedes Eigenwerts mit seiner algebraischen Vielfachheit übereinstimmt.*

Wenn eine $n \times n$-Matrix A mit Vielfachheit gezählt n reelle Eigenwerte hat, aber die geometrische Vielfachheit eines Eigenwerts nicht in jedem Fall mit seiner algebraischen Vielfachheit übereinstimmt, dann kann man die Matrix noch auf eine sogenannte Jordan'sche Normalform bringen. Wir wollen ohne Beweis angeben, was man darunter versteht.

Definition Es sei $m \geq 1$ und $\lambda \in \mathbb{R}$. Eine $m \times m$-Matrix der Form

$$J(m, \lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

heißt eine *Jordanblockmatrix.*

Definition Eine Matrix B heißt in *Jordan'scher Normalform*, wenn sie von der Gestalt

$$B = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_k} \end{pmatrix}$$

ist, wobei $J_1, \ldots, J_k$ Jordanblockmatrizen sind.

Ohne Beweis geben wir den folgenden Satz an.

Satz 6.7 (Jordan'sche Normalform) *Es sei A eine $n \times n$-Matrix, die mit Vielfachheit gezählt n reelle Eigenwerte hat. Dann ist A ähnlich zu einer Matrix in Jordan'scher Normalform.*

Beispiel 6.3 Eine 2×2-Matrix A, die mit Vielfachheit gezählt 2 reelle Eigenwerte hat, ist also ähnlich zu einer der folgenden Matrizen

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ oder } \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

wobei $\lambda_1, \lambda_2, \lambda \in \mathbb{R}$ die Eigenwerte sind. Man beachte dabei, dass auch $\lambda_1 = \lambda_2$ zugelassen ist.

Wir geben nun noch ein wichtiges Anwendungsbeispiel für die Diagonalisierbarkeit von Matrizen.

Beispiel 6.4 Wir betrachten die Schwingung eines Federpendels. Es sei eine Masse $m > 0$ an einer Feder aufgehängt. Zum Zeitpunkt $t = 0$ sei sie in senkrechter Richtung in die Position $y(0) = \alpha$ mit der Geschwindigkeit $\dot{y}(0) = \beta$ ausgelenkt. Dann wird die weitere Bewegung durch die Differentialgleichung

$$m \cdot \ddot{y} + r \cdot \dot{y} + k \cdot y = 0$$

mit den Anfangsbedingungen

$$y(0) = \alpha, \quad \dot{y} = \beta$$

bestimmt. Dabei ist $r \in \mathbb{R}$, $r \geq 0$, eine Konstante, die durch die Reibung bestimmt ist, und $k \in \mathbb{R}$, $k \geq 0$, die Federkonstante. Wir setzen

$$\mu = \frac{r}{2m}, \quad \omega = \sqrt{\frac{k}{m}}.$$

Dann schreibt sich die Differentialgleichung als

$$\ddot{y} + 2\mu\dot{y} + \omega^2 y = 0.$$

Daraus macht man üblicherweise durch $y_1 = y$ und $y_2 = \dot{y}$ ein lineares Differentialgleichungssystem erster Ordnung

$$\begin{aligned} \dot{y_1} &= y_2, \\ \dot{y_2} &= -\omega^2 y_1 - 2\mu y_2. \end{aligned}$$

Dieses Differentialgleichungssystem können wir auch wie folgt schreiben

$$\begin{pmatrix} \dot{y_1} \\ \dot{y_2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Die Matrix

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

bezeichnet man als die Koeffizientenmatrix dieses Differentialgleichungssystems.

Angenommen, unsere Koeffizientenmatrix ist eine Matrix D in Diagonalgestalt, d.h.

$$D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}.$$

Dann lautet das zugehörige Differentialgleichungssystem

$$\begin{aligned} \dot{y_1} &= d_1 y_1 \\ \dot{y_2} &= d_2 y_2. \end{aligned}$$

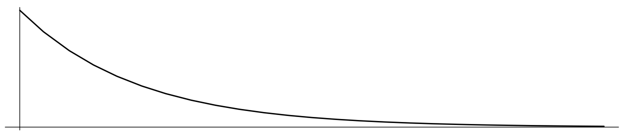


Abbildung 1: Starke Dämpfung: $y(t) = e^{-t}$

Dieses System hat die allgemeine Lösung

$$\begin{aligned} y_1(t) &= c_1 e^{d_1 t} \\ y_2(t) &= c_2 e^{d_2 t}. \end{aligned}$$

Wir sehen also, das die Diagonalisierung der Koeffizientenmatrix A dazu führt, dass wir das System einfach lösen können.

Deswegen versuchen wir nun, unsere Koeffizientenmatrix A zu diagonalisieren. Sie hat das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 + 2\mu\lambda + \omega^2.$$

Die Nullstellen des charakteristischen Polynoms sind also

$$\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega^2}.$$

Entscheidend für die Art der Bewegung ist die Diskriminante $\mu^2 - \omega^2$. Wir unterscheiden zwischen drei Fällen:

1. Fall: $\mu^2 - \omega^2 > 0$ (starke Dämpfung).

Dann gibt es zwei verschiedene negative reelle Eigenwerte λ_1, λ_2 und die Matrix A ist diagonalisierbar. Die allgemeine Lösung lautet

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}, \quad c_1, c_2 \in \mathbb{R}.$$

Die Lösungen klingen ohne Schwingungen, d.h. *aperiodisch*, ab.

2. Fall: $\mu^2 - \omega^2 = 0$ (aperiodischer Grenzfall).

Dann ist $\lambda = -\mu = -\omega$ ein 2-facher Eigenwert, die Matrix

$$A - \lambda E = \begin{pmatrix} \omega & 1 \\ -\omega^2 & -\omega \end{pmatrix}$$

hat den Rang 1, also hat der Eigenraum $\mathrm{Eig}(A, -\omega)$ die Dimension 1 und A ist nicht diagonalisierbar. Die allgemeine Lösung lautet in diesem Fall

$$y(t) = (c_1 + c_2 t)e^{-\mu t}, \quad c_1, c_2 \in \mathbb{R}.$$

Die Lösungen haben qualitativ das gleiche Verhalten wie im ersten Fall: Das Pendel schwingt noch einmal durch die Ruhelage hindurch, um dann langsam endgültig in die Ruhelage zurückzukehren.

Abbildung 2: Aperiodischer Grenzfall: $y(t) = (1 - 3t)e^{-t}$

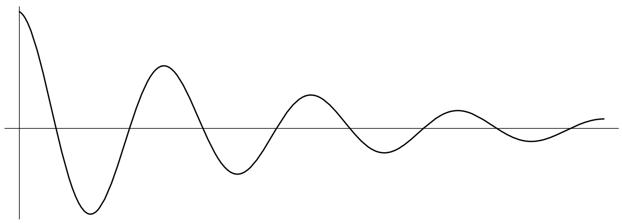


Abbildung 3: Schwache Dämpfung: $y(t) = e^{-(1/2)t} \cos 5t$

3. Fall: $\mu^2 - \omega^2 < 0$ (schwache Dämpfung)
Dann gibt es keine reellen Eigenwerte, dafür aber zwei verschiedene komplexe. Man kann A dann komplex diagonalisieren, dann zunächst komplexe und daraus reelle Lösungen berechnen. Als allgemeine Lösung erhält man

$$y(t) = e^{-\mu t}(c_1 \cos \omega t + c_2 \sin \omega t), \quad c_1, c_2 \in \mathbb{R}.$$

Das bedeutet, es treten gedämpfte (die Amplitude wird um den Faktor $e^{-\mu t}$ gedämpft) Schwingungen mit der *Eigenfrequenz* ω auf.

7 Orthogonale Abbildungen

Wir bringen nun das Skalarprodukt im $\mathbb{R}^n$ ins Spiel. Zunächst betrachten wir Basen mit besonderen Eigenschaften.

Definition Eine Teilmenge $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ des $\mathbb{R}^n$ heißt *orthogonal*, wenn die Vektoren $\mathbf{v}_i$ paarweise orthogonal sind, d.h. wenn

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{für alle } i \neq j, i, j = 1, \ldots, m.$$

Eine Teilmenge $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ heißt *orthonormal*, wenn S orthogonal ist und alle Vektoren auf die Länge 1 normiert sind, d.h.

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{für alle } i \neq j \text{ und } |\mathbf{v}_i| = 1$$

für $i, j = 1, \ldots, m$.
Eine Basis des $\mathbb{R}^n$ heißt *Orthonormalbasis*, wenn sie orthonormal ist.

Beispiel 7.1 Die Standardbasis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ des $\mathbb{R}^n$ ist eine Orthonormalbasis.

In vielen Fällen ist es nützlich, zu einer gegebenen Basis eines Unterraums von $\mathbb{R}^n$ eine Orthonormalbasis zu konstruieren. Dies kann mit dem *E. Schmidt'schen Orthonormalisierungsverfahren* geschehen, das wir nun darstellen. Zum Verständnis dieses Verfahrens benötigen wir die folgenden Sätze.

Satz 7.1 *Eine orthogonale Teilmenge $S = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ des $\mathbb{R}^n$ mit vom Nullvektor verschiedenen Elementen ist linear unabhängig.*

Beweis. Wir betrachten die Gleichung

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m = \mathbf{0}.$$

Wir müssen zeigen, dass hieraus $\alpha_1 = \cdots = \alpha_m = 0$ folgt. Dazu bilden wir das Skalarprodukt von beiden Seiten der Gleichung mit den Vektoren $\mathbf{v}_i$ von S. Da S orthogonal ist, gilt $\mathbf{v}_j \cdot \mathbf{v}_i =$ für $i \neq j$, also folgt

$$\alpha_i(\mathbf{v}_i \cdot \mathbf{v}_i) = \mathbf{0} \cdot \mathbf{v}_i = 0.$$

Da nach Voraussetzung $\mathbf{v}_i \neq \mathbf{0}$, folgt daraus $\alpha_i = 0$ für $i = 1, \dots, m$. □

Satz 7.2 *Es sei W ein Unterraum von $\mathbb{R}^n$ und $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ eine Orthonormalbasis von W. Es sei $\mathbf{u} \in \mathbb{R}^n$. Dann ist der Vektor*

$$\mathbf{v} := \mathbf{u} - (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 - \cdots - (\mathbf{u} \cdot \mathbf{v}_m)\mathbf{v}_m$$

orthogonal zu allen Vektoren von W.

Definition Der Vektor

$$\mathbf{u} - (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 - \cdots - (\mathbf{u} \cdot \mathbf{v}_m)\mathbf{v}_m$$

heißt die *zu W orthogonale Komponente von* $\mathbf{u}$.

Beweis. Wegen der Bilinearität des Skalarproduktes reicht es zu zeigen, dass der Vektor $\mathbf{v}$ orthogonal zu den Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_m$ ist. Es gilt aber

$$\mathbf{v} \cdot \mathbf{v}_i = \mathbf{u} \cdot \mathbf{v}_i - \mathbf{u} \cdot \mathbf{v}_i = 0,$$

da $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ für $i \neq j$ und $\mathbf{v}_i \cdot \mathbf{v}_j = 1$ für $i = j$, $i, j = 1, \dots, m$. □

Bemerkung 7.1 Gilt $\mathbf{u} \neq \mathbf{0}$ und $\mathbf{u} \notin W$, so ist $\mathbf{v} \neq \mathbf{0}$.

E. Schmidt'sches Orthonormalisierungsverfahren
Gegeben sei eine linear unabhängige Menge $S := \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ von Vektoren des $\mathbb{R}^n$ $(k \leq n)$. Wir konstruieren eine Orthonormalbasis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ von $\operatorname{span} S$.

Schritt 1. Setze
$$\mathbf{v}_1 := \frac{\mathbf{u}_1}{|\mathbf{u}_1|}.$$

Schritt 2. Es sei $W_1 := \operatorname{span}\{\mathbf{v}_1\}$. Wir betrachten die zu W_1 orthogonale Komponente $\mathbf{v}_2'$ von $\mathbf{u}_2$:

$$\mathbf{v}_2' := \mathbf{u}_2 - (\mathbf{u}_2 \cdot \mathbf{v}_1)\mathbf{v}_1.$$

Nach der Bemerkung ist $\mathbf{v}_2' \neq \mathbf{0}$. Also können wir setzen

$$\mathbf{v}_2 := \frac{\mathbf{v}_2'}{|\mathbf{v}_2'|}.$$

Schritt 3. Es sei $W_2 := \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Wir betrachten die zu W_2 orthogonale Komponente $\mathbf{v}_3'$ von $\mathbf{u}_3$:

$$\begin{aligned} \mathbf{v}_3' &:= \mathbf{u}_3 - (\mathbf{u}_3 \cdot \mathbf{v}_1)\mathbf{v}_1 - (\mathbf{u}_3 \cdot \mathbf{v}_2)\mathbf{v}_2, \\ \mathbf{v}_3 &:= \frac{\mathbf{v}_3'}{|\mathbf{v}_3'|} \end{aligned}$$

Schritt k. Es sei $W_{k-1} := \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$. Wir betrachten die zu W_{k-1} orthogonale Komponente $\mathbf{v}_k'$ von $\mathbf{u}_k$:

$$\begin{aligned} \mathbf{v}_k' &:= \mathbf{u}_k - (\mathbf{u}_k \cdot \mathbf{v}_1)\mathbf{v}_1 - \cdots - (\mathbf{u}_k \cdot \mathbf{v}_{k-1})\mathbf{v}_{k-1}, \\ \mathbf{v}_k &:= \frac{\mathbf{v}_k'}{|\mathbf{v}_k'|} \end{aligned}$$

Nach Satz 7.2 ist $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ orthonormal, nach Satz 7.1 also eine Orthonormalbasis von $\operatorname{span} S$.

Wir betrachten nun eine besondere Klasse von linearen Abbildungen, nämlich diejenigen, die das Skalarprodukt invariant lassen.

Definition Eine lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ heißt *orthogonal*, wenn sie das Skalarprodukt invariant lässt, d.h. wenn

$$f(\mathbf{x}) \cdot f(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Definition Eine $n \times n$-Matrix A heißt *orthogonal*, wenn gilt:

$$A^T A = E \quad (\text{also } A^T = A^{-1}).$$

Satz 7.3 *Für eine $n \times n$-Matrix A sind die folgenden Aussagen äquivalent:*

(a) *A ist orthogonal.*

(b) *Die durch $f(\mathbf{x}) = A\mathbf{x}$ definierte lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ ist orthogonal.*

(c) *Die Spaltenvektoren von A bilden eine Orthonormalbasis des $\mathbb{R}^n$.*

Beweis. (a) $\Rightarrow$ (b): Es sei A orthogonal. Dann gilt

$$\begin{aligned}(A\mathbf{x}) \cdot (A\mathbf{y}) &= (A\mathbf{x})^T(A\mathbf{y}) = \mathbf{x}^T A^T A\mathbf{y} \\ &= \mathbf{x}^T E\mathbf{y} = \mathbf{x}^T\mathbf{y} \\ &= \mathbf{x} \cdot \mathbf{y}\end{aligned}$$

(b) $\Rightarrow$ (c): Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{x}) = A\mathbf{x}$, orthogonal. Die Spalten der Matrix A sind die Bilder $f(\mathbf{e}_i)$ der Basisvektoren der Standardbasis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ des $\mathbb{R}^n$. Da f orthogonal ist, gilt

$$(A\mathbf{e}_i) \cdot (A\mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{falls } i \neq j. \end{cases}$$

Also bilden die Spaltenvektoren von A eine Orthonormalbasis des $\mathbb{R}^n$.

(c) $\Rightarrow$ (a): Es sei $\mathbf{a}_i := A\mathbf{e}_i$ der i-te Spaltenvektor von A. Dann gilt

$$A^T A = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T\mathbf{a}_1 & \cdots & \mathbf{a}_1^T\mathbf{a}_n \\ \vdots & \ddots & \vdots \\ \mathbf{a}_n^T\mathbf{a}_1 & \cdots & \mathbf{a}_n^T\mathbf{a}_n \end{pmatrix} = E.$$

□

Korollar 7.1 *Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung und $\mathcal{B}$ die Standardbasis von $\mathbb{R}^n$. Dann ist f genau dann orthogonal, wenn die Darstellungsmatrix $M_{\mathcal{B}}^{\mathcal{B}}(f)$ von f bezüglich der Standardbasis $\mathcal{B}$ orthogonal ist.*

Beweis. Die Spalten der Darstellungsmatrix $M_{\mathcal{B}}^{\mathcal{B}}(f)$ sind die Bilder $f(\mathbf{e}_i)$ der Basisvektoren $\mathbf{e}_i$ der Standardbasis $\mathcal{B}$. Damit folgt die Behauptung aus Satz 7.3. □

Wir notieren einige Eigenschaften von orthogonalen Abbildungen und Matrizen.

Satz 7.4 *Für orthogonale Abbildungen $f, g : \mathbb{R}^n \to \mathbb{R}^n$ gilt:*

(a) *f erhält Längen, d.h. $|f(\mathbf{x})| = |\mathbf{x}|$ für alle $\mathbf{x} \in \mathbb{R}^n$.*

(b) *f erhält Winkel, d.h. $\angle(f(\mathbf{x}), f(\mathbf{y})) = \angle(\mathbf{x}, \mathbf{y})$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.*

(c) *$\mathbf{x} \perp \mathbf{y} \Rightarrow f(\mathbf{x}) \perp f(\mathbf{y})$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.*

(d) *f ist bijektiv und f^{-1} ist auch orthogonal.*

(e) *$f \circ g$ ist ebenfalls orthogonal.*

(f) *Ist $\lambda \in \mathbb{R}$ Eigenwert von f, so ist $|\lambda| = 1$.*

Beweis. Die Aussagen (a) – (c) sind klar. Aus (a) folgt, dass f injektiv ist. Aus Satz 3.5 folgt damit (d).

(e) Für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt

$$(f \circ g)(\mathbf{x}) \cdot (f \circ g)(\mathbf{y}) = f(g(\mathbf{x})) \cdot f(g(\mathbf{y})) = g(\mathbf{x}) \cdot g(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}.$$

(f) Es sei $\mathbf{v}$ Eigenvektor von f zum Eigenwert λ. Dann gilt

$$|\mathbf{v}| = |f(\mathbf{v})| = |\lambda \mathbf{v}| = |\lambda||\mathbf{v}|.$$

Wegen $\mathbf{v} \neq \mathbf{0}$ folgt daraus $|\lambda| = 1$. □

Satz 7.5 *Für orthogonale Matrizen A, B gilt:*

(a) *A ist invertierbar und A^{-1} ist auch orthogonal.*

(b) *AB ist auch orthogonal.*

(c) $|\det A| = 1$.

Beweis. (a) Aus $A^T A = E$ folgt, dass A invertierbar ist und $A^{-1} = A^T$ gilt. Außerdem folgt $(A^{-1})^T A^{-1} = AA^T = A^T A = E$.

(b) $(AB)^T AB = B^T A^T AB = B^T B = E$.

(c) Aus $A^T A = E$ folgt $\det A^T \det A = \det A \det A = 1$. □

Wir betrachten nun orthogonale Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^n$ für $n = 1, 2, 3$.

a) Im Fall $n = 1$ gibt es nur die Möglichkeiten

$$f(x) = \pm x.$$

b) Im Fall $n = 2$ klassifizieren wir die orthogonalen 2×2-Matrizen.

Satz 7.6 *Ist A eine orthogonale 2×2-Matrix, so gibt es ein $\alpha \in [0, 2\pi)$, so dass*

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \textit{oder } A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

Beweis. Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Aus $A^T A = E$ folgt

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Wegen $a^2 + c^2 = 1$ und $b^2 + d^2 = 1$ gibt es $\alpha, \alpha' \in [0, 2\pi)$, so dass

$$a = \cos\alpha, \quad c = \sin\alpha, \quad b = \sin\alpha', \quad d = \cos\alpha'.$$

Wegen $ab + cd = 0$ gilt

$$0 = \cos\alpha \cdot \sin\alpha' + \sin\alpha \cdot \cos\alpha' = \sin(\alpha + \alpha').$$

Also ist $\alpha + \alpha'$ entweder ein geradzahliges oder ein ungeradzahliges Vielfaches von π. Im ersten Fall ist

$$b = \sin\alpha' = -\sin\alpha \text{ und } d = \cos\alpha' = \cos\alpha$$

und im zweiten Fall ist

$$b = \sin\alpha' = \sin\alpha \text{ und } d = \cos\alpha' = -\cos\alpha.$$

□

Im ersten Fall ist die zugehörige orthogonale Abbildung eine Drehung um den Winkel α, im zweiten Fall eine Spiegelung an der um den Winkel $\alpha/2$ gedrehten x-Achse. Im ersten Fall ist A ist dann und nur dann diagonalisierbar, wenn $\alpha = 0$ oder $\alpha = \pi$ ist. Im zweiten Fall ist A diagonalisierbar mit den Eigenwerten $+1$ und -1.

c) Es sei nun $f : \mathbb{R}^3 \to \mathbb{R}^3$ eine orthogonale Abbildung. Dann hat das charakteristische Polynom einer Darstellungsmatrix von f den Grad 3. Also hat das charakteristische Polynom mindestens eine relle Nullstelle. Also hat f einen Eigenwert λ_1. Nach Satz 7.4 gilt $\lambda_1 = \pm 1$. Es sei $\mathbf{v}_1 \in \mathbb{R}^3$ ein Eigenvektor zu dem Eigenwert λ_1 mit $|\mathbf{v}_1| = 1$. Dann betrachten wir eine beliebige Basis des $\mathbb{R}^3$, die $\mathbf{v}_1$ als ersten Basisvektor enthält. Mit Hilfe des E. Schmidt'schen Orthonormalisierungsverfahrens können wir zu dieser Basis eine Orthonormalbasis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ des $\mathbb{R}^3$ finden. Es sei W die von $\mathbf{v}_2$ und $\mathbf{v}_3$ aufgespannte Ebene. Aus Satz 7.4 folgt, dass $f(W) = W$. Es sei $A := M_{\mathcal{B}}^{\mathcal{B}}(f)$. Dann gilt

$$A = \begin{pmatrix} \lambda_1 & 0 \quad 0 \\ 0 & \multirow{2}{*}{$\boxed{A'}$} \\ 0 & \end{pmatrix}.$$

Aus Korollar 7.1 folgt, dass A' eine orthogonale 2×2-Matrix ist. Außerdem gilt

$$\det A = \lambda_1 \det A'.$$

Nun müssen wir Fallunterscheidungen machen.

Es sei $\det A = +1$. Ist $\lambda_1 = -1$, so muss $\det A' = -1$ gelten. Damit hat A' die Eigenwerte $\lambda_2 = +1$ und $\lambda_3 = -1$ und wir können $\mathbf{v}_2$ als Eigenvektor zu λ_2 und $\mathbf{v}_3$ als Eigenvektor zu λ_3 wählen. Damit gilt

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

Ist $\lambda_1 = +1$, so muss auch $\det A' = +1$ sein. Also gibt es nach Satz 7.6 ein $\alpha \in [0, 2\pi)$, so dass

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

Ist $\det A = -1$, so ergeben sich die Möglichkeiten

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ und } A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Diese Überlegungen haben eine erstaunliche Konsequenz:

Satz 7.7 (Satz vom Fußball) *Bei jedem Fußballspiel, in dem nur ein Ball benutzt wird, gibt es zwei Punkte auf der Oberfläche des Balles, die sich zu Beginn der ersten und der zweiten Halbzeit (wenn der Ball genau auf dem Anstoßpunkt liegt) an der gleichen Stelle im umgebenden Raum befinden.*

Beweis. Die beiden Positionen des Balles unterscheiden sich durch eine orthogonale Abbildung f des Balles mit der Determinante $+1$. Eine solche Abbildung hat stets den Eigenwert $+1$, d.h. f lässt eine Gerade fest. Die Gerade schneidet die Oberfläche des Fußballs in zwei Punkten. □

8 Hauptachsentransformation

Wir haben bisher hauptsächlich lineare Gleichungen und Systeme von linearen Gleichungen betrachtet. Die linke Seite einer Gleichung

$$a_1 x_1 + \cdots + a_n x_n = b$$

ist eine Funktion in n Variablen und wird als *Linearform* bezeichnet. Wir wollen nun Funktionen untersuchen, in denen auch Quadrate und Produkte von Variablen vorkommen.

Die Funktion

$$q(\mathbf{x}) = ax^2 + 2bxy + cy^2, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix},$$

heißt eine *quadratische Form in zwei Variablen* x und y.

Eine quadratische Form

$$ax^2 + 2bxy + cy^2$$

können wir auch so schreiben

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

nennen wir die *zu der quadratischen Form gehörige Matrix.* Man beachte, dass dies eine symmetrische Matrix ist, d.h. es gilt $A^T = A$. Damit gilt

$$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}.$$

Man kann auch quadratische Formen mit beliebig vielen Variablen erklären.

Definition Eine *quadratische Form* in den n Variablen $x_1, \ldots, x_n$ ist eine Funktion

$$q(\mathbf{x}) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

wobei A eine symmetrische $n \times n$-Matrix ist.

Wir können eine quadratische Form kurz als

$$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

schreiben. Durch Ausmultiplizieren ergibt sich

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= a_{11}x_1^2 + \cdots a_{nn}x_n^2 + \sum_{i \neq j} a_{ij}x_ix_j. \end{aligned}$$

Hierbei bezeichnet

$$\sum_{i \neq j} a_{ij}x_ix_j$$

die Summe aller Produkte $a_{ij}x_ix_j$ mit $i \neq j$. Diese Terme heißen *gemischte Terme* der quadratischen Form.

Es gilt nun der folgende Satz.

Satz 8.1 *Für eine symmetrische Matrix A gilt*

(a) *A hat nur reelle Eigenwerte.*

(b) *A ist diagonalisierbar. Genauer gilt: Es gibt eine orthogonale Matrix S, so dass $S^T AS = S^{-1}AS$ Diagonalgestalt hat.*

(c) *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis. Wir beweisen nur (c). Die Beweise von (a) und (b) sind schwieriger und wir verzichten auf sie.

Beweis von (c): Es seien $\mathbf{v}_1$ und $\mathbf{v}_2$ Eigenvektoren von A zu verschiedenen Eigenwerten λ_1 und λ_2. Dann gilt

$$(A\mathbf{v}_1) \cdot \mathbf{v}_2 = \lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2.$$

Auf der anderen Seite gilt

$$(A\mathbf{v}_1) \cdot \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 = \mathbf{v}_1^T A^T \mathbf{v}_2 = \mathbf{v}_1 \cdot A^T \mathbf{v}_2 = \mathbf{v}_1 \cdot A\mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2.$$

Subtraktion der beiden Gleichungen ergibt

$$(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

Da nach Voraussetzung $\lambda_1 - \lambda_2 \neq 0$ ist, folgt $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. □

Wir können zwar keinen vollständigen Beweis von Satz 8.1 angeben, aber ein Verfahren zur Diagonalisierung symmetrischer Matrizen.

Verfahren zur Diagonalisierung symmetrischer Matrizen

Schritt 1. Man bestimme für jeden Eigenraum von A eine Basis.

Schritt 2. Man wende auf jede dieser Basen das E. Schmidt'sche Orthonormalisierungsverfahren an.

Schritt 3. Man bilde eine Matrix S, deren Spalten die in Schritt 2 berechneten Vektoren sind. Diese Matrix ist orthogonal und diagonalisiert A.

Dass diese Methode das gewünschte Ergebnis liefert, folgt aus Satz 8.1 (c): Die Matrix S ist orthogonal, da die Basisvektoren aus Eigenräumen zu verschiedenen Eigenwerten orthogonal sind.

Aus Satz 8.1 ergibt sich der folgende Satz.

Satz 8.2 (Hauptachsentransformation) *Es sei $\mathbf{x}^T A\mathbf{x}$ eine quadratische Form in den Variablen $x_1, \dots, x_n$. Dann gibt es eine orthogonale Matrix S, so dass gilt: Definieren wir neue Variablen $y_1, \dots, y_n$ durch $\mathbf{x} = S\mathbf{y}$, so erhalten wir*

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \cdots \lambda_n y_n^2,$$

wobei $\lambda_1, \dots, \lambda_n$ *die Eigenwerte von* A *sind und*

$$D = S^T A S = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Beispiel 8.1 Anstelle eines vollständigen Beweises von Satz 8.1 betrachten wir das Beispiel $n = 2$ und

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

Wir untersuchen die Eigenwerte und Eigenvektoren der Matrix A. Das charakteristische Polynom lautet

$$P_A(\lambda) = \lambda^2 - 2a\lambda + a^2 - b^2 = (\lambda - a + b)(\lambda - a - b).$$

Die Eigenwerte sind also $\lambda_1 = a + b$ und $\lambda_2 = a - b$. Als Eigenvektoren berechnet man

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Sind x', y' die Koordinaten des Vektors $\mathbf{x}$ bezüglich der Basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, so gilt

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}, \\ q(\mathbf{x}) &= (a+b)x'^2 + (a-b)y'^2. \end{aligned}$$

In den neuen Koordinaten x', y' ist also der gemischte Term mit $x'y'$ verschwunden. Im $\mathbb{R}^2$ mit den Koordinaten x', y' betrachten wir die Kurve

$$C = \{\mathbf{x} \in \mathbb{R}^2 \mid q(\mathbf{x}) = 1\}.$$

Wir nehmen an, dass $\lambda_1 > 0$ und $\lambda_2 \neq 0$. Dann setzen wir

$$\alpha = \frac{1}{\sqrt{\lambda_1}} \text{ und } \beta = \frac{1}{\sqrt{|\lambda_2|}}.$$

Damit lautet die Gleichung von C in den Koordinaten x', y':

$$\frac{x'^2}{\alpha^2} \pm \frac{y'^2}{\beta^2} = 1.$$

Im Fall + beschreibt diese Gleichung eine Ellipse, α und β sind gerade die *Hauptachsen* (vgl. Abbildung 4).

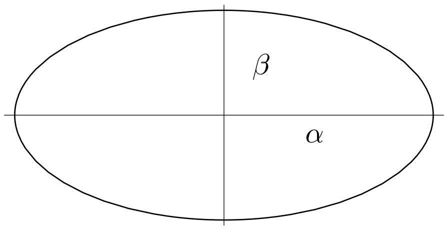


Abbildung 4: Ellipse $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$

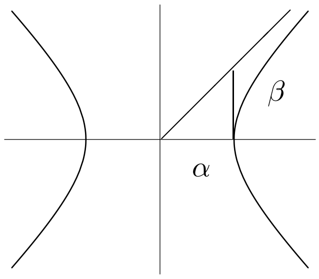


Abbildung 5: Hyperbel $\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1$

Im Fall $-$ können wir für die linke Seite der Gleichung auch schreiben

$$\frac{x'^2}{\alpha^2} - \frac{y'^2}{\beta^2} = (\frac{x'}{\alpha} + \frac{y'}{\beta})(\frac{x'}{\alpha} - \frac{y'}{\beta}).$$

Mit der Koordinatentransformation

$$\begin{aligned} \tilde{x} &= \frac{x'}{\alpha} + \frac{y'}{\beta}, \\ \tilde{y} &= \frac{x'}{\alpha} - \frac{y'}{\beta} \end{aligned}$$

schreibt sich die Gleichung von C als

$$\tilde{x}\tilde{y} = 1.$$

Diese Gleichung beschreibt eine Hyperbel, α und β sind wieder die *Hauptachsen* (vgl. Abbildung 5). Deswegen nennt man die obige Transformation auch *Hauptachsentransformation.*

Die Matrix

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

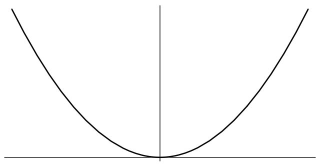


Abbildung 6: Parabel $x^2 - y = 0$

ist eine orthogonale Matrix. Wegen $\cos 45° = \frac{1}{\sqrt{2}}$ beschreibt S eine Drehung um $45°$. Das bedeutet, dass die durch

$$ax^2 + 2bxy + ay^2 = 1$$

gegebene Kurve nach einer Drehung des Koordinatensystems um $45°$ in Hauptachsenform vorliegt.

9 Kegelschnitte

Nun untersuchen wir Gleichungen der Gestalt

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0,$$

wobei $a, b, \ldots, f$ reelle Konstanten sind und mindestens eine der Zahlen a, b, c von Null verschieden ist. Eine solche Gleichung heißt *quadratische Gleichung* in x und y. Die Gleichung

$$ax^2 + 2bxy + cy^2$$

heißt die *zugehörige quadratische Form*.

Beispiel 9.1 (a) $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$, $\alpha, \beta > 0$. Diese Gleichung beschreibt eine *Ellipse* (vgl. Abbildung 4).

(b) $\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1$, $\alpha, \beta > 0$. Diese Gleichung beschreibt eine *Hyperbel* (vgl. Abbildung 5).

(c) $x^2 - \alpha y = 0$. Diese Gleichung beschreibt eine *Parabel* (vgl. Abbildung 6).

Die Kurven, die durch eine quadratische Gleichung in x und y beschrieben werden, heißen *Kegelschnitte*. Dies wollen wir nun begründen. Dazu betrachten wir einen Doppelkegel

$$K = \{(x', y', z') \in \mathbb{R}^3 \mid x'^2 + y'^2 = z'^2\}.$$

Diesen Kegel drehen wir nun um einen Winkel φ um die y'-Achse. Das bedeutet, dass wir die folgende Koordinatentransformation vornehmen:

$$\begin{aligned} x' &= (\cos\varphi)x + (\sin\varphi)z, \\ y' &= y, \\ z' &= (-\sin\varphi)x + (\cos\varphi)z. \end{aligned}$$

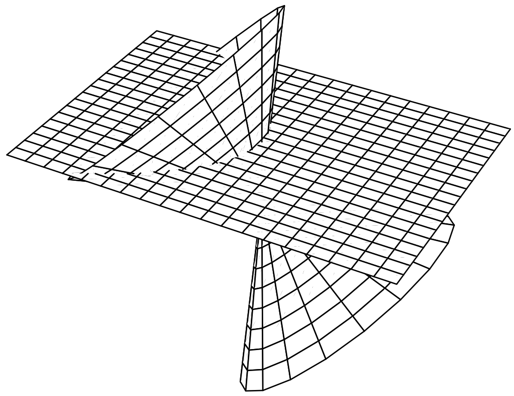


Abbildung 7: Die Parabel als Kegelschnitt

Setzt man dies in die Kegelgleichung ein, so erhält man

$$(\cos^2\varphi - \sin^2\varphi)x^2 + (4\cos\varphi\sin\varphi)zx + y^2 = (\cos^2\varphi - \sin^2\varphi)z^2.$$

Den gedrehten Kegel schneiden wir nun mit der Ebene

$$E = \{(x, y, z) \in \mathbb{R}^3 \mid z = 1\}.$$

Dann wird $K \cap E$ durch die Gleichung

$$(\cos^2\varphi - \sin^2\varphi)x^2 + (4\cos\varphi\sin\varphi)x + y^2 = (\cos^2\varphi - \sin^2\varphi)$$

in den Koordinaten (x, y) beschrieben.

Wir betrachten nun zunächst den Winkel $\varphi = \frac{\pi}{4}$. Dann reduziert sich die Gleichung zu

$$y^2 + 2x = 0.$$

Dies ist die Gleichung einer Parabel.

Für $0 \leq \varphi \leq \frac{\pi}{4}$ ist $a := \cos^2\varphi - \sin^2\varphi > 0$. Wir erhalten die Gleichung

$$ax^2 + (4\cos\varphi\sin\varphi)x + y^2 = a.$$

Durch quadratische Ergänzung erhalten wir

$$a\left(x + \frac{2}{a}\cos\varphi\sin\varphi\right)^2 + y^2 = \frac{1}{a}.$$

Nun führen wir die Koordinatentranslation

$$x' = x + \frac{2}{a}\cos\varphi\sin\varphi, \quad , y' = y$$

durch. Dies führt auf die Gleichung

$$\frac{x'^2}{\alpha^2} + \frac{y'^2}{\beta^2} = 1 \quad \text{mit } \alpha = \frac{1}{a}, \beta = \frac{1}{\sqrt{a}}.$$

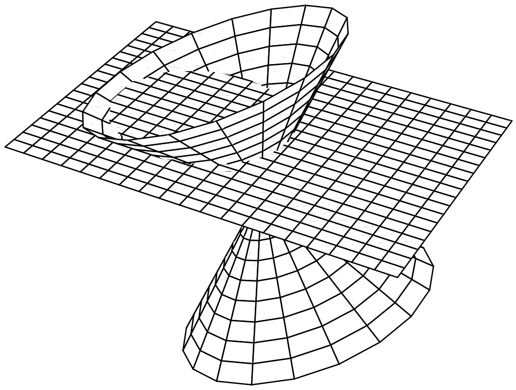


Abbildung 8: Die Ellipse als Kegelschnitt

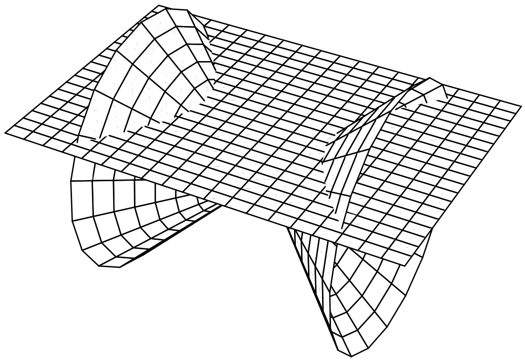


Abbildung 9: Die Hyperbel als Kegelschnitt

Dies ist die Gleichung einer Ellipse.

Für $\frac{\pi}{4} \leq \varphi \leq \frac{\pi}{2}$ ist $a := \cos^2 \varphi - \sin^2 \varphi > 0$ und eine analoge Rechnung führt auf die Gleichung

$$\frac{x'^2}{\alpha^2} - \frac{y'^2}{\beta^2} = 1 \quad \text{mit } \alpha = \frac{1}{a}, \beta = \frac{1}{\sqrt{a}}.$$

Dies ist die Gleichung einer Hyperbel.

Wir wollen nun eine quadratische Gleichung auf eine möglichst einfache Gestalt bringen. Aus Satz 8.2 folgt dass wir mit einer Koordinatentransformation den gemischten Term mit xy eliminieren können. Diese Koordinatentransformation erfolgt mit einer orthogonalen 2×2-Matrix S. Indem man eventuell die Spalten von S vertauscht, kann man annehmen, dass S eine Drehung beschreibt.

Kommt nun in der Gleichung ein Term mit x^2 und ein Term mit x vor, so können wir den Term mit x durch *quadratische Ergänzung* und eine anschließende Translation des Koordinatensystems eliminieren.

Beispiel 9.2 Wir betrachten die quadratische Gleichung

$$2x^2 - 12x + 4y + 26 = 0.$$

Diese Gleichung ist äquivalent zu

$$2(x^2 - 6x) + 4y + 26 = 0.$$

Durch quadratische Ergänzung der Klammer ergibt sich

$$2(x^2 - 6x + 9) + 4y + 26 - 18 = 0$$

oder

$$2(x-3)^2 + 4(y+2) = 0$$

oder

$$(x-3)^2 + 2(y+2) = 0.$$

In den Koordinaten

$$x' = x - 3, \quad y' = y + 2$$

schreibt sich diese Gleichung als

$$x'^2 + 2y' = 0.$$

Die Koordinatentransformation

$$x' = x - 3, \quad y' = y + 2$$

bedeutet eine Verschiebung (Translation) des Koordinatensystems.

Entsprechend können wir verfahren, wenn in der Gleichung Terme mit y^2 und y vorkommen. Wir sehen an dem Beispiel auch, dass wir mit einem Term mit x oder y die Konstante eliminieren können.

Zusammenfassend erhalten wir den folgenden Satz.

Satz 9.1 *Eine quadratische Gleichung*

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0,$$

die einen nicht leeren Kegelschnitt beschreibt und bei der mindestens eine der Zahlen a, b, c von Null verschieden ist, kann durch Drehungen und Translationen des Koordinatensystems auf eine der folgenden Normalformen gebracht werden:

	Gleichung	*Graph*
(a)	$x^2 = 0$	*(Doppel-)Gerade*
(b)	$x^2 + \alpha y^2 = 0,\ \alpha > 0$	*Punkt*
(c)	$x^2 - \alpha y^2 = 0,\ \alpha > 0$	*Geradenpaar*
(d)	$x^2 + \alpha y = 0,\ \alpha \neq 0$	*Parabel*
(e)	$\frac{x^2}{\alpha^2} = 1,\ \alpha > 0$	*Zwei parallele Geraden*
(f)	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1,\ \alpha, \beta > 0$	*Ellipse, Kreis*
(g)	$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1,\ \alpha, \beta > 0$	*Hyperbel*

Beweis. Wir reduzieren die Gleichung schrittweise zu einer der obigen Normalformen.

Schritt 1. Wir eliminieren den gemischten Term mit xy durch Hauptachsentransformation.

Schritt 2. Durch quadratische Ergänzung und Koordinatentranslation eliminieren wir den Term mit x, falls ein Term mit x^2 vorkommt, und den Term mit y, falls ein Term mit y^2 vorkommt.

Schritt 3. Bleibt ein Term mit x oder y stehen, so eliminieren wir mit einer weiteren Translation des Koordinatensystems den konstanten Term.

Damit können wir die Gleichung auf eine der folgenden drei Formen bringen

$$\begin{aligned} a'x'^2 + c'y'^2 + f' &= 0, \quad a' \neq 0 \text{ oder } c' \neq 0, \\ a'x'^2 + e'y' &= 0, \quad a', e' \neq 0, \\ c'y'^2 + d'x' &= 0, \quad c', d' \neq 0. \end{aligned}$$

Durch eine Drehung des Koordinatensystems um 90° können wir die dritte Gleichung auf die zweite zurückführen. Die Liste von Satz 9.1 ergibt sich nun, indem wir Fallunterscheidungen für die Koeffizienten a', c', e', f' durchführen. □

10 Quadriken

Wir betrachten nun quadratische Gleichungen in drei Variablen. Eine allgemeine *quadratische Gleichung* in x, y und z hat die Form

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + j = 0,$$

wobei $a, b, \ldots, f$ nicht alle Null sind. Der Ausdruck

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz$$

ist die *zugehörige quadratische Form.*

Der Graph einer quadratischen Gleichung in x, y, z heißt *Quadrik.* Die Gleichung hat dann die einfachste Form, wenn die Fläche sich in einer bestimmten Standardlage bezüglich der Koordinatenachsen befindet. Es gibt insgesamt sechs Typen von (nicht ausgearteten) Quadriken, die in den Abbildungen 10 bis 15 angegeben sind.

Beispiel 10.1 Wir wollen den Graphen der Gleichung

$$x^2 + y^2 + 2xy - 4\sqrt{2}xz + 4\sqrt{2}yz + 4x + 4y - 8\sqrt{2}z - 28 = 0$$

bestimmen. Dazu gehen wir in zwei Schritten vor:

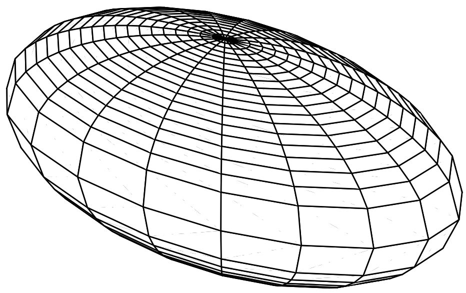


Abbildung 10: Ellipsoid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 1$

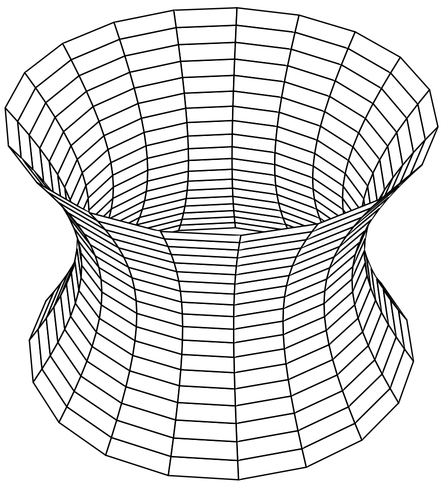


Abbildung 11: Einschaliges Hyperboloid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 1$

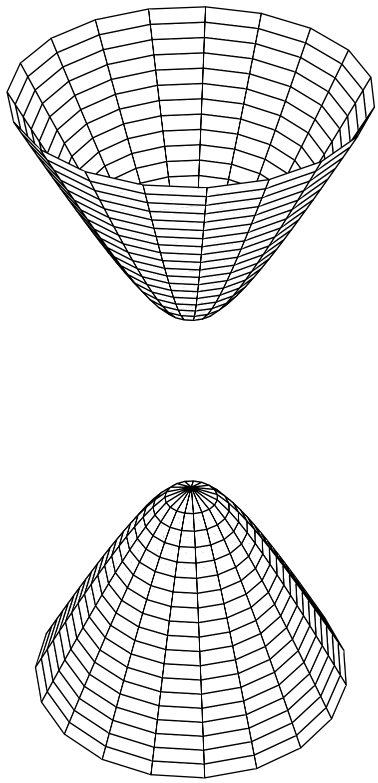


Abbildung 12: Zweischaliges Hyperboloid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = -1$

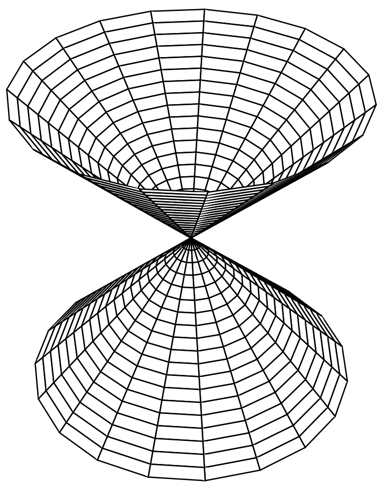


Abbildung 13: Elliptischer Kegel $z^2 = \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$

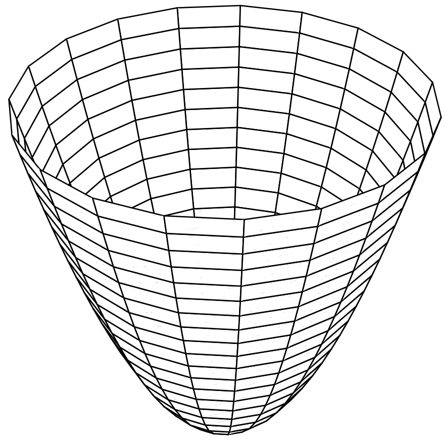


Abbildung 14: Elliptisches Paraboloid $z = \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$

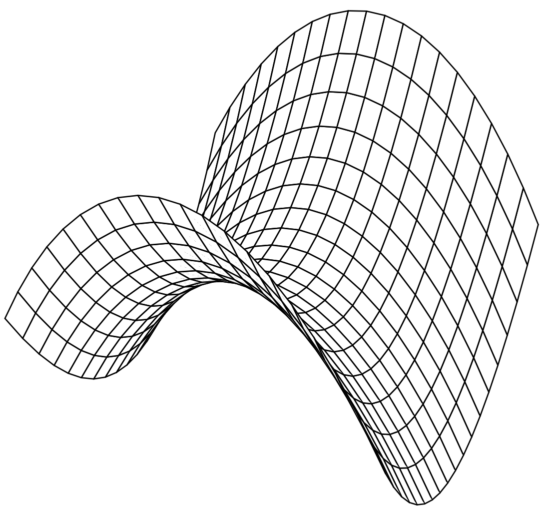


Abbildung 15: Hyperbolisches Paraboloid $z = \frac{y^2}{\beta^2} - \frac{x^2}{\alpha^2}$

1) Wir bestimmen zunächst eine orthogonale Matrix S, die die zugehörige quadratische Form auf Hauptachsengestalt bringt. Die zugehörige quadratische Form hat die Matrixform $\mathbf{x}^T A\mathbf{x}$ mit

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

Das charakteristische Polynom dieser Matrix lautet

$$\begin{aligned} P_A(\lambda) &= \begin{vmatrix} 1-\lambda & 1 & -2\sqrt{2} \\ 1 & 1-\lambda & 2\sqrt{2} \\ -2\sqrt{2} & 2\sqrt{2} & -\lambda \end{vmatrix} \\ &= -\lambda^3 + 2\lambda^2 + 16\lambda - 32 = (-\lambda+2)(\lambda+4)(\lambda-4). \end{aligned}$$

Die Eigenwerte von A sind also $\lambda_1 = 2$, $\lambda_2 = 4$ und $\lambda_3 = -4$. Eine Lösung von

$$\begin{pmatrix} -1 & 1 & -2\sqrt{2} \\ 1 & -1 & 2\sqrt{2} \\ -2\sqrt{2} & 2\sqrt{2} & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist der Vektor

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$\mathbf{v}_1$ ist also ein Eigenvektor zum Eigenwert $\lambda_1 = 2$. Eine Lösung des Gleichungssystems

$$\begin{pmatrix} -3 & 1 & -2\sqrt{2} \\ 1 & -3 & 2\sqrt{2} \\ -2\sqrt{2} & 2\sqrt{2} & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist der Vektor

$$\mathbf{v}_2 = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -\sqrt{2} \end{pmatrix},$$

$\mathbf{v}_2$ ist also ein Eigenvektor zum Eigenwert $\lambda_2 = 4$. Entsprechend sieht man, dass der Vektor

$$\mathbf{v}_3 = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ -\sqrt{2} \end{pmatrix}$$

ein Eigenvektor zum Eigenwert $\lambda_3 = -4$ ist. Also diagonalisiert die Matrix

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

die Matrix A. Definieren wir die neuen Variablen x', y', z' durch $\mathbf{x} = S\mathbf{x}'$, so lautet die ursprüngliche Gleichung in den neuen Koordinaten x', y', z':

$$2x'^2 + 4y'^2 - 4z'^2 + 4\sqrt{2}x' + 8y' + 8z' - 28 = 0.$$

2) Nun versuchen wir, mit Hilfe von quadratischer Ergänzung die linearen Terme zu eliminieren. Quadratische Ergänzung liefert:

$$2(x' + \sqrt{2})^2 + 4(y' + 1)^2 - 4(z' - 1)^2 - 32 = 0.$$

Mit der Koordinatentransformation

$$\tilde{x} := x' + \sqrt{2}, \quad , \tilde{y} := y' + 1, \quad \tilde{z} := z' - 1$$

lautet diese Gleichung

$$2\tilde{x}^2 + 4\tilde{y}^2 - 4\tilde{z}^2 = 32$$

oder

$$\frac{\tilde{x}^2}{16} + \frac{\tilde{y}^2}{8} - \frac{\tilde{z}^2}{8} = 1.$$

Dies ist die Gleichung eines einschaligen Hyperboloids in Standardlage.

11 Quadratische Formen

Wir betrachten nun quadratische Formen in n Variablen. Es sei A eine symmetrische $n \times n$-Matrix und $\mathbf{x} \in \mathbb{R}^n$.

Definition Eine quadratische Form $\mathbf{x}^T A\mathbf{x}$ heißt *positiv definit*, wenn gilt:

$$\mathbf{x}^T A\mathbf{x} > 0 \text{ für alle } \mathbf{x} \neq \mathbf{0}.$$

Entsprechend heißt $\mathbf{x}^T A\mathbf{x}$ *negativ definit*, wenn $\mathbf{x}^T A\mathbf{x} < 0$ für alle $\mathbf{x} \neq \mathbf{0}$ gilt.

Eine quadratische Form $\mathbf{x}^T A\mathbf{x}$ heißt *indefinit*, wenn $\mathbf{x}^T A\mathbf{x}$ sowohl positive als auch negative Werte annimmt.

Satz 11.1 *Eine quadratische Form $\mathbf{x}^T A\mathbf{x}$ ist genau dann positiv definit, wenn die Matrix A nur positive Eigenwerte besitzt.*

Beweis. a) Es sei $\mathbf{x}^T A\mathbf{x}$ positiv definit. Es sei λ ein Eigenwert von A und $\mathbf{x}$ ein zugehöriger Eigenvektor. Dann ist $\mathbf{x} \neq \mathbf{0}$ und $A\mathbf{x} = \lambda\mathbf{x}$, also

$$0 < \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T \lambda\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x} = \lambda|\mathbf{x}|^2.$$

Wegen $|\mathbf{x}|^2 > 0$ folgt $\lambda > 0$.

b) Es sei nun umgekehrt A eine Matrix mit positiven Eigenwerten λ_1, $\ldots$, λ_n. Nach Satz 8.2 gibt es eine orthogonale Matrix S, so dass gilt: Definieren wir neue Variablen $y_1, \ldots, y_n$ durch $\mathbf{x} = S\mathbf{y}$, so erhalten wir

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \cdots \lambda_n y_n^2.$$

Es gilt aber

$$\lambda_1 y_1^2 + \cdots \lambda_n y_n^2 > 0$$

für alle $\mathbf{y} \neq \mathbf{0}$ und $\mathbf{y} \neq \mathbf{0}$ genau dann, wenn $\mathbf{x} \neq \mathbf{0}$. □

Wir geben nun für quadratische Formen in zwei Variablen ein Kriterium an, mit dem man entscheiden kann, ob eine quadratische Form positiv definit, negativ definit oder indefinit ist, ohne die Eigenwerte der Matrix A auszurechnen.

Lemma 11.1 *Es sei $n = 2$ und*

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Dann ist die quadratische Form $\mathbf{x}^T A\mathbf{x}$ genau dann positiv definit, wenn $a > 0$ und $\det A = ac - b^2 > 0$. Entsprechend ist $\mathbf{x}^T A\mathbf{x}$ genau dann negativ definit, wenn $a < 0$ und $\det A = ac - b^2 > 0$. Die quadratische Form $\mathbf{x}^T A\mathbf{x}$ ist genau dann indefinit, wenn $\det A = ac - b^2 < 0$.

Beweis. Es gilt

$$\mathbf{x}^T A\mathbf{x} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2.$$

Wir ergänzen die quadratische Form und schreiben

$$\mathbf{x}^T A\mathbf{x} = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2.$$

Angenommen, die quadratische Form $\mathbf{x}^T A\mathbf{x}$ ist positiv definit. Setzen wir $y = 0$, so sehen wir $a > 0$. Setzen wir $x = -(b/a)y$, so sehen wir $c - b^2/a > 0$ oder $ac - b^2 > 0$.

Es sei umgekehrt $a > 0$ und $c - b^2/a > 0$. Dann ist $\mathbf{x}^T A\mathbf{x}$ eine Summe von Quadraten, also gilt $\mathbf{x}^T A\mathbf{x} \geq 0$. Ist $\mathbf{x}^T A\mathbf{x} = 0$, dann müssen auch die beiden Quadrate Null sein. Daraus folgt aber $x = 0$ und $y = 0$. Also ist die quadratische Form $\mathbf{x}^T A\mathbf{x}$ positiv definit.

Die anderen Aussagen folgen analog. □

Beispiel 11.1 Für $n = 2$ ist die quadratische Form $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2}$ positiv definit, die quadratische Form $-\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2}$ negativ definit und die quadratische Form $\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2}$ indefinit.

Für $n = 3$ ist die quadratische Form $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2}$ positiv definit und die quadratische Form $-\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2}$ negativ definit. Die quadratischen Formen $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2}$ und $-\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2}$ sind indefinit.

Lemma 11.1 lässt sich wie folgt verallgemeinern. Für eine quadratische Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

bezeichnet man die Determinanten der Untermatrizen

$$A_1 = (a_{11}), \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \dots,$$

$$A_n = A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

die aus den ersten r Zeilen und r Spalten von A (für $r = 1, \dots, n$) gebildet werden, als die *Hauptminoren.* Damit gilt der folgende Satz, den wir nicht beweisen wollen.

Satz 11.2 (Hurwitzkriterium) *Eine quadratische Form $\mathbf{x}^T A\mathbf{x}$ ist genau dann positiv definit, wenn alle Hauptminoren der Matrix A positiv sind.*

Literatur

[1] H. Anton: Lineare Algebra. Spektrum Akademischer Verlag, Heidelberg Berlin, 1998, ISBN 3-8274-0324-3.

[2] G. Fischer: Lineare Algebra. 12., verb. Aufl., Vieweg Verlag, Wiesbaden, 2000, ISBN 3-528-87217-9.

Inhaltsverzeichnis