INGENIEURMATHEMATIK I

LINEARE ALGEBRA

L. G. LUCHT

Technische Universität Clausthal

Wintersemester 2000/2001

Inhalt

10. Vektorräume

10.1. **Definition und grundlegende Eigenschaften.** Vom anschaulichen Beispiel der Ortsvektoren in der Ebene oder im Raum gelangt man durch Abstraktion zu einer axiomatisch definierten Rechenstruktur, dem *Vektorraum.* Dabei bleibt die Natur der Rechengrößen und der sie verknüpfenden Rechenoperationen offen. Kennzeichnend für die Rechenstruktur sind allein die in den Axiomen formulierten Rechenregeln.

Gegeben sei eine Menge X, deren Elemente *Vektoren* genannt und hier mit $\vec{a}, \vec{b}, \ldots, \vec{z}$ bezeichnet werden. Ferner seien in X die beiden folgenden *lineare Operationen* definiert:

Die *Addition* ordnet je zwei Vektoren $\vec{x}, \vec{y} \in X$ eindeutig einen mit $\vec{x} + \vec{y}$ bezeichneten Vektor aus X (*Summenvektor*) zu.

Die *Multiplikation mit Skalaren* ordnet je einer reellen Zahl c (*Skalar*) und einem Vektor $\vec{x} \in X$ einen mit $c\vec{x}$ bezeichneten Vektor aus X zu.

Definition 10.1. *Eine mit den linearen Operationen versehene Menge X heißt ein Vektorraum, wenn folgende Axiome erfüllt sind:*

(10.1) $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$ *(Assoziativität).*

(10.2) $\vec{x} + \vec{y} = \vec{y} + \vec{x}$ *(Kommutativität).*

(10.3) *Es gibt mindestens einen Vektor $\vec{0} \in X$ mit $\vec{x} + \vec{0} = \vec{x}$ für alle $\vec{x} \in X$.*

(10.4) *Zu jedem $\vec{x} \in X$ existiert mindestens ein Vektor $\vec{x}\,' \in X$ mit $\vec{x} + \vec{x}\,' = \vec{0}$.*

(10.5) $(cd)\vec{x} = c(d\vec{x})$ *(Assoziativität).*

(10.6) $c(\vec{x} + \vec{y}) = c\vec{x} + c\vec{y}, \quad (c + d)\vec{x} = c\vec{x} + d\vec{x}$ *(Distributivität).*

(10.7) $1\vec{x} = \vec{x}$.

Wegen (10.1) kommt es in mehrgliedrigen Summen von Vektoren nicht auf die Art der Klammersetzung an, so daß weiterhin die Klammern fortgelassen werden können. In (10.6) steckt schon die übliche Konvention, daß die Multiplikation (mit Skalaren) Vorrang hat vor der Addition.

Beispiel 10.2. (a) Die von einem festen Anfangspunkt ausgehenden Ortsvektoren (gerichteten Strecken) der Ebene oder des Raumes bilden je einen Vektorraum unter den üblichen linearen Operationen. Dabei werden Vektoren nach dem Parallelogramm-Prinzip addiert.

(b) Die Menge $\mathbb{R}^n$ aller n-Tupel reeller Zahlen ist ein Vektorraum, wenn man für n-Tupel $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$ mit Skalaren c die linearen Operationen durch

$$\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n) \quad \text{und} \quad c\vec{x} = (cx_1, \ldots, cx_n)$$

definiert. Speziell für $n = 1$ bildet also auch $\mathbb{R}$ unter den üblichen Addition und Multiplikation einen Vektorraum.

(c) Es sei F die Menge aller Funktionen $f : \mathbb{R} \to \mathbb{R}$ oder die der Folgen $f : \mathbb{N}_0 \to \mathbb{R}$ mit den linearen Operationen

$$(f+g)(t) = f(t) + g(t) \quad \text{und} \quad (cf)(t) = c\,f(t)\,.$$

Dann bildet F ein Vektorraum (Funktionenraum bzw. Folgenraum).

Satz 10.3. (a) *Der Vektor $\vec{0}$ aus Axiom* (10.3) *ist eindeutig bestimmt. Er heißt der Nullvektor in X.*

(b) *Der Vektor $\vec{x}\,'$ aus* (10.4) *ist eindeutig bestimmt durch den Vektor $\vec{x}$. Er heißt der zu $\vec{x}$ entgegengesetzte (oder negative) Vektor und wird mit $-\vec{x}$ bezeichnet.*

(c) *Bei gegebenen Vektoren $\vec{a}, \vec{b} \in X$ hat die Gleichung $\vec{a} + \vec{x} = \vec{b}$ genau eine Lösung $\vec{x} \in X$, nämlich den Differenzvektor $\vec{x} = \vec{b} - \vec{a} := \vec{b} + \vec{-a}$. Hierdurch wird die Subtraktion in X definiert.*

In Beispiel 10.2 ist der durch den Anfangspunkt repräsentierte (entartete) Ortsvektor der Nullvektor. Das n-Tupel $(0, \ldots, 0)$ ist der Nullvektor des $\mathbb{R}^n$. Der Nullvektor des Funktionenraumes ist die durch

$$f(t) = 0 \quad \text{für alle } t$$

definierte *Nullfunktion*.

Folgerung 10.4. *Aus $\vec{a} + \vec{x} = \vec{a}$ für nur einen Vektor $\vec{a}$ folgt bereits $\vec{x} = \vec{0}$. Genau dann gilt $c\vec{x} = \vec{0}$, wenn $c = 0$ oder $\vec{x} = \vec{0}$ gilt. Für alle $\vec{x} \in X$ gilt $(-1)\vec{x} = -\vec{x}$.*

10.2. **Unterräume.**

Definition 10.5. *Eine Teilmenge $U \subseteq X$ heißt Unterraum von X, wenn die folgenden Bedingungen erfüllt sind:*

(a) *$U \neq \emptyset$ (leere Menge).*

(b) *Aus $\vec{x}, \vec{y} \in U$ folgt $\vec{x} + \vec{y} \in U$.*

(c) *Aus $\vec{x} \in U$ folgt $c\vec{x} \in U$ für jeden beliebigen Skalar c.*

Ein Unterraum von X ist also stets eine unter den linearen Operationen von X abgeschlossene nicht-leere Teilmenge von X.

Satz 10.6. *Jeder Unterraum von X ist unter den linearen Operationen aus X selbst wieder ein Vektorraum.*

Ganz X ist ein Unterraum von sich selbst. Auch die nur aus dem Nullvektor bestehende Teilmenge $\{\vec{0}\}$ ist ein Unterraum von X (Null-Raum).

Satz 10.7. *Der Durchschnitt beliebig vieler Unterräume von X ist wieder ein Unterraum von X.*

Nach diesem Satz macht die folgende Definition Sinn.

Definition 10.8. *Es sei $M \subseteq X$. Dann heißt*

$$[M] := \bigcap \{U : U \ \textit{Unterraum von}\ X \ \textit{mit}\ M \subseteq U\}$$

der von M erzeugte (oder aufgespannte) Unterraum.

Es ist also $[M]$ der kleinste Unterraum von X, der M als Teilmenge enthält. Speziell ergibt sich $[\emptyset] = \{\vec{0}\}$, und genau dann ist M Unterraum von X, wenn $[M] = M$ gilt.

Definition 10.9. *Jeder Vektor der Form*

$$c_1\vec{a}_1 + \cdots + c_n\vec{a}_n$$

mit beliebigen Skalaren $c_1, \dots, c_n$ wird eine Linearkombination der Vektoren $\vec{a}_1, \dots, \vec{a}_n$ genannt. Unter einer Linearkombination einer nicht-leeren Teilmenge M von X versteht man eine Linearkombination von endlich vielen Vektoren aus M.

Satz 10.10. *Es sei M eine nicht-leere Teilmenge von X. Dann ist $[M]$ genau die Menge aller Linearkombinationen von M.*

Es seien U und V zwei Unterräume von X. Dann ist die Menge aller Vektoren der Form $\vec{u} + \vec{v}$ mit $\vec{u} \in U$ und $\vec{v} \in V$ wieder ein Unterraum von X, der der *Summenraum* von U und V genannt und mit $U + V$ bezeichnet wird:

$$U + V := \{\vec{u} + \vec{v} : \vec{u} \in U \ \text{und}\ \vec{v} \in V\}.$$

10.3. Basis und Dimension.

Definition 10.11. *Vektoren $\vec{a}_1, \dots, \vec{a}_n \in X$ werden linear unabhängig genannt, wenn aus $c_1\vec{a}_1 + \cdots + c_n\vec{a}_n = \vec{0}$ stets $c_1 = \ldots = c_n = 0$ folgt. Andernfalls heißen die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ linear abhängig. Eine Teilmenge M von X heißt linear unabhängig, wenn je endlich viele Vektoren aus M linear unabhängig sind.*

Jede Teilmenge einer linear unabhängigen Menge ist wieder linear unabhängig; jede Obermenge einer linear abhängigen Menge ist selbst linear abhängig. Die leere Menge ist linear unabhängig, die Menge $\{\vec{0}\}$ jedoch linear abhängig.

Satz 10.12. *Die nur aus dem Vektor $\vec{x}$ bestehende Menge $\{\vec{x}\}$ ist genau dann linear abhängig, wenn $\vec{x} \neq \vec{0}$ gilt. Die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ sind genau dann linear abhängig, wenn sich unter ihnen mindestens einer als Linearkombination der übrigen Vektoren darstellen läßt.*

Definition 10.13. *Eine Teilmenge B von X heißt eine Basis des Vektorraumes X, wenn sie folgende Eigenschaften besitzt:*

(a) *B ist linear unabhängig.*

(b) *B spannt den Raum X auf, d.h. $[B] = X$.*

Die leere Menge ist die einzige Basis des Nullraumes. Im $\mathbb{R}^n$ bilden die n-Tupel

$$\vec{e}_1 = (1,0,\dots,0),\ \vec{e}_2 = (0,1,\dots,0),\ \dots\ ,\ \vec{e}_n = (0,0,\dots,1)$$

eine Basis, die auch die *kanonische Basis* des $\mathbb{R}^n$ genannt wird. Allgemein kann man beweisen, daß jeder Vektorraum mindestens eine Basis besitzt.

Satz 10.14. *Es sei $X \neq \{\vec{0}\}$. Dann sind folgende Aussagen äquivalent:*

(a) *B ist eine Basis von X.*

(b) *B ist eine minimale Teilmenge von X mit $[B] = X$.*

(c) *B ist eine maximale linear unabhängige Teilmenge von X.*

(d) *Jeder Vektor aus X kann auf genau eine Weise als Linearkombination von B dargestellt werden.*

Satz 10.15. *Es sei $\{\vec{a}_1,\dots,\vec{a}_n\}$ eine Basis von X, und es gelte*

$$\vec{b} = b_1\vec{a}_1 + \cdots + b_k\vec{a}_k + \cdots + b_n\vec{a}_n \quad \textit{mit} \quad b_k \neq 0.$$

Dann ist auch $\{\vec{a}_1,\dots,\vec{a}_{k-1},\vec{b},\vec{a}_{k+1},\dots,\vec{a}_n\}$ eine Basis von X.

Satz 10.16. *Es sei $\{\vec{a}_1,\dots,\vec{a}_n\}$ eine Basis von X, und die Vektoren $\vec{b}_1,\dots,\vec{b}_k \in X$ seien linear unabhängig. Dann gilt $k \leq n$, und bei geeigneter Numerierung der Basisvektoren ist auch $\{\vec{b}_1,\dots,\vec{b}_k,\vec{a}_{k+1},\dots,\vec{a}_n\}$ eine Basis von X.*

Es folgt: Je zwei Basen eines Vektorraumes X bestehen entweder beide aus unendlich vielen Vektoren, oder sie enthalten beide dieselbe Anzahl von Vektoren. Die somit allen Basen von X gemeinsame (evtl. unendliche) Anzahl von Vektoren wird die *Dimension* von X genannt.

Definition 10.17. *Man setzt $\dim\{\vec{0}\} = 0$ sowie $\dim X = \infty$, wenn X keine endliche Basis besitzt. Andernfalls bezeichnet $\dim X$ die Anzahl der Vektoren einer beliebigen Basis.*

Insbesondere gilt hiernach $\dim \mathbb{R}^n = n$.

Satz 10.18. *Es sei U ein Unterraum von X. Dann gilt $\dim U \leq \dim X$. Besitzt X endliche Dimension, so folgt aus $\dim U = \dim X$ sogar $U = X$. Jede Basis von U kann zu einer Basis von X verlängert werden.*

Satz 10.19. *Es seien U und V endlich-dimensionale Unterräume von X. Dann gilt*

$$\dim U + \dim V = \dim(U+V) + \dim(U \cap V).$$

11. Lineare Abbildungen

11.1. **Koordinaten.** Es sei X ein Vektorraum der endlichen Dimension n und $\{\vec{a}_1, \ldots, \vec{a}_n\}$ eine fest gewählte Basis von X. Dann besitzt jeder Vektor $\vec{x} \in X$ eine eindeutig bestimmte Basisdarstellung

$$\vec{x} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n \,.$$

Die durch den Vektor $\vec{x}$ und die Basis eindeutig bestimmten Skalare $x_1, \ldots, x_n$ heißen die *Koordinaten* von $\vec{x}$ bezüglich der Basis $\{\vec{a}_1, \ldots, \vec{a}_n\}$. Umgekehrt bestimmt jedes Zahlen-n-Tupel eindeutig einen Vektor $\vec{x} \in X$ mit dem entsprechenden Koordinaten-n-Tupel. Hinsichtlich der Basis $\{\vec{a}_1, \ldots, \vec{a}_n\}$ ist demnach eine bijektive Abbildung

$$X \;\leftrightarrow\; \mathbb{R}^n \quad \text{mit} \quad \vec{x} \;\leftrightarrow\; (x_1, \ldots, x_n)$$

gegeben.

Satz 11.1. *Aus* $\vec{x} \;\leftrightarrow\; (x_1, \ldots, x_n)$ *und* $\vec{y} \;\leftrightarrow\; (y_1, \ldots, y_n)$ *folgt*

$$\vec{x} + \vec{y} \;\leftrightarrow\; (x_1 + y_1, \ldots, x_n + y_n)\,, \quad c\vec{x} \;\leftrightarrow\; (cx_1, \ldots, cx_n).$$

Die linearen Operationen in X gehen also bei dieser Beziehung in die entsprechenden Operationen der Koordinaten-n-Tupel (aufgefaßt als Vektoren des $\mathbb{R}^n$) über.

Satz 11.2. *Genau dann sind* $\vec{x}_1, \ldots, \vec{x}_k \in X$ *linear unabhängig, wenn die entsprechenden Koordinaten-n-Tupel, aufgefaßt als Vektoren des* $\mathbb{R}^n$*, linear unabhängig sind.*

Gegeben seien k Vektoren $\vec{x}_1, \ldots, \vec{x}_k \in X$. Gesucht ist die Maximalzahl linear unabhängiger Vektoren unter ihnen, also die Dimension des von $\vec{x}_1, \ldots, \vec{x}_k$ aufgespannten Unterraums.

Es bedeute $(x_{\kappa,1}, \ldots, x_{\kappa,n})$ das Koordinaten-n-Tupel des Vektors $\vec{x}_\kappa$ $(\kappa = 1, \ldots, k)$. Schreibt man diese Koordinaten-n-Tupel untereinander, so erhält man ein rechteckiges Zahlenschema, das als *Matrix* bezeichnet wird:

$$\begin{pmatrix} x_{1,1} & x_{1,2} & \ldots & x_{1,n} \\ x_{2,1} & x_{2,2} & \ldots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k,1} & x_{k,2} & \ldots & x_{k,n} \end{pmatrix}. \tag{1}$$

Sind Verwechslungen nicht möglich, schreiben wir die Doppelindizes in den Matrixelementen auch ohne Komma, also kurz $x_{\kappa\nu}$ statt $x_{\kappa,\nu}$.

Gleichwertig mit der oben gestellten Frage ist: Welches ist die Maximalzahl linear unabhängiger Zeilen dieser Matrix?

Unterwirft man die Matrix (1) in geeigneter Weise den folgenden *elementaren Umformungen*

(1) Vertauschung zweier Zeilen,

(2) Vertauschung zweier Spalten,

(3) Ersetzung der i-ten Zeile durch die Summe der i-ten und der mit einer Zahl c multiplizierten k-ten Zeile $(i \neq k)$,

so kann man stets zu einer Matrix der folgenden Form gelangen:

$$
\begin{pmatrix}
x'_{1,1} & x'_{1,2} & \dots & x'_{1,r-1} & x'_{1,r} & \dots & x'_{1,n} \\
0 & x'_{2,2} & \dots & x'_{2,r-1} & x'_{2,r} & \dots & x'_{2,n} \\
0 & 0 & \dots & x'_{3,r-1} & x'_{3,r} & \dots & x'_{3,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\
0 & 0 & \dots & 0 & x'_{r,r} & \dots & x'_{r,n} \\
0 & 0 & \dots & \dots & \dots & \dots & 0 \\
\vdots & \vdots & & & & & \vdots \\
0 & 0 & \dots & \dots & \dots & \dots & 0
\end{pmatrix} \tag{2}
$$

mit $x'_{\varrho,\varrho} \neq 0$ $(\varrho = 1, \dots, r)$. Unterhalb der r-ten Zeile und links von den Diagonalelementen $x'_{\varrho,\varrho}$ stehen lauter Nullen. Die Ausgangsfrage wird nun durch den nächsten Satz beantwortet.

Satz 11.3. *Die Maximalzahl linear unabhängiger Zeilen einer Matrix wird durch die Anwendung der elementaren Umformungen nicht geändert.*

Die hiernach gemeinsame Maximalzahl linear unabhängiger Zeilen der Matrizen (1) *und* (2) *ist gleich der Anzahl r derjenigen Zeilen der Matrix* (2)*, die nicht aus lauter Nullen bestehen.*

Die Zeilen der Matrizen (1) *und* (2) *spannen denselben Unterraum auf (dabei sind eventuelle Spaltenvertauschungen zu berücksichtigen). Die ersten r Zeilen der Matrix* (2) *bilden eine Basis dieses Unterraumes. Ebenso bilden diejenigen der Vektoren $\vec{x}_1, \dots, \vec{x}_k$ eine Basis, die bei der Berücksichtigung der evtl. vorgenommenen Zeilenvertauschungen den ersten r Zeilen der Matrix* (2) *entsprechen.*

Beispiel 11.4. Hinsichtlich einer Basis von X seien die Vektoren

$$\vec{x}_1 \leftrightarrow (1, 1, 2, -1),\ \vec{x}_2 \leftrightarrow (1, 2, -1, 0),\ \vec{x}_3 \leftrightarrow (2, 3, 6, 3),\ \vec{x}_4 \leftrightarrow (2, 4, 3, 4)$$

gegeben. Gesucht sind die Maximalzahl linear unabhängiger Vektoren unter ihnen sowie der von ihnen erzeugte Unterraum.

Aufstellung der Matrix und elementare Umformungen werden in folgenden Rechenschema ersichtlich:

1	1	2	−1	$\vec{z}_1$
1	2	−1	0	$\vec{z}_2$
2	3	6	3	$\vec{z}_3$
2	4	3	4	$\vec{z}_4$
1	1	2	−1	$\vec{z}_1$
0	1	−3	1	$\vec{z}_2 - \vec{z}_1$
0	1	2	5	$\vec{z}_3 - 2\vec{z}_1$
0	2	−1	6	$\vec{z}_4 - 2\vec{z}_1$
1	1	2	−1	$\vec{z}_1$
0	1	−3	1	$\vec{z}_2 - \vec{z}_1$
0	0	5	4	$\vec{z}_3 - 2\vec{z}_1 - (\vec{z}_2 - \vec{z}_1)$
0	0	5	4	$\vec{z}_4 - 2\vec{z}_1 - 2(\vec{z}_2 - \vec{z}_1)$
1	1	2	−1	$\vec{z}_1$
0	1	−3	1	$\vec{z}_2 - \vec{z}_1$
0	0	5	4	$\vec{z}_3 - \vec{z}_2 - \vec{z}_1$
0	0	0	0	$\vec{z}_4 - \vec{z}_3 - \vec{z}_2 + \vec{z}_1$

Ergebnis: Es gilt $\dim U = 3$, und $\{\vec{z}_1, \vec{z}_2, \vec{z}_3\}$ ist eine Basis von U.

Die Koordinaten eines Vektors hängen wesentlich von der Wahl der Basis ab: Geht man zu einer anderen Basis über, so entsprechen demselben Vektor im allgemeinen auch andere Koordinaten. Eine Ausnahme macht lediglich der Nullvektor, der hinsichtlich jeder Basis lauter Nullen als Koordinaten besitzt.

11.2. **Basistransformation.** Es seien $B = \{\vec{a}_1, \dots, \vec{a}_n\}$ und $B^* = \{\vec{a}_1^*, \dots, \vec{a}_n^*\}$ zwei Basen des Vektorraumes X. Die Vektoren der Basis B^* können dann eindeutig als Linearkombinationen von B dargestellt werden:

$$\vec{a}_\mu^* = \sum_{\nu=1}^{n} a_{\mu,\nu}\,\vec{a}_\nu \qquad (\mu = 1, \dots, n).$$

Die hierbei auftretenden Koeffizienten bilden eine n-reihige quadratische Matrix

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

Der Übergang von der Basis B zu der Basis B' wird eine *Basistransformation* genannt. Sie ist durch die *Transformationsmatrix* A eindeutig bestimmt.

Satz 11.5. *Eine quadratische Matrix ist genau dann die zu einer Basistransformation gehörende Transformationsmatrix, wenn alle ihre Zeilen linear unabhängig sind.*

Satz 11.6. *Der Vektor $\vec{x}$ besitze hinsichtlich der Basis B die Koordinaten $(x_1, \dots, x_n)$ und hinsichtlich der Basis B^* die Koordinaten $(x_1^*, \dots, x_n^*)$. Dann gilt*

$$x_\nu = \sum_{\mu=1}^{n} a_{\mu,\nu}\, x_\mu^* \qquad (\nu = 1, \dots, n). \tag{3}$$

Durch diese Gleichungen wird die der Basistransformation entsprechende *Koordinatentransformation* beschrieben. Sie drücken die Koordinaten x_ν durch die Koordinaten x_ν^* mit Hilfe der Transformationsmatrix aus.

Beispiel 11.7. Die Transformationsmatrix, die der Basistransformation $B \to B^*$ zugeordnet ist, sei gegeben durch

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

Der Vektor $\vec{x}$ habe hinsichtlich der Basis B^* die Koordinaten $(-1, 2, 1, 1)$. Dann hat er nach (3) hinsichtlich der Basis B die Koordinaten

$$\begin{array}{cc} \cdot & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix} \\ \begin{pmatrix} -1\,, & 2\,, & 1\,, & 1 \end{pmatrix} & \begin{pmatrix} -1\,, & \frac{11}{6}\,, & 0\,, & \frac{1}{6} \end{pmatrix}. \end{array}$$

Dabei wird die Berechnung der Koordinaten $\big(-1, \frac{11}{6}, 0, \frac{1}{6}\big)$ von $\vec{x}$ bezüglich der Basis B praktisch durchgeführt durch Bildung des „inneren Produkts“ von $(-1, 2, 1, 1)$ mit der jeweiligen Spalte der Matrix A, etwa $\frac{11}{6} = -1 \cdot 0 + 2 \cdot 1 + 1 \cdot (-\frac{1}{2}) + 1 \cdot \frac{1}{6}$.

Eine Deutung des vorstehenden Beispiels wird durch die Polynome

$$b_1(t) = 1\,, \quad b_2(t) = t\,, \quad b_3(t) = t^2\,, \quad b_4(t) = t^3$$

sowie

$$b_1^*(t) = \tbinom{t}{0} := 1\,, \quad b_2^*(t) = \tbinom{t}{1} := t\,, \quad b_3^*(t) = \tbinom{t}{2} := \tfrac{t(t-1)}{2}\,, \quad b_4^*(t) = \tbinom{t}{3} := \tfrac{t(t-1(t-2)}{6}$$

geliefert. Sie bilden jeweils eine Basis des Vektorraums $\mathbb{R}_3[t]$ aller reellen Polynome vom Grad ≤ 3. Die gegebene Transformationsmatrix A beschreibt die Basistransformation. Hat ein Polynom $x(t)$ die Darstellung

$$x(t) = -b_1^*(t) + 2b_2^*(t) + 3b_3^*(t) + b_4^*(t)$$

bezüglich der Basis B^*, so läßt sich auch leicht direkt die Basisdarstellung von $x(t)$ bezüglich der Basis B verifizieren:

$$x(t) = -b_1(t) + \tfrac{11}{6}\, b_2(t) + 0 \cdot b_3(t) + \tfrac{1}{6}\, b_4(t).$$

11.3. **Struktursätze.** Es seien X und Y zwei Vektorräume, und $\varphi : X \to Y$ sei eine beliebige Abbildung von X in Y. Sie ordnet jedem Vektor $\vec{x} \in X$ eindeutig einen mit $\varphi(\vec{x}) \in Y$ bezeichneten Bildvektor zu.

Definition 11.8. *$\varphi : X \to Y$ heißt eine lineare Abbildung, wenn φ die Linearitätseigenschaften besitzt:*

$$\varphi(\vec{x} + \vec{x}') = \varphi(\vec{x}) + \varphi(\vec{x}') \quad \textit{und} \quad \varphi(c\vec{x}) = c\,\varphi(\vec{x}).$$

Die Linearitätseigenschaften besagen die Vertauschbarkeit der Abbildung mit den linearen Operationen. Spezielle lineare Abbildungen sind die *Nullabbildung* 0, die jedem Vektor aus X auf den Nullvektor von Y abbildet, und die durch $\varepsilon(\vec{x}) = \vec{x}$ erklärte *Identität* $\varepsilon : X \to Y = X$.

Satz 11.9. *Für jede lineare Abbildung $\varphi : X \to Y$ gilt*

$$\varphi(\vec{0}) = \vec{0} \quad \textit{und} \quad \varphi(-\vec{x}) = -\varphi(\vec{x}).$$

Für endlich-dimensionale Vektorräume X liefert der folgende Satz einen Überblick über alle möglichen linearen Abbildungen von X in Y.

Satz 11.10. *Es sei $\{\vec{a}_1, \dots, \vec{a}_n\}$ eine Basis von X, und $\vec{b}_1, \dots, \vec{b}_n$ seien beliebige (nicht notwendig verschiedene) Vektoren aus Y. Dann gibt es genau eine lineare Abbildung $\varphi : X \to Y$ mit $\varphi(\vec{a}_\nu) = \vec{b}_\nu$ für $\nu = 1, \dots, n$.*

Ist M eine Teilmenge von X, so wird mit $\varphi(M)$ die Menge aller Bildvektoren $\varphi(\vec{x})$ von Vektoren $\vec{x} \in M$ bezeichnet.

Satz 11.11. *Es sei $\varphi : X \to Y$ eine lineare Abbildung, und U sei ein Unterraum von X. Dann ist $\varphi(U)$ ein Unterraum von Y, und es gilt $\dim \varphi(U) \le \dim U$.*

Eine lineare Abbildung bildet also Unterräume auf Unterräume ab. Speziell ist auch $\varphi(X)$ ein Unterraum von Y, dessen Dimension der *Rang* von φ genannt wird.

Definition 11.12. *Für lineare Abbildungen $\varphi : X \to Y$ wird $\operatorname{Rg} \varphi = \dim \varphi(X)$ gesetzt.*

Satz 11.13. *Es seien X, Y endlich-dimensionale Vektorräume und $\varphi : X \to Y$ eine lineare Abbildung. Dann sind die folgenden Aussagen äquivalent:*

(a) *φ ist surjektiv.*

(b) *Ist B eine Basis von X, so gilt $[\varphi(B)] = Y$.*

(c) $\operatorname{Rg} \varphi = \dim Y$.

Ist N eine Teilmenge von Y, so wird die Teilmenge $\varphi^{-1}(N) := \{\vec{x} \in X : \varphi(\vec{x}) \in N\}$ von X das *Urbild* von N bei der Abbildung φ genannt.

Satz 11.14. *Es sei $\varphi : X \to Y$ eine lineare Abbildung, und V sei ein Unterraum von Y. Dann ist $\varphi^{-1}(V)$ ein Unterraum von X.*

Speziell ist hiernach das Urbild $\varphi^{-1}(\{\vec{0}\})$ des Nullraums ein Unterraum von X. Er heißt der *Kern* von φ und seine Dimension der *Defekt* von φ. Der Rang mißt also die Dimension des Bildraums, der Defekt den Dimensionsverlust unter φ.

Definition 11.15. *Für lineare Abbildungen* $\varphi : X \to Y$ *wird* $\operatorname{Kern}\varphi := \varphi^{-1}(\{\vec{0}\})$ *und* $\operatorname{Def}\varphi := \dim(\operatorname{Kern}\varphi)$ *gesetzt.*

Satz 11.16. *Der Raum* X *sei endlich-dimensional. Dann gilt für jede lineare Abbildung* $\varphi : X \to Y$

$$\operatorname{Rg}\varphi + \operatorname{Def}\varphi = \dim X.$$

Satz 11.17. *Es sei* X *endlich-dimensional,* $\varphi : X \to Y$ *linear und* $M \subseteq X$*. Dann sind folgende Aussagen äquivalent:*

(a) φ *ist injektiv.*

(b) $\operatorname{Kern}\varphi = \{\vec{0}\}$.

(c) $\operatorname{Def}\varphi = 0$.

(d) *Genau dann ist* M *linear unabhängig, wenn* $\varphi(M)$ *linear unabhängig ist.*

(e) $\operatorname{Rg}\varphi = \dim X$.

12. Matrizenrechnung

Die Vektorräume dieses Abschnitts sind endlich-dimensional, und es ist in ihnen jeweils eine Basis festgelegt. Dann werden die Vektoren durch Koordinaten-n-Tupel, lineare Abbildungen durch Matrizen umkehrbar eindeutig beschrieben. Im Fall $\dim X = n$, $\dim Y = m$ wird also in $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ gerechnet, und der linearen Abbildung $\varphi : X \to Y$ entspricht eine (n, m)-Matrix $A = (a_{\nu\mu})$. Sie besteht aus n Zeilen und m Spalten und ordnet jedem Vektor $\vec{x} \in \mathbb{R}^n$ den Vektor

$$\begin{aligned}\vec{y} = \vec{x}\,A = (x_1, \dots, x_n) &\begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,m} \end{pmatrix} \\ &= (a_{1,1}x_1 + \dots + a_{n,1}x_n\,,\ \dots\ ,\ a_{1,m}x_1 + \dots + a_{n,m}x_n)\end{aligned}$$

zu.

Definition 12.1. *Unter der Transposition eines Vektors* $\vec{x} = (x_1, \dots, x_n)$ *oder einer Matrix* $A = (a_{\nu,\mu})$ *mit* $1 \le \mu \le m$, $1 \le \nu \le n$ *versteht man den Übergang zu*

$$\vec{x}^{\,T} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \textit{bzw.} \qquad A^T = (a_{\mu,\nu}) = \begin{pmatrix} a_{1,1} & \dots & a_{n,1} \\ \vdots & & \vdots \\ a_{1,m} & \dots & a_{n,m} \end{pmatrix}.$$

Die Gleichung $\vec{y} = \vec{x}\,A$ geht bei Transposition in $\vec{y}^{\,T} = A^T\,\vec{x}^{\,T}$ über. Die Reihenfolge der Faktoren ist zu beachten, insbesondere im Fall $m = n$.

Definition 12.2. *Es sei* $L(X,Y) = \{\varphi : X \to Y \;\; \text{linear}\}$ *die Menge aller linearen Abbildungen von* X *in* Y*. Durch*

$$(\varphi + \psi)(\vec{x}) := \varphi(\vec{x}) + \psi(\vec{x}) \qquad \text{und} \qquad (c\varphi)(\vec{x}) := c\,\varphi(\vec{x}) \tag{4}$$

werden auf $L(X,Y)$ *eine Addition von Abbildungen* $\varphi, \psi \in L(X,Y)$ *und eine Multiplikation von Abbildungen* $\varphi \in L(X,Y)$ *mit Skalaren* $c \in \mathbb{R}$ *punktweise erklärt.*

Satz 12.3. *Mit* φ *und* ψ *sind auch* $\varphi + \psi$ *und* $c\varphi$ *lineare Abbildungen von* X *in* Y*. Die Menge* $L(X,Y)$ *aller linearen Abbildungen von* X *in* Y *ist hinsichtlich der in* (4) *definierten linearen Operationen selbst ein Vektorraum. Es gilt* $\dim L(X,Y) = \dim X \cdot \dim Y$.

Insbesondere gehört die Nullabbildung 0 mit $0(\vec{x}) = \vec{0}$ für alle $\vec{x} \in \mathbb{R}^n$ zu $L(X,Y)$. Zu ihr gehört die Nullmatrix, deren sämtliche Einträge Null sind. Zu $\varphi \in L(X,Y)$ ist $-\varphi \in L(X,Y)$ mit $\varphi + (-\varphi) = 0$ die additiv inverse Abbildung. Gehört zu φ die Matrix A, so gehört zu $-\varphi$ die Matrix $-A$, deren Einträge die negativen der Einträge von A sind. Die zu (4) gehörenden linearen Matrixoperationen sind gegeben durch

$$A + B = \begin{pmatrix} a_{1,1} + b_{1,1} & \dots & a_{1,m} + b_{1,m} \\ \vdots & & \vdots \\ a_{n,1} + b_{1,1} & \dots & a_{n,m} + b_{n,m} \end{pmatrix} \qquad \text{und} \qquad cA = \begin{pmatrix} c\,a_{1,1} & \dots & c\,a_{1,m} \\ \vdots & & \vdots \\ c\,a_{n,1} & \dots & c\,a_{n,m} \end{pmatrix}.$$

Die Begriffe Rang, Kern, Defekt übertragen sich von linearen Abbildungen auf die zugeordneten Matrizen.

Beispiel 12.4. Gegeben ist die Matrix

$$A = \begin{pmatrix} 0 & 1 & -1 & -1 \\ 2 & -3 & -1 & -5 \\ 3 & -7 & 1 & -5 \end{pmatrix},$$

deren Rang, Kern, Defekt bestimmt werden. Dazu bedienen wir uns wieder der elementaren Umformungen. Statt $\vec{x}\,A = \vec{0}$ ist es zweckmäßig, die transponierte Form $A^T\,\vec{x}^{\,T} = \vec{0}^{\,T}$ zu verwenden und gleich den Kern von A zu bestimmen. Dies geschieht im nachstehenden Rechenschema.

x_1	x_2	x_3		
0	2	3	0	$\vec{z}_1$
1	-3	-7	0	$\vec{z}_2$
-1	-1	1	0	$\vec{z}_3$
-1	-5	-5	0	$\vec{z}_4$
1	-3	-7	0	$\vec{z}_2$
0	2	3	0	$\vec{z}_1$
0	-4	-6	0	$\vec{z}_2 + \vec{z}_3$
0	-8	-12	0	$\vec{z}_2 + \vec{z}_4$
1	-3	-7	0	$\vec{z}_2$
0	2	3	0	$\vec{z}_1$
0	0	0	0	$2\vec{z}_1 + \vec{z}_2 + \vec{z}_3$
0	0	0	0	$4\vec{z}_1 + \vec{z}_2 + \vec{z}_4$
1	0	$-5/2$	0	$\vec{z}_2 + 3/2\,\vec{z}_1$
0	1	$3/2$	0	$1/2\,\vec{z}_1$
0	0	0	0	$2\vec{z}_1 + \vec{z}_2 + \vec{z}_3$
0	0	0	0	$4\vec{z}_1 + \vec{z}_2 + \vec{z}_4$

Die Anzahl linear unabhängiger Zeilen ist also 2. Folglich gilt $\operatorname{Rg} A^T = \operatorname{Rg} A = 2$. Wegen $\dim \mathbb{R}^3 = 3$ ist $\operatorname{Def} A = 1$. Demnach wird der Kern von A von einem linear unabhängigen Vektor erzeugt, besteht also aus allen Vielfachen dieses Vektors. Indem wir zum Beispiel $x_3 = 2$ setzen und nach x_2 und x_1 auflösen, erweist sich der Vektor $(5, -3, 2)$ als eine nicht-triviale Lösung. Es folgt $\operatorname{Kern} A = [\{(5, -3, 2)\}] = \{t\,(5, -3, 2) \in \mathbb{R}^3 : t \in \mathbb{R}\}$.

Bemerkung 12.5. *Für (n, m)-Matrizen A gilt stets $\operatorname{Rg} A = \operatorname{Rg} A^T$. Zur Lösung linearer Gleichungssysteme rechnet man zweckmäßig mit Spaltenvektoren. Im folgenden wird gleich das transponierte System $A\,\vec{x} = \vec{0}$ betrachtet. Dabei ist A eine (n, m)-Matrix und $\vec{x} \in \mathbb{R}^n$ eine Spaltenvektor.*

Satz 12.6. *Es sei A eine (m, n)-Matrix und $A\,\vec{x} = \vec{0}$ ein Gleichungssystem von m Gleichungen für n Unbekannte. Dann gilt*

(a) *Genau dann hat das System nur die triviale Lösung $\vec{x} = \vec{0}$, wenn $\operatorname{Rg} A = n$ gilt.*

(b) *Die allgemeine Lösung des Systems enthält $\operatorname{Def} A = n - \operatorname{Rg} A$ freie Parameter.*

(c) *Insbesondere hat das System für $m < n$ stets nicht-triviale Lösungen.*

12.1. **Inhomogene lineare Gleichungssysteme.**

Definition 12.7. *Es sei A eine (m, n)-Matrix und $\vec{b} \in \mathbb{R}^m$ eine Spaltenvektor. Das lineare Gleichungssystem $A\,\vec{x} = \vec{b}$ heißt homogen, wenn $\vec{b} = \vec{0}$ ist. Andernfalls heißt es inhomogen.*

Satz 12.6 gibt über die allgemeine Lösung homogener linearer Gleichungssysteme Auskunft. Ein inhomogenes lineares Gleichungssystem muß nicht lösbar sein, nämlich genau dann nicht, wenn der Vektor $\vec{b}$ nicht in $\varphi(\mathbb{R}^n)$ liegt, wobei φ die der Matrix A^T entsprechende lineare Abbildung $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ ist. Die Lösbarkeitsbedingung lautet also $\vec{b} \in \varphi(\mathbb{R}^n)$. Das bedeutet, daß $\vec{b}$ in dem von den n Spaltenvektoren

$$\vec{s}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, \vec{s}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

aufgespannten Unterraum des $\mathbb{R}^m$ liegt. Wir setzen

$$A_{\text{erw}} = (A, \vec{b}) := \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{pmatrix}$$

und nennen A_{erw} die *erweiterte Koeffizientenmatrix* des linearen Gleichungssystems $A\,\vec{x} = \vec{b}$. Die Lösbarkeitsbedingung lautet damit

$$\operatorname{Rg} A = \operatorname{Rg} A_{\text{erw}}\,. \tag{5}$$

Es sei nun $\vec{v}$ eine Lösung von $A\,\vec{x} = \vec{b}$ und $\vec{u}$ eine beliebige Lösung des homogenen Systems $A\,\vec{x} = \vec{0}$. Dann ist $\vec{u} + \vec{v}$ eine Lösung des Systems $A\,\vec{x} = \vec{b}$. Sind umgekehrt $\vec{v}, \vec{w}$ zwei Lösungen von $A\,\vec{x} = \vec{b}$, so ist $\vec{v} - \vec{w}$ eine Lösung des zugehörigen homogenen Systems $A\,\vec{x} = \vec{0}$. Hieraus kommt der folgende Satz.

Satz 12.8. *Es sei A eine (m, n)-Matrix, $\vec{b} \in \mathbb{R}^m$ ein Spaltenvektor und $A\,\vec{x} = \vec{0}$ ein System von m linearen Gleichungen für n Unbekannte.*

(a) *Das System ist genau dann lösbar, wenn* (5) *gilt.*

(b) *Im Fall der Lösbarkeit enthält die allgemeine Lösung $n - \operatorname{Rg} A$ freie Parameter. Sie läßt sich darstellen in der Form*

$$\vec{x} = \vec{u}_0 + \vec{u}$$

mit einer speziellen Lösung $\vec{u}_0$ des inhomogenen Systems und der allgemeinen Lösung $\vec{u}$ des zugehörigen homogenen Systems.

Beispiel 12.9. Gegeben seien zwei lineare Gleichungssysteme $A\,\vec{x} = \vec{b}$ mit $\vec{b} \in \{\vec{b}_1, \vec{b}_2\}$:

$$A = \begin{pmatrix} 1 & -4 & 2 & 0 \\ 2 & -3 & -1 & -5 \\ 3 & -7 & 1 & -5 \\ 0 & 1 & -1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 10 \\ 10 \\ 5 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 14 \\ 16 \\ 2 \end{pmatrix}.$$

Da beide Systeme die gleiche Koeffizientenmatrix A besitzen, ist die simultane Lösung der Systeme im folgenden Rechenschema möglich.

x_1	x_2	x_3	x_4			
1	−4	2	0	0	2	$\vec{z}_1$
2	−3	−1	−5	10	14	$\vec{z}_2$
3	−7	1	−5	10	16	$\vec{z}_3$
0	1	−1	−1	5	2	$\vec{z}_4$
1	−4	2	0	0	2	$\vec{z}_1$
0	5	−5	−5	10	10	$\vec{z}_2 - \vec{z}_1$
0	5	−5	−5	10	10	$\vec{z}_3 - 3\vec{z}_1$
0	1	−1	−1	5	2	$\vec{z}_4$
1	−4	2	0	0	2	$\vec{z}_1$
0	1	−1	−1	5	2	$\vec{z}_4$
0	0	0	0	−15	0	$\vec{z}_2 - 2\vec{z}_1 - 5\vec{z}_4$
0	0	0	0	−15	0	$\vec{z}_3 - 3\vec{z}_1 - 5\vec{z}_4$

Offenbar ist die Lösbarkeitsbedingung (5), $\operatorname{Rg} A = \operatorname{Rg} A_{\text{erw}}$, nicht erfüllt für das System $A\,\vec{x} = \vec{b}_1$. Dieses System besitzt keine Lösung (die beiden letzten Zeilen lauten $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = -15$, was unmöglich ist).

Dagegen ist das System $A\,\vec{x} = \vec{b}_2$ lösbar, denn hier gilt $\operatorname{Rg} A = \operatorname{Rg} A_{\text{erw}}$. Eine weitere einfache Umformung ergibt

1	0	−2	−4		10
0	1	−1	−1		2
0	0	0	0		0
0	0	0	0		0

Eine spezielle Lösung des reduzierten Systems ist etwa

$$(x_1, x_2, x_3, x_4) = (10, 2, 0, 0),$$

und für $x_3 = 1, x_4 = 0$ bzw. $x_3 = 0, x_4 = 1$ ergeben sich zwei linear unabhängige Lösungen des homogenen Systems zu

$$(2, 1, 1, 0) \quad \text{bzw.} \quad (4, 1, 0, 1).$$

Damit lautet die allgemeine Lösung von $A\,\vec{x} = \vec{b}_2$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ 1 \\ 0 \\ 1 \end{pmatrix} \qquad (s, t \in \mathbb{R}).$$

12.2. **Matrizenprodukt und Inversion.** Sind $\varphi : X \to Y$ und $\psi : Y \to Z$ zwei lineare Abbildungen, so wird durch

$$(\psi \circ \varphi)(\vec{x}) := \psi(\varphi(\vec{x}))$$

eine Produktabbildung $\psi \circ \varphi : X \to Z$ definiert (*Verkettung* oder *Hintereinanderschaltung*).

Satz 12.10. *Mit φ und ψ ist auch $\psi \circ \varphi$ eine lineare Abbildung. Sinngemäß gelten die Gleichungen*

$$\begin{aligned} (\chi \circ \psi) \circ \varphi &= \chi \circ (\psi \circ \varphi), & (c\psi) \circ \varphi &= \psi \circ (c\varphi) = c(\psi \circ \varphi) \\ \psi \circ (\varphi_1 + \varphi_2) &= (\psi \circ \varphi_1) + (\psi \circ \varphi_2), & (\psi_1 + \psi_2) \circ \varphi &= (\psi_1 \circ \varphi) + (\psi_2 \circ \varphi). \end{aligned}$$

Es seien jetzt $\{\vec{a}_1, \ldots, \vec{a}_n\}$, $\{\vec{b}_1, \ldots, \vec{b}_r\}$ und $\{\vec{c}_1, \ldots, \vec{c}_s\}$ Basen der endlich-dimensionalen Vektorräume X, Y bzw. Z. Hinsichtlich der Basen von X und Y entspreche φ die Matrix $A = (a_{\nu,\varrho})$, und hinsichtlich der Basen von Y und Z entspreche ψ die Matrix $B = (b_{\varrho,\sigma})$. Der Produktabbildung $\psi \circ \varphi$ entspricht dann schließlich hinsichtlich der Basen von X und Z eine Matrix $C = (c_{\nu,\sigma})$, deren Elemente sich durch die Elemente von A und B ausdrücken lassen.

Satz 12.11. *Es gilt* $\quad c_{\nu,\sigma} = \sum_{\varrho=1}^{r} a_{\nu,\varrho}\, b_{\varrho,\sigma} \qquad (1 \le \nu \le n\,,\ 1 \le \sigma \le s).$

Diese Matrix C wird das Produkt der Matrizen A und B genannt und mit $A\,B$ bezeichnet. Der Produktabbildung entspricht also die *Produktmatrix* $A\,B$. Man achte jedoch auf die Reihenfolge! Bei den zugeordneten Matrizen ist die Reihenfolge der Faktoren gerade entgegengesetzt wie bei den Abbildungen.

Die Produktabbildung $\psi \circ \varphi$ kann nur dann gebildet werden, wenn der Bildraum Y von φ gleichzeitig der Originalraum von ψ ist. Dem entspricht: Die Produktmatrix $A\,B$ kann nur dann gebildet werden, wenn die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt (beide sind nämlich gleich der Dimension von Y). Die Produktmatrix läßt sich leicht berechnen: Das Matrixelement $c_{\nu,\sigma}$ ist das innere Produkt der ν-ten Zeile von A mit der σ-ten Spalte von B.

Beispiel 12.12. Gegeben seien die Matrizen

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

Dann ist die Produktmatrix

$$A\,B = \begin{pmatrix} 3 & -7 & 0 & 7 \\ 0 & 4 & -1 & 12 \end{pmatrix}$$

Im Spezialfall einer linearen Abbildung $\varphi : X \to X$ des n-dimensionalen Vektorraums X in sich braucht man nur eine Basis $\{\vec{a}_1, \ldots, \vec{a}_n\}$ von X zur Beschreibung der Abbildung durch eine n-reihige quadratische Matrix $A = (a_{\mu\nu})$. Deren Elemente $a_{\mu\nu}$ sind durch

$$\varphi(\vec{a}_\mu) = \sum_{\nu=1}^{n} a_{\mu\nu}\, \vec{a}_\nu \qquad (1 \le \mu \le n)$$

bestimmt. Insbesondere existiert zu bijektiven Abbildungen $\varphi \in L(X, X)$ stets die Umkehrabbildung φ^{-1}. Es gilt dann $\varphi^{-1} \circ \varphi = \varphi \circ \varphi^{-1} = \varepsilon$ (Identität).

Satz 12.13. *Die Dimension von X sei endlich. Eine Abbildung $\varphi \in L(X, X)$ ist genau dann bijektiv, wenn sie injektiv ist, und auch genau dann, wenn sie surjektiv ist. Die Umkehrabbildung φ^{-1} einer bijektiven Abbildung $\varphi \in L(X, X)$ ist wieder linear und bijektiv. Es gilt*

$$(\varphi^{-1})^{-1} = \varphi \qquad \text{und} \qquad (\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}.$$

Ist A die einer bijektiven Abbildung $\varphi : X \to X$ hinsichtlich einer Basis zugeordnete Matrix, so wird die der inversen Abbildung φ^{-1} zugeordnete Matrix mit A^{-1} bezeichnet und die *inverse Matrix* von A genannt. Der Identität ε von X entspricht hinsichtlich einer beliebigen Basis die *Einheitsmatrix*

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

die in der Hauptdiagonalen lauter Einsen und sonst lauter Nullen aufweist. Der Abbildungsgleichung $\varphi^{-1} \circ \varphi = \varphi \circ \varphi^{-1} = \varepsilon$ entspricht dann die Matrizengleichung $A\,A^{-1} = A^{-1}\,A = E$. Zur Berechnung von A^{-1} aus A sind die $n = \dim X$ Gleichungssysteme $A\,\vec{x} = \vec{e}_\nu$ für $1 \le \nu \le n$ zu lösen. Dies kann wieder mit elementaren Umformungen simultan geschehen.

Definition 12.14. *Eine (n, n)-Matrix A heißt regulär, wenn $\operatorname{Rg} A = n$ gilt, wenn also alle Zeilen bzw. Spalten von A linear unabhängig sind. Andernfalls heißt sie singulär.*

Beispiel 12.15. Die im folgenden Rechenschema angegebene Matrix A ist regulär und besitzt die angegebene Inverse:

	1	2	3	1	0	0	
A	-1	0	1	0	1	0	E
	0	2	1	0	0	1	
	1	2	3	1	0	0	
	0	2	4	1	1	0	
	0	2	1	0	0	0	
	1	0	-1	0	-1	0	
	0	2	4	1	1	0	
	0	0	3	1	1	-1	
	1	0	0	$1/3$	$-2/3$	$-1/3$	
	0	2	0	$-1/3$	$-1/3$	$4/3$	
	0	0	3	1	1	-1	
	1	0	0	$1/3$	$-2/3$	$-1/3$	
E	0	1	0	$-1/6$	$-1/6$	$2/3$	A^{-1}
	0	0	1	$1/3$	$1/3$	$-1/3$	

Ergebnis: Die Matrizen

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1/3 & -2/3 & -1/3 \\ -1/6 & -1/6 & 2/3 \\ 1/3 & 1/3 & -1/3 \end{pmatrix}$$

sind regulär und zueinander invers.

Es seien jetzt $\{\vec{a}_1, \ldots, \vec{a}_n\}$ und $\{\vec{a}^*_1, \ldots, \vec{a}^*_n\}$ zwei Basen von X, $\{\vec{b}_1, \ldots, \vec{b}_r\}$ und $\{\vec{b}^*_1, \ldots, \vec{b}^*_r\}$ zwei Basen von Y. Ferner sei $\varphi : X \to Y$ eine lineare Abbildung. Hinsichtlich der Basen $\{\vec{a}_\nu\}$ und $\{\vec{b}_\varrho\}$ entspricht dann φ eine Matrix A. Im allgemeinen entspricht aber derselben Abbildung φ hinsichtlich der anderen Basen $\{\vec{a}^*_\nu\}$ und $\{\vec{b}^*_\varrho\}$ auch eine andere Matrix A^*. Der Zusammenhang zwischen den Matrizen A und A^* kann mit Hilfe der Transformationsmatrizen, die zu den entsprechenden Basistransformationen gehören, hergestellt werden: Es gelte nämlich

$$\vec{a}^*_\mu = \sum_{\nu=1}^{n} s_{\mu,\nu}\, \vec{a}_\nu \quad \text{und} \quad \vec{b}^*_\sigma = \sum_{\varrho=1}^{r} t_{\sigma,\varrho}\, \vec{b}_\varrho\,,$$

es seien also $S = (s_{\mu,\nu})$ und $T = (t_{\sigma,\varrho})$ die Transformationsmatrizen, die als solche regulär sind.

Satz 12.16. *Es gilt* $A^* = S\,A\,T^{-1}$.

Im Fall $\varphi : X \to X$ ist bereits hinsichtlich der einen Basis $\{\vec{a}_\nu\}$ eine Matrix A und hinsichtlich einer anderen Basis $\{\vec{a}^*_\nu\}$ eine Matrix A^* zugeordnet. Man hat es dann nur mit einer Basistransformation und folglich auch nur mit einer Transformationsmatrix S zu tun. In diesem Fall gilt

$$A^* = S\,A\,S^{-1}. \tag{6}$$

13. Determinanten

13.1. **Determinantenformen.** Determinanten von zwei- und dreireihigen quadratischen Matrizen messen, abgesehen vom Vorzeichen, den Inhalt des von den Zeilenvektoren erzeugten Parallelogramms in der Ebene bzw. Parallelotops im Raum. Es sei nun $\{\vec{e}_1, \ldots, \vec{e}_n\}$ die kanonische Basis von $X = \mathbb{R}^n$, und Δ sei eine Abbildung, die je n Vektoren $\vec{x}_1, \ldots, \vec{x}_n \in X$ eindeutig einen mit $\Delta(\vec{x}_1, \ldots, \vec{x}_n)$ bezeichneten Skalar zuordnet.

Definition 13.1. Δ *heißt eine Determinantenform, wenn folgende Bedingungen erfüllt sind:*

(a) $\Delta(\ldots, \vec{x}_\nu + \vec{x}\,'_\nu, \ldots) = \Delta(\ldots, \vec{x}_\nu, \ldots) + \Delta(\ldots, \vec{x}\,'_\nu, \ldots)$.

(b) $\Delta(\ldots, c\vec{x}_\nu, \ldots) = c\,\Delta(\ldots, \vec{x}_\nu, \ldots)$.

(c) $\Delta(\vec{x}_1, \ldots, \vec{x}_n) = 0$, *wenn die Vektoren* $\vec{x}_1, \ldots, \vec{x}_n$ *linear abhängig sind.*

(d) $\Delta(\vec{e}_1, \ldots, \vec{e}_n) \neq 0$.

Die Bedingungen (a) und (b) drücken aus, daß Δ bzgl. jeder Matrixzeile linear ist. Die Bedingungen (c) und (d) enthalten notwendige Bedingungen für einen vernünftigen dimensionsangepaßten Inhaltsbegriff. In den folgenden Sätzen bedeutet Δ stets eine Determinantenform.

Satz 13.2. *Es seien $\mu, \nu \in \{1, \ldots, n\}$ verschiedene Indizes, c ein Skalar und π eine Permutation der Zahlen $1, \ldots, n$. Dann gilt*

(a) $\Delta(\ldots, \vec{x}_\mu, \ldots, \vec{x}_\nu, \ldots) = \Delta(\ldots, \vec{x}_\mu + c\vec{x}_\nu, \ldots, \vec{x}_\nu, \ldots)$,

(b) $\Delta(\ldots, \vec{x}_\mu, \ldots, \vec{x}_\nu, \ldots) = -\Delta(\ldots, \vec{x}_\nu, \ldots, \vec{x}_\mu, \ldots)$,

(c) $\Delta(\vec{x}_{\pi 1}, \ldots, \vec{x}_{\pi n}) = (\operatorname{sgn} \pi)\, \Delta(\vec{x}_1, \ldots, \vec{x}_n)$.

Zur Erinnerung: Jede Permutation π der Zahlen $1, \ldots, n$ läßt sich durch eine Folge von Vertauschungen zweier Zahlen realisieren. Dabei ist die Anzahl der verwendeten Vertauschungen entweder immer gerade ($\operatorname{sgn} \pi = 1$) oder immer ungerade ($\operatorname{sgn} \pi = -1$).

Es sei $\{\vec{a}_1, \ldots, \vec{a}_n\}$ eine Basis des $\mathbb{R}^n$. Wir schreiben jeden Zeilenvektor $\vec{x}_\nu = (x_{\nu 1}, \ldots, x_{\nu n})$ in der Form

$$\vec{x}_\nu = x_{\nu 1}\vec{a}_1 + \cdots + x_{\nu n}\vec{a}_n$$

und wenden Satz 13.2 sukzessive an. Es folgt

Satz 13.3. *Es gilt*

$$\Delta(\vec{x}_1, \ldots, \vec{x}_n) = \left(\sum_\pi (\operatorname{sgn} \pi)\, x_{1,\pi 1} \cdots x_{n,\pi n} \right) \Delta(\vec{a}_1, \ldots, \vec{a}_n),$$

wobei über alle Permutationen π der Zahlen $1, \ldots, n$ zu summieren ist. Für linear unabhängige Vektoren $\vec{x}_1, \ldots, \vec{x}_n$ gilt stets $\Delta(\vec{x}_1, \ldots, \vec{x}_n) \neq 0$.

Setzt man speziell $\vec{a}_1 = \vec{e}_1, \ldots, \vec{a}_n = \vec{e}_n$, so sind durch Satz 13.3 alle Determinantenformen bestimmt. Insbesondere notieren wir die

Definition 13.4. *Die Determinantenform $\Delta : X^n \to \mathbb{R}$ mit $\Delta(\vec{e}_1, \ldots, \vec{e}_n) = 1$ heißt Determinante. Man schreibt*

$$\det(\vec{x}_1, \ldots, \vec{x}_n) = \begin{vmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \ldots & x_{nn} \end{vmatrix} = \sum_\pi (\operatorname{sgn} \pi)\, x_{1\pi(1)} \cdots x_{n\pi(n)} \,.$$

Nach Satz 13.3 ist die Determinante unabhängig von der Basis $\{\vec{a}_1, \ldots, \vec{a}_n\}$ des Vektorraums $X = \mathbb{R}^n$. Sie hängt nur ab von der linearen Abbildung $\varphi : X \to X$ mit $\varphi(\vec{a}_\nu) = \vec{x}_\nu$ für $\nu = 1, \ldots, n$.

Beispiel 13.5. *Es gibt* $n!$ *Permutationen der Zahlen* $1, \ldots, n$. *Speziell folgt für*

$$n = 2: \qquad \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}\,,$$

$$n = 3: \qquad \begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} = \begin{cases} x_{11}x_{22}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} \\ -x_{11}x_{23}x_{32} - x_{12}x_{21}x_{33} - x_{13}x_{22}x_{31} \end{cases}$$

(Regel von Sarrus).

Satz 13.6. *Für quadratische Matrizen* A, B *gilt*

$$\det(A\,B) = \det A \cdot \det B\,.$$

Proof. Die Matrizen A, B repräsentieren lineare Abbildungen φ, ψ des Raumes $\mathbb{R}^n$ in sich mit $\vec{a}_\nu = \vec{e}_\nu\, A$, $\vec{b}_\nu = \vec{e}_\nu\, B$ und $\vec{c}_\nu = \vec{e}_\nu\, A\, B$.

(a) Es seien $\vec{a}_1, \ldots, \vec{a}_n$ linear unabhängig. Dann folgt

$$\det(A\,B) = \det(\vec{c}_1, \ldots, \vec{c}_n) = \frac{\Delta(\vec{c}_1, \ldots, \vec{c}_n)}{\Delta(\vec{a}_1, \ldots, \vec{a}_n)} \cdot \frac{\Delta(\vec{a}_1, \ldots, \vec{a}_n)}{\Delta(\vec{e}_1, \ldots, \vec{e}_n)}$$

$$= \frac{\Delta(\varphi(\vec{a}_1), \ldots, \varphi(\vec{a}_n))}{\Delta(\vec{a}_1, \ldots, \vec{a}_n)} \cdot \frac{\Delta(\vec{a}_1, \ldots, \vec{a}_n)}{\Delta(\vec{e}_1, \ldots, \vec{e}_n)} = \det B \cdot \det A\,.$$

(b) Sind $\vec{a}_1, \ldots, \vec{a}_n$ linear abhängig, so sind auch $\vec{c}_1, \ldots, \vec{c}_n$ linear abhängig, und es gilt $\det A = \det A\,B = 0$.

In jedem Fall folgt die Behauptung. □

13.2. **Berechnung von Determinanten.**

Satz 13.7. *Es sei* A *eine* n*-reihige quadratische Matrix und* c *ein Skalar. Dann gilt:*

(a) $\det A^T = \det A$.

(b) *Vertauscht man in* A *zwei Zeilen oder Spalten, so ändert die Determinante ihr Vorzeichen.*

(c) *Addiert man zu einer Zeile (Spalte) eine Linearkombination der übrigen Zeilen (Spalten), so ändert sich die Determinante nicht.*

(d) *Multipliziert man die Elemente einer Zeile (Spalte) mit* c*, so wird die Determinante mit* c *multipliziert.*

(e) *Sind in* A *zwei Zeilen (Spalten) gleich, so gilt* $\det A = 0$.

(f) $\det(cA) = c^n \det A$.

(g) $\det(A\,B) = (\det A)(\det B)$.

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

(i) $\det E = 1$.

Die praktische Berechnung von Determinanten geschieht auf der Grundlage dieses Satzes mit Hilfe elementarer Umformungen. Dazu wird A in die Gestalt A' überführt:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad \rightsquigarrow \quad A' = \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ 0 & a'_{22} & \dots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a'_{nn} \end{pmatrix},$$

wobei A' unterhalb der Hauptdiagonalen nur Nullen enthält.

Satz 13.8. *Werden bei den elementaren Umformungen von A insgesamt k Zeilen- oder Spaltenvertauschungen vorgenommen, so gilt*

$$\det A = (-1)^k \, a'_{11} a'_{22} \cdots a'_{nn} \,.$$

Beispiel 13.9. Das folgende Rechenschema zeigt die praktische Berechnung einer Determinante für $n = 4$.

1	1	2	1		1	1	2	1
2	3	3	0		0	1	−1	−2
1	−1	3	3		0	0	−1	−2
−1	0	−1	5		0	0	0	4
1	1	2	1		1	1	2	1
0	1	−1	−2		0	1	−1	−2
0	−2	1	2		0	0	−1	−2
0	1	1	6		0	0	2	8

Ergebnis:

$$\det A = \begin{vmatrix} 1 & 1 & 2 & 1 \\ 2 & 3 & 3 & 0 \\ 1 & -1 & 3 & 3 \\ -1 & 0 & -1 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 4 \end{vmatrix} = -4 \,.$$

Definition 13.10. *Werden in einer n-reihigen quadratischen Matrix A die i-te Zeile und die k-te Spalte gestrichen, so bleibt eine $(n-1)$-reihige quadratische Matrix A_{ik} übrig. Die mit dem Vorzeichenfaktor $(-1)^{i+k}$ versehene Determinante dieser Matrix heißt die Adjunkte des Matrixelements a_{ik}:*

$$\operatorname{Adj} a_{ik} = (-1)^{i+k} \det A_{ik} \,.$$

Satz 13.11. *Es gilt der Entwicklungssatz von Laplace:*

$$\sum_{\nu=1}^{n} a_{i\nu} \operatorname{Adj} a_{k\nu} = \sum_{\nu=1}^{n} a_{\nu i} \operatorname{Adj} a_{\nu k} = \begin{cases} \det A & \text{\textit{für} } i = k \\ 0 & \text{\textit{für} } i \neq k\,. \end{cases}$$

Dieser Satz führt die Berechnung einer n-reihigen Determinante auf die Berechnung von $(n-1)$-reihigen Determinanten zurück. Man spricht von der *Entwicklung der Determinante nach der i-ten Zeile bzw. Spalte*. Es ist dann nämlich

$$\det A = \sum_{\nu=1}^{n} a_{i\nu} \operatorname{Adj} a_{i\nu} = \sum_{\nu=1}^{n} a_{\nu i} \operatorname{Adj} a_{\nu i}\,.$$

Der in den Adjunkten auftretende Vorzeichenfaktor folgt dem (Schachbrett-) Muster

$$\begin{array}{cccc} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots} \end{array}$$

Beispiel 13.12. Entwicklung nach der ersten Zeile liefert

$$\begin{vmatrix} a_2 & 0 & 0 & b_2 \\ 0 & a_1 & b_1 & 0 \\ 0 & b_1 & a_1 & 0 \\ b_2 & 0 & 0 & a_2 \end{vmatrix} = a_2 \begin{vmatrix} a_1 & b_1 & 0 \\ b_1 & a_1 & 0 \\ 0 & 0 & a_2 \end{vmatrix} - b_2 \begin{vmatrix} 0 & a_1 & b_1 \\ 0 & b_1 & a_1 \\ b_2 & 0 & 0 \end{vmatrix}$$

und weiter durch Entwicklung nach den letzten Zeilen

$$= a_2^2 \begin{vmatrix} a_1 & b_1 \\ b_1 & a_1 \end{vmatrix} - b_2^2 \begin{vmatrix} a_1 & b_1 \\ b_1 & a_1 \end{vmatrix} = \left(a_2^2 - b_2^2\right)\left(a_1^2 - b_1^2\right).$$

Im Zusammenhang mit linearen Gleichungssystemen $A\,\vec{x}^{\,T} = \vec{b}^{\,T}$ mit der quadratischen Koeffizientenmatrix $A = (a_{\mu\nu})$ und Spaltenvektoren $\vec{x}^{\,T}$, $\vec{b}^{\,T}$ verdient der Fall $\det A \neq 0$ besondere Aufmerksamkeit, da es dann eine eindeutig bestimme Lösung $\vec{x}^{\,T}$ gibt. Es bezeichne A_i diejenige Matrix, die aus A durch Ersdetzen der i-ten Spalte durch den Spaltenvektor $\vec{b}^{\,T}$ hervorgeht.

Satz 13.13. *Es gilt die Cramersche Regel*

$$x_i = \frac{\det A_i}{\det A} \qquad (i = 1, \dots, n).$$

Für invertierbare Matrizen ergibt sich daraus noch der folgende Satz.

Satz 13.14. *Es sei A eine invertierbare Matrix, und es sei B erklärt durch*

$$B = (b_{\mu\nu}) \qquad \textit{mit} \qquad b_{\mu\nu} = \operatorname{Adj} a_{\mu\nu} \qquad (1 \leq \mu, \nu \leq n).$$

Dann besteht

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

14. Eigenwerte und Eigenvektoren

14.1. **Ähnliche Matrizen.** Es sei X ein endlich-dimensionaler Vektorraum, $n = \dim X$ und $\varphi : X \to X$ eine lineare Abbildung, der hinsichtlich einer Basis von X die n-reihige quadratische Matrix A entspricht. Wir untersuchen die für Anwendungen wichtige Frage, ob durch eine geeignete Wahl der Basis von X die φ zugeordnete Matrix eine besonders einfache Gestalt besitzt.

Definition 14.1. *Zwei n-reihige quadratische Matrizen A, B heißen ähnlich, wenn sie hinsichtlich zweier geeigneten Basen zur selben linearen Abbildung $\varphi : X \to X$ gehören.*

Nach Gleichung (6) läßt sich Ähnlichkeit wie folgt charakterisieren.

Satz 14.2. *Genau dann sind A, B ähnlich, wenn es eine Matrix S gibt mit* $\det S \neq 0$ *und*

$$B = S\,A\,S^{-1}.$$

Die Menge aller n-reihigen quadratischen Matrizen zerfällt hinsichtlich der Ähnlichkeitsbeziehung in Klassen zueinander ähnlicher Matrizen. Die eingangs gestellte Frage kann also auch so formuliert werden: Gibt es in jeder Ähnlichkeitsklasse eine Matrix möglichst einfacher Gestalt?

Definition 14.3. *Ein Vektor $\vec{x} \in X$ heißt Eigenvektor der linearen Abbildung $\varphi : X \to X$, wenn*

$$\vec{x} \neq \vec{0} \quad \text{und} \quad \varphi(\vec{x}) = c\,\vec{x} \tag{7}$$

mit einem geeigneten Skalar c gilt. Dann heißt c der zum Eigenvektor $\vec{x}$ gehörige Eigenwert (oder auch ein Eigenwert) von φ.

Offenbar bedeutet die Bestimmung der Eigenvektoren von φ dasselbe wie die nicht-triviale Lösung von $(\varphi - c\varepsilon)(\vec{x}) = \vec{0}$, also des homogenen linearen Gleichungssystems

$$\vec{x}(A - cE) = \vec{0} \quad \text{oder gleichwertig} \quad (A^T - cE)\vec{x}^{\,T} = \vec{0}^{\,T},$$

wenn A die zu φ gehörende Matrix ist. Wegen $\det A = \det A^T$ ist die nicht-triviale Lösbarkeitsbedingung dafür

$$\det(A - cE) = 0\,. \tag{8}$$

Definition 14.4. *Das Polynom*

$$f(t) = \det(\varphi - t\varepsilon) = \det(A - cE)$$

heißt das charakteristische Polynom der linearen Abbildung $\varphi : X \to X$ oder auch der zugehörigen Matrix A.

Satz 14.5. *Ähnliche Matrizen besitzen dasselbe charakteristische Polynom. Genau dann ist ein Skalar c Eigenwert von φ bzw. A, wenn c Nulstelle des charakteristischen Polynoms ist, wenn also*

$$f(c) = 0$$

gilt. Die zu einem Eigenwert c von φ bzw. A gehörenden Eigenvektoren sind genau die nicht-trivialen Lösungen des homogenen Gleichungssystems, dessen Koeffizientenmatrix

$$A^T - cE$$

ist (man beachte die Transposition).

Beispiel 14.6. Wir bestimmen Eigenwerte und Eigenvektoren der Matrix

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

Zur Eigenwertbestimmung berechnen wir das charakteristische Polynom

$$\det(A - tE) = \begin{vmatrix} 2-t & -1 & 2 \\ 0 & 1-t & 2 \\ -3 & 3 & -1-t \end{vmatrix} = (2-t)\begin{vmatrix} 1-t & 2 \\ 3 & -1-t \end{vmatrix} - 3\begin{vmatrix} -1 & 2 \\ 1-t & 2 \end{vmatrix}$$

$$= (t-2)(t-1)(t+1).$$

Die Matrix A besitzt also genau die Eigenwerte $c_1 = 2$, $c_2 = 1$ und $c_3 = -1$.

Zur Berechnung der zugehörigen Eigenvektoren ist jeweils $(A^T - cE)\vec{x}^{\,T} = \vec{0}^{\,T}$ nicht-trivial zu lösen.

$$c_1 = 2: \quad \begin{array}{rrr|r} x & y & z & \\ \hline 0 & 0 & -3 & 0 \\ -1 & -1 & 3 & 0 \\ 2 & 2 & -3 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline \end{array} \qquad \vec{s}_1 = (1, -1, 0)$$

$$c_2 = 1: \quad \begin{array}{rrr|r} x & y & z & \\ \hline 1 & 0 & -3 & 0 \\ -1 & 0 & 3 & 0 \\ 2 & 2 & -2 & 0 \\ \hline 1 & 0 & -3 & 0 \\ 0 & 1 & 2 & 0 \\ \hline \end{array} \qquad \vec{s}_2 = (3, -2, 1)$$

$$c_3 = -1: \quad \begin{array}{rrr|r} x & y & z & \\ \hline 3 & 0 & -3 & 0 \\ -1 & 2 & 3 & 0 \\ 2 & 2 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ \hline \end{array} \qquad \vec{s}_3 = (1, -1, 1)$$

Eigenvektoren zu c_1, c_2, c_3 sind also $\vec{s}_1 = (1, -1, 0)$, $\vec{s}_2 = (3, -2, 1)$ bzw. $\vec{s}_3 = (1, -1, 1)$.

Zu einem Eigenwert c können mehrere linear unabhängige Eigenvektoren gehören. Ihre Anzahl ist gleich dem Defekt der Abbildung $\varphi - c\varepsilon$ bzw. der Matrix $A - cE$. Sie kann demnach nicht die Vielfachheit übersteigen, mit der der Eigenwert c als Nullstelle des charakteristischen Polynoms auftritt.

Satz 14.7. *Es seien $\vec{x}_1, \dots, \vec{x}_k$ Eigenvektoren mit den zugehörigen Eigenwerten $c_1, \dots, c_k$ von φ bzw. A. Sind diese Eigenwerte paarweise verschieden, so sind die Eigenvektoren $\vec{x}_1, \dots, \vec{x}_k$ linear unabhängig.*

14.2. **Diagonalmatrizen.**

Definition 14.8. *Eine n-reihige quadratische Matrix D heißt Diagonalmatrix, wenn sie außerhalb der Hauptdiagonalen nur Nullen enthält, wenn sie also die Gestalt*

$$D = \begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_n \end{pmatrix}$$

besitzt.

Satz 14.9. *Einer linearen Abbildung $\varphi : X \to X$ entspricht genau dann bzgl. einer geeigneten Basis eine Diagonalmatrix, wenn es eine Basis von X gibt, die aus lauter Eigenvektoren von φ besteht. Hinsichtlich dieser Basis aus Eigenvektoren enspricht φ diejenige Diagonalmatrix, in deren Hauptdiagonale die Eigenwerte von φ entsprechend ihrer Vielfachheit auftreten.*

Beispiel 14.10. Die Eigenvektoren $\vec{x}_1, \vec{x}_2, \vec{x}_3$ aus Beispiel 14.6 sind linear unabhängig. Die Basistransformation, die den Übergang von der gegebenen Basis $\vec{a}_1, \vec{a}_2, \vec{a}_3$ zur Basis $\vec{x}_1, \vec{x}_2, \vec{x}_3$ beschreibt, wird nach Satz 11.6 durch die Transformationsmatrix

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

vermittelt. Ihre Zeilen sind die Koordinatentripel der Eigenvektoren $\vec{x}_1, \vec{x}_2, \vec{x}_3$ hinsichtlich der Basis $\vec{a}_1, \vec{a}_2, \vec{a}_3$. Für die Inverse von S rechnet man nach

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

Führt man durch die Transformation $\vec{x} = \vec{y}\, S$ die neuen Koordinaten hinsichtlich der Basis $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ ein, so geht $\vec{x}_\nu\, A = c_\nu\, \vec{x}_\nu$ über in $\vec{y}_\nu\, S\, A = c_\nu\, \vec{y}_\nu\, S$ oder äquivalent dazu in

$$\vec{y}_\nu (S\, A\, S^{-1}) = c_\nu\, \vec{y}_\nu\,.$$

Rechnung ergibt in der Tat

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

und zu $\vec{y}_\nu\, D = c_\nu\, \vec{y}_\nu$ gehören für $\nu = 1, 2, 3$ nun die Eigenvektoren $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$, bezogen auf die Basis $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$.

Satz 14.11. *Eine quadratische Matrix A ist genau dann zu einer Diagonalmatrix D ähnlich, wenn es eine Basis $\{\vec{x}_1, \ldots, \vec{x}_n\}$ von X gibt, die aus lauter Eigenvektoren von A besteht. In der dann gültigen Gleichung*

$$D = S\,A\,S^{-1} \tag{9}$$

sind die Zeilen der Transformationsmatrix S die Koordinatenzeilen der Eigenvektoren $\vec{x}_1, \ldots, \vec{x}_n$.

Bemerkung 14.12. *Eine Anwendung von* (9) *besteht in der schnellen Matrixpotenzierung: Ist A zu einer Diagonalmatrix D ähnlich, so ist nämlich A^q zu D^q ähnlich, und es gilt*

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

Demnach gilt $A^q = S^{-1}\,D^q\,S$ mit der Diagonalmatrix D^q, deren Hauptdiagonalelemente gerade die q-ten Potenzen der Hauptdiagonalemente von D sind.

Bemerkung 14.13. *Nicht jede (reelle) Matrix A ist zu einer (reellen) Diagonalmatrix ähnlich, wie die Matrizen*

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

lehren. Für die charakteristischen Polynome gilt

$$f(t) = t^2 + 1 \neq 1 \quad \text{bzw.} \quad g(t) = (t-1)^2.$$

A hat keine reellen Eigenwerte und Eigenvektoren und B hat den doppelten Eigenwert $c_1 = c_2 = 1$, aber nur einen zugehörigen Eigenvektor $\vec{x_1} = (0, 1)$. Demnach ist B zu keiner reellen Diagonalmatrix ähnlich. In der Tat besitzt A die komplexen Eigenwerte $\pm i$ mit den komplexen Eigenvektoren $(1, -i)$ bzw. $(1, i)$, aber für B ändert sich auch im Komplexen nichts.

15. Euklidische Räume

15.1. **Das Skalarprodukt.** Ziel dieses Abschnitts ist die Längen- und Winkelmessung in Vektorräumen.

Definition 15.1. *Es sei X ein Vektorraum, und $\beta : X^2 \to \mathbb{R}$ habe die Eigenschaften:*

(a) $\beta(\vec{x}_1 + \vec{x}_2, \vec{y}) = \beta(\vec{x}_1, \vec{y}) + \beta(\vec{x}_2, \vec{y})$,

(b) $\beta(c\,\vec{x}, \vec{y}) = c\,\beta(\vec{x}, \vec{y})$,

(c) $\beta(\vec{x}, \vec{y}) = \beta(\vec{y}, \vec{x})$ *(Symmetrie),*

(d) *aus* $\vec{x} \neq \vec{0}$ *folgt* $\beta(\vec{x}, \vec{x}) > 0$ *(Positivität).*

Dann heißt β ein Skalarprodukt und X ein euklidischer Raum.

Die Eigenschaften (a) und (b) beinhalten die Linearität von β bezüglich des ersten Arguments. Wegen (c) ist β auch linear bezüglich des zweiten Arguments, insgesamt also *bilinear*.

Beispiel 15.2. *Linearität und Symmetrie der nachstehenden Funktionen sind klar. Die Positivität bei* (b) *und* (c) *ist eine Übungsaufgabe.*

(a) In $\mathbb{R}^3$ wird ein Skalarprodukt erklärt durch
$$\beta(\vec{x},\vec{y}) = x_1y_1 + x_2y_2 + x_3y_3$$
für $\vec{x} = (x_1, x_2, x_3), \vec{y} = (y_1, y_2, y_3)$.

(b) In $\mathbb{R}^2$ wird ein Skalarprodukt erklärt durch
$$\beta(\vec{x},\vec{y}) = 3x_1y_1 - 4x_1y_2 - 4x_2y_2 + 7x_2y_2$$
für $\vec{x} = (x_1, x_2), \vec{y} = (y_1, y_2)$.

(c) Auf dem Vektorraum aller auf dem Intervall $[a, b] \subseteq \mathbb{R}$ stetigen Funktionen wird ein Skalarprodukt erklärt durch
$$\beta(f, g) = \int_a^b f(t)\, g(t)\, dt\,.$$

Satz 15.3. *Es gilt die Schwarzsche Ungleichung*
$$\beta(\vec{x},\vec{y})^2 \le \beta(\vec{x},\vec{x})\,\beta(\vec{y},\vec{y})$$
mit Gleichheit genau dann, wenn $\vec{x}$ und $\vec{y}$ linear abhängig sind.

Proof. Für $\vec{y} = \vec{0}$ gilt $\beta(\vec{x},\vec{y}) = \beta(\vec{y},\vec{y}) = 0$ wegen Definition 15.1 (b) und (d). In diesem Fall trifft die Behauptung zu. Nun sei $\vec{y} \neq \vec{0}$, also auch $\beta(\vec{y},\vec{y}) > 0$. Dann gilt für alle $t \in \mathbb{R}$
$$\begin{aligned} 0 &\le \beta(\vec{y},\vec{y})\beta(\vec{x}-t\vec{y},\vec{x}-t\vec{y}) \\ &= t^2\beta(\vec{y},\vec{y})^2 - 2t\beta(\vec{y},\vec{y})\beta(\vec{x},\vec{y}) + \beta(\vec{x},\vec{x})\beta(\vec{y},\vec{y}) \\ &= \big(t\beta(\vec{y},\vec{y}) - \beta(\vec{x},\vec{y})\big)^2 + \beta(\vec{x},\vec{x})\beta(\vec{y},\vec{y}) - \beta(\vec{x},\vec{y})^2. \end{aligned}$$
Die behauptete Ungleichung folgt speziell bei Wahl von $t = \frac{\beta(\vec{x},\vec{y})}{\beta(\vec{y},\vec{y})}$. In ihr besteht genau dann Gleichheit, wenn es ein $t \in \mathbb{R}$ gibt mit $\vec{x}-t\vec{y} = \vec{0}$, wenn also $\vec{x}$ und $\vec{y}$ linear unabhängig sind. □

Wir vereinfachen die Schreibweise, indem wir
$$\vec{x}\,\vec{y} := \beta(\vec{x},\vec{y})$$
setzen, sofern im euklidischen Raum X das Skalarprodukt feststeht. Dann nimmt die Schwarzsche Ungleichung die übliche Form an:
$$(\vec{x}\,\vec{y})^2 \le (\vec{x}\,\vec{x})(\vec{y}\,\vec{y}).$$

Definition 15.4. *Es sei X ein euklidischer Raum. Dann heißt*

$$|\vec{x}| = \sqrt{\vec{x}\,\vec{x}}$$

der Betrag oder die Länge von $\vec{x} \in X$. Man nennt $\vec{x}$ normiert oder einen Einheitsvektor, wenn $|\vec{x}| = 1$ gilt. Für $\vec{x} \neq \vec{0}$ und $\vec{y} \neq \vec{0}$ heißt

$$\cos(\vec{x}, \vec{y}) = \frac{\vec{x}\,\vec{y}}{|\vec{x}||\vec{y}|}$$

der Cosinus des Winkels zwischen $\vec{x}$ und $\vec{y}$.

Satz 15.5. *Der Betrag besitzt die Eigenschaften*

(a) $|\vec{x}| \geq 0\,,\ |\vec{x}| = 0 \iff \vec{x} = \vec{0}$ *(Definitheit),*

(b) $|c\vec{x}| = |c||\vec{x}|$ *(Multiplikativität),*

(c) $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$ *(Dreiecksungleichung).*

Aus der Schwarzschen Ungleichung folgt wegen $|\vec{x}\,\vec{y}| \leq |\vec{x}||\vec{y}|$

$$-1 \leq \cos(\vec{x}, \vec{y}) \leq 1.$$

Die Definition 15.4 ergibt $\vec{x}\,\vec{y} = |\vec{x}||\vec{y}|\cos(\vec{x}, \vec{y})$ sowie den Kosinussatz

$$|\vec{x} - \vec{y}|^2 = |\vec{x}|^2 + |\vec{y}|^2 - |\vec{x}||\vec{y}|\cos(\vec{x}, \vec{y}). \tag{10}$$

Im Fall $\cos(\vec{x}, \vec{y}) = 0$ ist dies der Satz von Pythagoras.

Definition 15.6. *Es sei X ein euklidischer Raum. Die Vektoren $\vec{x}, \vec{y} \in X$ heißen orthogonal, wenn $\vec{x}\,\vec{y} = 0$ gilt. Eine nicht-leere Teilmenge $M \subseteq X$ heißt ein Orthogonalsystem, wenn sie nicht den Nullvektor enthält und wenn je zwei verschiedene Vektoren aus M orthogonal sind. Es heißt ein Orthonormalsystem, wenn zusätzlich seine Vektoren normiert sind, also die Länge 1 haben. M heißt eine Orthonormalbasis von X, wenn M ein Orthonormalsystem ist, das zugleich Basis von X ist.*

Die Eigenschaft einer Menge $\{\vec{x}_1, \vec{x}_2, \ldots\}$, ein Orthonormalsystem zu sein, drückt man mit Hilfe des *Kronecker-Symbols*

$$\delta_{\mu\nu} := \begin{cases} 1 & \text{für } \mu = \nu \\ 0 & \text{für } \mu \neq \nu \end{cases}$$

aus: $\vec{x}_\mu\,\vec{x}_\nu = \delta_{\mu\nu}$.

Satz 15.7. *Jedes Orthogonalsystem ist linear unabhängig.*

Satz 15.8. *Ist $\{\vec{e}_1, \ldots, \vec{e}_n\}$ Orthonormalbasis eines euklidischen Raumes, so nimmt das Skalarprodukt eine einfache Gestalt an, nämlich*

$$\vec{x}\,\vec{y} = x_1 y_1 + \cdots + x_n y_n$$

mit den Koordinaten $(x_1, \dots, x_n)$ *von* $\vec{x}$ *und* $(y_1, \dots, y_n)$ *von* $\vec{y}$ *bezüglich* $\{\vec{e}_1, \dots, \vec{e}_n\}$. *Ferner gilt*

$$x_\nu = \vec{x}\,\vec{e}_\nu \qquad (\nu = 1, \dots, n).$$

Die Frage nach der Existenz von Orthonormalsystemen und Orthonormalbasen beantwortet in endlich-dimensionalen Vektorräumen das *Orthonormalisierungsverfahren von E. Schmidt*:

Es seien $\{\vec{a}_1, \dots, \vec{a}_n\}$ linear unabhängig. Dann läßt sich daraus ein Orthonormalsystem $\{\vec{e}_1, \dots, \vec{e}_n\}$ wie folgt induktiv konstruieren.

$$\begin{aligned}
\vec{e}_1 &= \frac{1}{|\vec{a}_1|}\,\vec{a}_1\,,\\
\vec{e}\,'_{k+1} &= \vec{a}_{k+1} - \sum_{\nu=1}^{k} (\vec{a}_{k+1}\,\vec{e}_\nu)\,\vec{e}_\nu\,,\\
\vec{e}_{k+1} &= \frac{1}{|\vec{e}\,'_{k+1}|}\,\vec{e}\,'_{k+1}\,.
\end{aligned}$$

Dabei spannen die Systeme $\{\vec{a}_1, \dots, \vec{a}_n\}$ und $\{\vec{e}_1, \dots, \vec{e}_n\}$ denselben Unterraum auf.

Beispiel 15.9. Das Orthonormalisierungsverfahren für die Vektoren

$$\vec{a}_1 = (1,2,1), \quad \vec{a}_2 = (-1,3,1), \quad \vec{a}_3 = (0,5,-4)$$

liefert

$$\begin{aligned}
\vec{e}_1 &= \frac{1}{\sqrt{6}}\,(1,2,1),\\
\vec{e}\,'_2 &= (-1,3,1) - \Big((-1,3,1)\frac{1}{\sqrt{6}}\,(1,2,1)\Big)\frac{1}{\sqrt{6}}\,(1,2,1) = (-2,1,0),\\
\vec{e}_2 &= \frac{1}{\sqrt{5}}\,(-2,1,0),\\
\vec{e}\,'_3 &= (0,5,-4) - \Big((0,5,-4)\frac{1}{\sqrt{6}}\,(1,2,1)\Big)\frac{1}{\sqrt{6}}\,(1,2,1)\\
&\qquad\qquad - \Big((0,5,-4)\frac{1}{\sqrt{5}}\,(-2,1,0)\Big)\frac{1}{\sqrt{5}}\,(-2,1,0) = (1,2,-5),\\
\vec{e}_3 &= \frac{1}{\sqrt{30}}\,(1,2,-5).
\end{aligned}$$

Daraus kommt die Orthonormalbasis

$$\vec{e}_1 = \frac{1}{\sqrt{6}}\,(1,2,1), \quad \vec{e}_2 = \frac{1}{\sqrt{5}}\,(-2,1,0), \quad \vec{e}_3 = \frac{1}{\sqrt{30}}\,(1,2,-5)$$

des $\mathbb{R}^3$ mit $[\{\vec{e}_1\}] = [\{\vec{a}_1\}]$ und $[\{\vec{e}_1, \vec{e}_2\}] = [\{\vec{a}_1, \vec{a}_2\}]$.

Im Vektorraum $\mathbb{R}^3$ (und nur hier) läßt sich ein zu $\vec{e}_1, \vec{e}_2$ orthonormaler Vektor auch aus dem Vektorprodukt $\vec{e}_1 \times \vec{e}_2$ gewinnen. Rechnung ergibt hier den Vektor $-\vec{e}_3$.

15.2. **Selbstadjungierte Abbildungen.**

Definition 15.10. *Es sei X ein euklidischer Raum. Eine lineare Abbildung $\varphi : X \to X$ heißt selbstadjungiert, wenn*

$$\varphi(\vec{x})\,\vec{y} = \vec{x}\,\varphi(\vec{y})$$

für alle $\vec{x}, \vec{y} \in X$ gilt.

Satz 15.11. *Es sei X ein endlich-dimensionaler euklidischer Raum, $\{\vec{e}_1, \ldots, \vec{e}_n\}$ eine Orthonormalbasis von X, und der linearen Abbildung $\varphi : X \to X$ entspreche die Matrix A. Genau dann ist φ selbstadjungiert, wenn*

$$A = A^T$$

gilt.

Proof. Die Matrixelemente $a_{\mu\nu}$ von A lassen sich in der Gestalt

$$a_{\mu\nu} = \varphi(\vec{e}_\mu)\,\vec{e}_\nu\,, \quad a_{\nu\mu} = \varphi(\vec{e}_\nu)\,\vec{e}_\mu \qquad (1 \le \mu, \nu \le n)$$

darstellen. Daraus folgt die Behauptung. □

Definition 15.12. *Eine quadratische Matrix A heißt symmetrisch, wenn $A = A^T$ gilt.*

Satz 15.13. *Eine symmetrische Matrix besitzt lauter reelle Eigenwerte und ist zu einer Diagonalmatrix ähnlich.*

Proof. Es seien $c, \overline{c} \in \mathbb{C}$ konjugiert komplexe Nullstellen des reellen charakteristischen Polynoms (Eigenwerte) und $\vec{x}, \vec{y}$ zugehörige konjugiert komplexe Eigenvektoren von A, also $y_\nu = \overline{x}_\nu$ für $\nu = 1, \ldots, n$. Dann gilt

$$(\vec{x}\,A)\,\vec{y} = c\,\vec{x}\,\vec{y}\,, \quad (\vec{y}\,A)\,\vec{x} = \overline{c}\,\vec{y}\,\vec{x}\,.$$

Wegen der Symmetrie der reellen Matrix A stimmen die linken Seiten überein, und aus $\vec{x}\,\vec{y} = \vec{y}\,\vec{x} = x_1^2 + \cdots + x_n^2 > 0$ folgt $c = \overline{c}$. Also gilt $c \in \mathbb{R}$. Durch Induktion nach n läßt sich zeigen, daß die zugehörigen Eigenvektoren von A ein Orthogonalsystem bilden. □

Beispiel 15.14. Die Matrix

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

ist symmetrisch. Ihr charakteristisches Polynom ist $\det(A - tE) = -t(t-1)(t-3)$, also sind die Eigenwerte $c_1 = 0$, $c_2 = 1$, $c_3 = 3$. Die zugehörigen Eigenvektoren ergeben sich (etwa) zu

$$\vec{x}_1 = (1, -1, -1), \quad \vec{x}_2 = (1, 1, 0), \quad \vec{x}_3 = (1, -1, 2).$$

In Übereinstimmung mit Satz 15.13 sind sie paarweise orthogonal. Ihre Normierung liefert eine Orthonormalbasis des $\mathbb{R}^3$, nämlich

$$\vec{s}_1 = \frac{1}{\sqrt{3}}\,\vec{x}_1\,, \quad \vec{s}_2 = \frac{1}{\sqrt{2}}\,\vec{x}_2\,, \quad \vec{s}_3 = \frac{1}{\sqrt{6}}\,\vec{x}_3\,.$$

Es sei S die Transformationsmatrix mit den Zeilen $\vec{s}_1$, $\vec{s}_2$, $\vec{s}_3$. Dann hat A bezüglich der Orthonormalbasis $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$ die Gestalt

$$D = S\,A\,S^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

16. Orthogonale Abbildungen

16.1. **Maßstabstreue.**

Definition 16.1. *Es seien X und Y euklidische Räume und $\varphi : X \to Y$ linear. Gilt*

$$\varphi(\vec{x})\,\varphi(\vec{x}\,') = \vec{x}\,\vec{x}\,' \tag{11}$$

für alle $\vec{x}, \vec{x}\,' \in X$, so heißt φ eine orthogonale Abbildung.

Satz 16.2. *Folgende Aussagen sind äquivalent:*

(a) *φ ist orthogonal,*

(b) *Aus $|\vec{x}| = 1$ folgt $|\varphi(\vec{x})| = 1$,*

(c) *Für alle $\vec{x} \in X$ gilt $|\varphi(\vec{x})| = |\vec{x}|$,*

(d) *Mit $\{\vec{e}_1, \dots, \vec{e}_n\} \subseteq X$ ist auch $\{\varphi(\vec{e}_1), \dots, \varphi(\vec{e}_n)\} \subseteq Y$ ein Orthonormalsystem.*

Eine Konsequenz dieses Satzes besteht in der *Maßstabstreue*, d.h. unter orthogonalen Abbildungen bleiben Längen und Winkel erhalten.

Folgerung 16.3. *Orthogonale Abbildungen sind injektiv.*

Es ist also stets $\dim \varphi(X) = \dim X$. Daher genügt es, bijektive lineare Abbildungen von X auf $Y = X$ zu betrachten. Setzt man dann $\vec{y} = \varphi(\vec{x}\,')$, so geht (11) über in

$$\varphi(\vec{x})\,\vec{y} = \vec{x}\,\varphi^{-1}(\vec{x}). \tag{12}$$

Satz 16.4. *Es sei $\varphi : X \to X$ linear und bijektiv. Dann sind äquivalent:*

(a) *φ ist orthogonal,*

(b) *φ^{-1} ist orthogonal,*

(c) *Für alle $\vec{x}, \vec{y} \in X$ gilt (12).*

Durch (12) wird die folgende Erklärung nahegelegt.

Definition 16.5. *Eine quadratische Matrix A heißt orthogonal, wenn A injektiv ist und $A^{-1} = A^T$ gilt.*

Folgerung 16.6. *Es sei X ein endlich-dimensionaler euklidischer Vektorraum und φ : $X \to X$ linear und injektiv. Genau dann ist φ orthogonal, wenn φ bezüglich einer Orthonormalbasis von X eine orthogonale Matrix entspricht.*

Satz 16.7. *Eine quadratische Matrix $A = (a_{\mu\nu})$ ist genau dann orthogonal, wenn ihre Zeilen bzw. Spalten ein Orthonormalsystem bilden. Das heißt*

$$\sum_{\varrho} a_{\mu\varrho}\, a_{\nu\varrho} = \sum_{\varrho} a_{\varrho\mu}\, a_{\varrho\nu} = \delta_{\mu\nu} \qquad \textit{(Kronecker-Symbol).}$$

Beispiel 16.8. Eine orthogonale Abbildung des $\mathbb{R}^3$ in sich wird vermittelt durch

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

Es handelt sich dabei um eine Drehung des $\mathbb{R}^3$ um die x-Achse mit dem Drehwinkel $\frac{\pi}{4}$. Ersetzt man in A das Matrixelement 1 durch -1, so wird zusätzlich an der (y,z)-Ebene gespiegelt.

Satz 16.9. *Für jeden Eigenwert $c \in \mathbb{C}$ einer orthogonalen Abbildung φ gilt $|c| = 1$, im Fall endlicher Dimension* $\det \varphi = \pm 1$.

Der folgende Satz ermöglicht die geometrische Deutung injektiver linearer Abbildungen endlich-dimensionaler euklidischer Vektorräume in sich.

Satz 16.10. *Es sei X ein endlich-dimensionaler euklischer Vektorraum und $\varphi : X \to X$ linear und injektiv. Dann kann φ eindeutig dargestellt werden in der Gestalt*

$$\varphi = \psi_2 \circ \psi_1\,,$$

wobei ψ_2 orthogonal ist und ψ_1 selbstadjungiert mit lauter positiven Eigenwerten.

Für die φ zugeordnete quadratische Matrix A heißt das: Es gibt eine eindeutige Zerlegung von A der Gestalt

$$A = B_1 B_2$$

mit orthogonaler Matrix B_2 und symmetrischer Matrix B_1, die nur positive Eigenwerte besitzt. Jede injektive lineare Abbildung des endlich-dimensionalen Vektorraums X in sich ist also aus einer Drehung oder Drehspiegelung und einer Stauchung zusammengesetzt.

Die Ermittlung der Matrizen B_1, B_2 erfordert eine konstruktive Beweisskizze. Die Matrix

$$B = A\,A^T$$

ist symmetrisch. Ihre Eigenwerte sind positiv, denn aus $\vec{x}\,B = c\,\vec{x}$ mit $\vec{x} \neq \vec{0}$ folgt

$$c\,|\vec{x}|^2 = c\,\vec{x}\,\vec{x} = \vec{x}\,B\,\vec{x}^{\,T} = (\vec{x}\,A)(\vec{x}\,A)^T = |\vec{x}\,A|^2 > 0.$$

Es sei nun $\{\vec{e}_1, \dots, \vec{e}_n\}$ eine zu den Eigenwerten $c_1, \dots, c_n > 0$ von B gehörende Orthonormalbasis. Hinsichtlich dieser Basis hat dann B Diagonalgestalt D mit den Hauptdiagonalelementen $c_1, \dots, c_n$. Hinsichtlich der neuen Basis setzen wir

$$B_1 = \begin{pmatrix} \sqrt{c_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{c_n} \end{pmatrix}.$$

Dann ist B_1 symmetrisch mit lauter positiven Eigenwerten, und diese Eigenschaften bleiben beim Übergang zur alten Orthonormalbasis erhalten. Mit den auf die neue Basis bezogenen Matrizen setzen wir

$$B_2 = B_1^{-1}\,A\,.$$

Zu zeigen bleibt, daß B_2 orthogonal ist. In der Tat gilt, bezogen auf die neue Basis ,

$$B_2\,B_2^T = B_1^{-1}\,A\,A^T\,(B_1^{-1})^T = B_1^{-1}\,D\,(B_1^{-1})^T = D\big(B_1^T\,B_1\big)^{-1} = D\,D^{-1} = E\,.$$

Beispiel 16.11. Gesucht ist die Zerlegung nach Satz 16.10 von

$$A = \begin{pmatrix} 11/2 & 7/2 \\ -3/2 & 9/2 \end{pmatrix}.$$

Rechnung liefert

$$B = A\,A^T = \begin{pmatrix} 85/2 & 15/2 \\ 15/2 & 45/2 \end{pmatrix} = \frac{5}{2}\begin{pmatrix} 17 & 3 \\ 3 & 9 \end{pmatrix}$$

sowie

$$\det(B - tE) = \frac{25}{4}\begin{vmatrix} 17 - 2/5\,t & 3 \\ 3 & 9 - 2/5\,t \end{vmatrix} = \frac{25}{4}\,(u^2 - 26u + 144),$$

wobei $u = \frac{2}{5}\,t$ gesetzt wurde. Daraus ergeben sich die Eigenwerte von B zu $u_1 = 13, u_2 = 8$, also $c_1 = 45, c_2 = 20$, sowie zugehörige normierte Eigenvektoren zu $\vec{e}_1 = \frac{1}{\sqrt{10}}\,(3,1)$ und $\vec{e}_2 = \frac{1}{\sqrt{10}}\,(-1,3)$. Die orthogonale Transformationsmatrix lautet damit

$$S = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}.$$

Ergebnis: Die gesuchten Matrizen B_1, B_2 mit $A = B_1 B_2$ sind

$$B_1 = S^T\begin{pmatrix} 3\sqrt{5} & 0 \\ 0 & 2\sqrt{5} \end{pmatrix} S = \frac{1}{\sqrt{20}}\begin{pmatrix} 29 & 3 \\ 3 & 21 \end{pmatrix},$$

$$B_2 = B_1^{-1}A = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

16.2. **Drehungen.**

Beispiel 16.12. In $\mathbb{R}^2$ ist einer orthogonalen Abbildung φ mit $\det\varphi = 1$ hinsichtlich einer Orthonormalbasis stets eine Matrix der Gestalt

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

zugeordnet. Der Betrag des Drehwinkels α ist allein durch die Abbildung φ bestimmt und insbesondere unabhängig von der Basis. Dagegen ist das Vorzeichen von α erst eindeutig festgelegt, wenn in $\mathbb{R}^2$ zusätzlich ein Drehsinn (eine *Orientierung* der Ebene) gegeben ist.

Satz 16.13. *Jede orthogonale Abbildung* $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ *mit* $\det\varphi = 1$ *besitzt den Eigenwert* 1. *Für* $\varphi \neq \varepsilon$ *stellt sie eine Drehung dar. Dabei ist die Drehachse die Fixgerade* g *unter* φ, *also der eindimensionale Unterraum* g *der Lösungen* $\vec{x} \in \mathbb{R}^3$ *von* $\varphi(\vec{x}) = \vec{x}$.

Beispiel 16.14. Ist φ eine orthogonale Abbildung des $\mathbb{R}^3$ in sich mit $\det\varphi = 1$, so läßt sich φ hinsichtlich jeder Orthonormalbasis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, bei der $\vec{e}_1$ Eigenvektor zum Eigenwert 1 ist, durch eine Matrix der Gestalt

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

beschreiben. Hier bewirkt die Abbildung φ eine Drehung um die durch $\vec{e}_1$ bestimmte Drehachse um den Winkel $\pm\alpha$, dessen Vorzeichen von der Orientierung des Raumes abhängt.

Definition 16.15. *Die Summe der Hauptdiagonalelemente einer quadratischen Matrix $A = (a_{\mu\nu})$ heißt die Spur von A,*

$$\operatorname{Spur} A = a_{11} + \cdots + a_{nn}\,.$$

Abgesehen vom Vorzeichen ist $\operatorname{Spur} A$ der Koeffizient von t^{n-1} des charakteristischen Polynoms

$$f(t) = \det(A - tE) = (-1)^n\big(t^n - \operatorname{Spur} A\; t^{n-1} + - \cdots + (-1)^n \det A\big)$$

der Matrix A. Da ähnliche Matrizen dasselbe charakteristische Polynom besitzen, stimmen ihre Spuren überein. Hieraus folgt der

Satz 16.16. *Es sei φ eine Drehung des $\mathbb{R}^3$ mit dem Drehwinkel α, und hinsichtlich einer beliebigen Basis des $\mathbb{R}^3$ sei φ die Matrix A zugeordnet. Dann gilt*

$$\cos\alpha = \frac{\operatorname{Spur} A - 1}{2}\,.$$

Beispiel 16.17. Die Matrix

$$A = \frac{1}{2}\begin{pmatrix} 1 & -\gamma-1 & \gamma \\ \gamma+1 & \gamma & -1 \\ \gamma & 1 & \gamma+1 \end{pmatrix} \qquad \text{mit} \qquad \gamma = \frac{\sqrt{5}-1}{2}$$

erfüllt $A\,A^T = E$, ist also orthogonal. Wegen $\det A = 1$ bewirkt sie eine Drehung des $\mathbb{R}^3$. Als Eigenvektor zum Eigenwert $c = 1$ ergibt sich

$$\vec{a} = (\gamma, 0, 1).$$

Der Drehwinkel α um die durch $\vec{a}$ charakterisierte Drehachse ist nach Satz 16.16 gegeben durch $\cos\alpha = \frac{\gamma}{2}$. Daraus folgt

$$\alpha = 72\,^\circ.$$

Für die praktische Rechnung ist die Beziehung $\gamma^2 + \gamma = 1$ nützlich.

17. Hauptachsentransformation

17.1. Quadratische Formen.

Definition 17.1. *Jede Funktion $P : \mathbb{R}^n \to \mathbb{R}$ der Gestalt*

$$P(x_1, \ldots, x_n) = a_0 + \sum_{\nu=1}^{n} a_\nu\, x_\nu + \sum_{\mu,\nu=1}^{n} a_{\mu\nu}\, x_\mu\, x_\nu$$

mit gegebenen Koeffizienten $a_\nu, a_{\mu\nu} \in \mathbb{R}$ heißt ein (reelles) quadratisches Polynom.

Es handelt sich also um eine Funktion von n reellen Variablen, die aus einem konstanten Anteil, einem linearen Anteil und einem quadratischen Anteil zusammengesetzt ist. Mit $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ läßt sie sich in der vektoriellen Gestalt

$$P(\vec{x}) = a_0 + \vec{a}\,\vec{x}^{\,T} + \vec{x}\,A\,\vec{x}^{\,T} \tag{13}$$

mit $a_0 \in \mathbb{R}$, $\vec{a} \in \mathbb{R}^n$ und einer reellen (n, n)-Matrix A schreiben. Wegen $x_\mu\, x_\nu = x_\nu\, x_\mu$ kann stets $a_{\mu\nu} = a_{\nu\mu}$ erreicht werden, also eine symmetrische Matrix $A = A^T$.

Definition 17.2. *Es sei $\vec{a} \in \mathbb{R}^n$ und A eine reelle symmetrische (n, n)-Matrix. Dann heißen*

$$\vec{a}\,\vec{x}^{\,T} \quad \textit{und} \quad \vec{x}\,A\,\vec{x}^{\,T} \qquad (\vec{x} \in \mathbb{R}^n)$$

eine Linearform bzw. eine quadratische Form.

Beispiel 17.3. (a) Die Rotationsenergie $T = \frac{1}{2}\,\vec{v}\,J\,\vec{v}^{\,T}$ eines starren Körpers ist eine quadratische Form in den Koordinaten des Drehgeschwindigkeitsvektors $\vec{v}$. Dabei ist die quadratische Matrix J der sog. *Trägheitstensor*, bezogen auf den Massenmittelpunkt.

(b) Die Gleichungen

$$36x^2 + 29y^2 - 24xy - 180 = 0$$

und

$$3x^2 + 3y^2 - 4z^2 + 10xy + 6\sqrt{2}\,x + 10\sqrt{2} + 14 = 0$$

beschreiben eine (gedrehte) Ellipse in der (x, y)-Ebene bzw. ein (verschobenes und gedrehtes) einschaliges Hyperboloid im (x, y, z)-Raum.

Satz 17.4. *Zu jeder quadratischen Form $P(\vec{x}) = \vec{x}\,A\,\vec{x}^{\,T}$ mit symmetrischer n-reihiger Matrix A existiert eine Orthonormalbasis $\{\vec{s}_1, \ldots, \vec{s}_n\}$ von Eigenvektoren $\vec{s}_1, \ldots, \vec{s}_n$ von A. Die Matrix A ist zu einer Diagonalmatrix D aus den zugehörigen Eigenwerten ähnlich, und mit der Transformationsmatrix S, deren Zeilen die Vektoren $\vec{s}_1, \ldots, \vec{s}_n$ sind, gilt*

$$D = S\,A\,S^T.$$

Definition 17.5. *Mit den Bezeichnungen aus Satz 17.4 heißt die Abbildung*

$$\vec{x} \mapsto \vec{y} = \vec{x}\,S \tag{14}$$

eine Hauptachsentransformation der quadratischen Form $\vec{x}\,A\,\vec{x}^{\,T}$ oder des quadratischen Polynoms $P(\vec{x})$ aus (13).

Die Hauptachsentransformation (14) ist bis auf Umorientierung der orthonormalen Eigenvektoren $\vec{s}_1, \ldots, \vec{s}_n$ von A oder deren Permutation eindeutig bestimmt. Das System der Eigenvektoren stellt dabei das Hauptachsensystem der *Eigenrichtungen* von A dar. Nach der Transformation (14) hat das Polynom (13) wegen $S^{-1} = S^T$ die einfache Gestalt

$$Q(\vec{y}) = P(\vec{y}\, S^T) = \vec{y}\, D\, \vec{y}^T + \vec{a}\, S\, \vec{y}^T + a_0 \,. \tag{15}$$

Hieraus ist die Art der Nullstellenmenge von $Q(\vec{y})$, genannt *Quadrik*, leicht erkennbar. Von Entartungsfällen abgesehen, beschreibt sie für $n = 2$ eine ebene Kurve, für $n = 3$ eine räumliche Fläche.

Unter der *Normalform* einer Quadrik versteht man die Darstellung (15), wenn sich der nicht-quadratische Anteil $\vec{a}\, S\, \vec{y}^T + a_0$ durch keine affine Substitution (Verschiebung oder Stauchung) mehr verkürzen läßt.

Anhand der Normalform lassen sich Quadriken klassifieren. Zum Beispiel ergeben sich in der Ebene Ellipsen, Hyperbeln und Parabeln. Dabei kommt es entscheidend auf die Diagonalmatrix D, also die Eigenwerte von A, an.

17.2. **Definite Matrizen.**

Definition 17.6. *Die symmetrische (n, n)-Matrix A heißt positiv (negativ) definit, wenn für alle $\vec{x} \neq \vec{0}$ des $\mathbb{R}^n$ für die quadratische Form $P(\vec{x}) = \vec{x}\, A\, \vec{x}^T$ stets $P(\vec{x}) > 0$ (stets $P(\vec{x}) < 0$) gilt. Sie heißt indefinit, wenn $P(\vec{x})$ sowohl positive als auch negative Werte annimmt. Sie heißt positiv (negativ) semidefinit, wenn für alle $\vec{x} \in \mathbb{R}^n$ stets $P(\vec{x}) \geq 0$ (stets $P(\vec{x}) \leq 0$) gilt.*

Satz 17.7. *Es sei A eine symmetrische (n, n)-Matrix.*

(a) *Genau dann ist A positiv definit, wenn A nur positive Eigenwerte besitzt.*
(b) *Genau dann ist A negativ definit, wenn A nur negative Eigenwerte besitzt.*
(c) *Genau dann ist A indefinit, wenn A positive und negative Eigenwerte besitzt.*
(d) *Ist A semidefinit, so besitzt A nur Eigenwerte, die entweder alle ≥ 0 sind, oder die alle ≤ 0 sind.*

Während die Eigenwertbestimmung oft nur mühsam (numerisch) durchführbar ist, gestattet das folgende nützliche Kriterium oftmals eine Entscheidung.

Satz 17.8. *Es sei A eine symmetrische (n, n)-Matrix und*

$$A_k := \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix} \qquad (k = 1, \ldots, n)$$

die k-reihige quadratische Untermatrix von A, die nach Streichen der ν-ten Zeilen und Spalten für $\nu = k+1, \ldots, n$ übrig bleibt. Genau dann ist A positiv definit, wenn $\det A_k > 0$ *für alle $k = 1, \ldots, n$ gilt. Genau dann ist A negativ definit, wenn $(-1)^k$* $\det A_k > 0$ *für alle $k = 1, \ldots, n$ gilt.*

Etwa sind die Matrizen aus Beispiel 17.3 (b),

$$\begin{pmatrix} 36 & -12 \\ -12 & 29 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 3 & 5 & 0 \\ 5 & 3 & 0 \\ 0 & 0 & -4 \end{pmatrix},$$

positiv definit bzw. indefinit.

17.3. **Beispiele.**

Beispiel 17.9. Die ebene Kurve

$$36x^2 + 29y^2 - 24xy - 180 = 0$$

ist auf Normalform zu transformieren.

Mit

$$A = \begin{pmatrix} 36 & -12 \\ -12 & 29 \end{pmatrix}$$

ist die vektorielle Form der Kurve $(x, y)\, A\, (x, y)^T = 180$. Aus dem charakteristischen Polynom $\det(A - tE) = t^2 - 65\,t + 900 = (t - 20)(t - 45)$ von A ergeben sich die Eigenwerte von A zu

$$c_1 = 20\,, \quad c_2 = 45\,,$$

und zugehörige Eigenvektoren sind

$$\vec{e}_1 = \frac{1}{5}\,(3, 4)\,, \quad \vec{e}_2 = \frac{1}{5}\,(-4, 3)\,.$$

Für die orthogonale Transformationsmatrix S und ihre Inverse $S^{-1} = S^{\,T}$ folgt

$$S = \frac{1}{5}\begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix}, \quad S^{-1} = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}.$$

Wegen $\det S = 1$ bewirkt S eine Drehung um den Winkel α mit $\cos\alpha = \frac{3}{5}$. Zur Transformation der Quadrik auf Hauptachsen mittels $(x', y') = (x, y)\, S^{\,T}$ schreiben wir

$$\big((x, y)\, S^{\,T}\big)\,\big(S\, A\, S^{\,T}\big)\,\big((x, y)\, S^{\,T}\big)^T = 180$$

mit der Diagonalmatrix

$$D = S\, A\, S^{\,T} = \begin{pmatrix} 20 & 0 \\ 0 & 45 \end{pmatrix}$$

und erhalten $20\, x'^2 + 45\, y'^2 = 180$, also die Ellipse

$$\frac{x'^2}{9} + \frac{y'^2}{4} = 1$$

mit den Halbachsen 3 und 2 . Sie ist gegen die x-Achse um $\arccos\frac{3}{5} \approx 53{,}1^\circ$ gedreht.

Beispiel 17.10. Die räumliche Fläche

$$3x^2 + 3y^2 - 4z^2 + 10xy + 6\sqrt{2}\,x + 10\sqrt{2}\,y + 14 = 0$$

ist auf Normalform zu transformieren.

Mit den Größen

$$A = \begin{pmatrix} 3 & 5 & 0 \\ 5 & 3 & 0 \\ 0 & 0 & -4 \end{pmatrix}, \quad \vec{a} = 2\sqrt{2}\,(3,5,0), \quad a_0 = 14$$

ist die vektorielle Form der Flächengleichung $(x,y,z)\,A\,(x,y,z)^T + \vec{a}\,(x,y,z)^T + a_0 = 0$. Wegen $\det A = 64 \neq 0$ ist A invertierbar, und das Gleichungssystem

(*) $$2\,\vec{m}\,A + \vec{a} = \vec{0}$$

ist eindeutig nach $\vec{m} = (-\sqrt{2}, 0, 0)$ auflösbar. Mit $\vec{x}\,' := \vec{x} - \vec{m}$ läßt sich die Quadrik in Mittelpunktslage verschieben. Es folgt

$$\vec{x}\,'\,A\,(\vec{x}\,')^T + (2\,\vec{m}\,A + \vec{a})(\vec{x}\,')^T + \vec{m}\,A\,\vec{m}^{\,T} + \vec{a}\,\vec{m}^{\,T} + a_0 = 0\,.$$

Kurze Rechnung liefert

$$\vec{x}\,'\,A\,(\vec{x}\,')^T + 8 = 0\,.$$

Im (x,y,z)-System liegt der Mittelpunkt der Quadrik also im Punkt $M = (\sqrt{2}, 0, 0)$. Das charakteristische Polynom $\det(A - tE) = -(t+4)(t+2)(t-8)$ von A liefert die Eigenwerte

$$c_1 = 8\,, \quad c_2 = -2\,, \quad c_3 = -4$$

von A, und zugehörige normierte Eigenvektoren ergeben sich zu

$$\vec{e}_1 = \frac{1}{\sqrt{2}}\,(1,1,0)\,, \quad \vec{e}_2 = \frac{1}{\sqrt{2}}\,(-1,1,0)\,, \quad \vec{e}_3 = (0,0,1)\,.$$

Daraus ergibt sich für die orthogonale Transformationsmatrix S und ihre Inverse $S^{-1} = S^{\,T}$

$$S = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Zur Transformation der Quadrik auf Hauptachsen mittels $\vec{x}\,'' := \vec{x}\,'\,S^{\,T}$ schreiben wir

$$(\vec{x}\,''\,S^{\,T})(S\,A\,S^{\,T})(\vec{x}\,''\,S^{\,T})^T + 8 = 0$$

und beachten, daß $D = S\,A\,S^{\,T}$ die Diagonalmatrix aus den Eigenwerten c_1, c_2, c_3 ist. Als Normalform ergibt sich nun $\vec{x}\,''\,D\,(\vec{x}\,'')^T + 8 = 0$ oder

$$-x''^2 + \frac{y''^2}{4} + \frac{z''^2}{2} = 1\,.$$

Es handelt sich um ein sog. einschaliges Hyperboloid mit dem Mittelpunkt in $M = (\sqrt{2}, 0, 0)$ im (x,y,z)-System und den Hauptachsen $(1,1,0)$, $(-1,1,0)$ und $(0,0,1)$. Die Drehachse ist die Parallele zur z-Achse durch den Punkt M, und der Drehwinkel ist gegeben durch $\cos\alpha = \frac{1}{2}\,(\operatorname{Spur} S - 1) = \frac{1}{\sqrt{2}}$, also $\alpha = 45^\circ$.