Dominik Eberlein · Christian Schwarz

Lineare Algebra

Dominik Eberlein
Eberlein@mathematik.tu-muenchen.de

Christian Schwarz
C.Schwarz@monet.m.isar.de

Dieses Skript stellt eine Ausarbeitung des Stoffes der Lineare Algebra-Vorlesung dar, die im Wintersemester 1994/95 und im Sommersemester 1995 von Prof. Meyberg an der Technischen Universität München gehalten wurde.

Copyright © Dominik Eberlein und Christian Schwarz 1995, 1996.
Version 1.7.0 vom 31. Oktober 1996, Ausdruck vom 31. Oktober 1996.

Inhaltsverzeichnis

0 Grundlagen 1
- 0.1 Mengen, Abbildungen und Relationen . . . 1
  - 0.1.1 Mengen nach Cantor . . . 1
  - 0.1.2 Mengenoperationen . . . 2
  - 0.1.3 Familien von Mengen . . . 2
  - 0.1.4 Abbildungen . . . 3
  - 0.1.5 Mächtigkeit von Mengen . . . 3
  - 0.1.6 Ordnungen und Relationen . . . 4
  - 0.1.7 Das Auswahlaxiom und das Lemma von Zorn . . . 6
- 0.2 Gruppen . . . 6
- 0.3 Ringe und Körper . . . 7

1 Matrizen und lineare Gleichungssysteme 8
- 1.1 Matrizen . . . 8
- 1.2 Matrizenmultiplikation . . . 10
- 1.3 Lineare Gleichungssysteme . . . 11
  - 1.3.1 Aussehen . . . 11
  - 1.3.2 Sonderformen linearer Gleichungssysteme . . . 11
  - 1.3.3 Der Gauß-Algorithmus . . . 12
  - 1.3.4 Homogene, lineare Gleichungssysteme . . . 13
- 1.4 Invertierbare Matrizen und Elementarmatrizen . . . 13
  - 1.4.1 Invertierbare Matrizen . . . 13
  - 1.4.2 Elementarmatrizen . . . 14
  - 1.4.3 Darstellung invertierbarer Matrizen durch Elementarmatrizen . . . 16

2 Vektorräume 17
- 2.1 Vektoren und Vektorräume . . . 17
- 2.2 Unterräume, Linearkombinationen, lineare Hülle . . . 18
  - 2.2.1 Unterräume . . . 18
  - 2.2.2 Linearkombinationen, lineare Hülle . . . 19
- 2.3 Basis und Dimension . . . 21
  - 2.3.1 Basis . . . 21
  - 2.3.2 Dimension . . . 23
- 2.4 Der Vektorraum der Polynome . . . 24
  - 2.4.1 Definitionen und Eigenschaften . . . 24
  - 2.4.2 Polynomdivision . . . 26
  - 2.4.3 Ideale . . . 27

3 Lineare Abbildungen 29
- 3.1 Definitionen und elementare Eigenschaften . . . 29
  - 3.1.1 Definitionen . . . 29
  - 3.1.2 Elementare Eigenschaften . . . 30
  - 3.1.3 Lineare Fortsetzung, Konstruktionssatz . . . 32
  - 3.1.4 Der Rang einer Abbildung und Rangformeln . . . 33
  - 3.1.5 Der Rang einer Matrix . . . 35
- 3.2 Darstellung von linearen Abbildungen durch Matrizen . . . 36
  - 3.2.1 Darstellungssatz . . . 36
  - 3.2.2 Basiswechsel . . . 37
  - 3.2.3 Invariante Unterräume . . . 39
- 3.3 Der Dualraum und duale Abbildungen . . . 40
- 3.4 Ergänzungen . . . 42
  - 3.4.1 Isomorphie . . . 42
  - 3.4.2 Faktorräume, Quotientenräume . . . 43
  - 3.4.3 Der Homomorphiesatz . . . 44

4 Determinanten 46
- 4.1 Definitionen und elementare Eigenschaften . . . 46
  - 4.1.1 Rekursive Definition der Determinante . . . 46
  - 4.1.2 Determinantenberechnung . . . 46
  - 4.1.3 Rechenregeln . . . 47
  - 4.1.4 Der Entwicklungssatz nach Laplace . . . 50
- 4.2 Die vollständige Entwicklung der Determinante . . . 51
  - 4.2.1 Permutationen . . . 51
  - 4.2.2 Der Entwicklungssatz von Leibniz . . . 51
- 4.3 Anwendungen von Determinanten . . . 52
- 4.4 Determinante eines Endomorphismus . . . 53

5 Eigenwerte 54
- 5.1 Eigenwerte und Eigenvektoren . . . 54
  - 5.1.1 Definitionen . . . 54
  - 5.1.2 Elementare Eigenschaften . . . 55
  - 5.1.3 Bemerkungen . . . 56
- 5.2 Die Bestimmung von Eigenwerten . . . 56
- 5.3 Trigonalisierung . . . 59
- 5.4 Diagonalisierung . . . 61
- 5.5 Der Satz von Cayley und Hamilton . . . 62

6 Euklidische und unitäre Vektorräume 65
- 6.1 Das Skalarprodukt . . . 65
- 6.2 Orthogonalisierung und Orthonormalisierung . . . 67
  - 6.2.1 Die Norm . . . 67
  - 6.2.2 Das Schmidtsche Orthogonalisierungsverfahren . . . 68
- 6.3 Adjungierte Abbildungen . . . 70
- 6.4 Orthogonale und unitäre Endomorphismen . . . 72
  - 6.4.1 Definitionen und Eigenschaften . . . 72
  - 6.4.2 Anwendungen in der Geometrie . . . 74
- 6.5 Selbstadjungierte Endomorphismen . . . 75

6.6 Schiefadjungierte Endomorphismen . . . 76
6.7 Normale Endomorphismen . . . 77
6.8 Orthogonale Trigonalisierung . . . 78
6.9 Orthogonale Diagonalisierung . . . 79
6.10 Komplexifizierung . . . 80
6.10.1 Fortsetzung des Vektorraumes . . . 80
6.10.2 Fortsetzung des Skalarproduktes . . . 80
6.10.3 Fortsetzung des Endomorphismus . . . 81
6.10.4 Spektralsatz für komplexifizierte Endomorphismen . . . 82
6.11 Symmetrische Matrizen und quadratische Formen . . . 85
6.11.1 Basiswechsel und Koordinatentransformation . . . 85
6.11.2 Quadratische Formen . . . 87
6.11.3 Hauptachsentransformation . . . 87
6.11.4 Quadriken . . . 88
6.11.5 Nichtorthogonale Diagonalisierung symmetrischer Matrizen . . . 90
6.11.6 Positive Definitheit von Matrizen . . . 92

7 Die Jordansche Normalform **95**
7.1 Formale Polynome . . . 95
7.2 Primärzerlegung . . . 99
7.3 Nilpotente Endomorphismen . . . 101
7.4 Jordan-Normalform . . . 103
7.5 Komplexifizierung . . . 104
7.6 Anwendungen . . . 105

8 Unendliche Vektorräume **109**
8.1 Definitionen . . . 109
8.2 Normierte Vektorräume . . . 111
8.3 Hilberträume . . . 112
8.4 Vollständige Orthogonalsysteme (Hilbertraumbasen) . . . 117

9 Multilineare Algebra **119**
9.1 Bilineare Abbildungen und Tensorprodukte . . . 119
9.1.1 Bilineare Abbildungen . . . 119
9.1.2 Bilineare Tensorprodukte . . . 126
9.2 Multilineare Abbildungen und Tensorprodukte . . . 128
9.2.1 Multilineare Abbildungen . . . 128
9.2.2 Multilineare Tensorprodukte . . . 129
9.3 Tensorielle Produkte linearer Abbildungen . . . 129
9.4 Tensormultiplikation . . . 129
9.5 Algebren . . . 130

Namen- und Sachverzeichnis **134**

Kapitel 0

Grundlagen

0.1 Mengen, Abbildungen und Relationen

0.1.1 Mengen nach Cantor

Definition. Eine Menge A heißt *Teilmenge* einer Menge B, wenn für jedes Element $x \in A$ auch $x \in B$ gilt. Wir schreiben dann $A \subseteq B$ und es ist $B \supseteq A \iff A \subseteq B$.
Gilt für jedes $x \in A$, daß x auch Element einer Menge B ist und gibt es ein $y \in B$ derart, daß $y \notin A$, so heißt A *echte Teilmenge* von B und wir schreiben $A \subsetneq B$.
Ist jedes Element $x \in A$ auch Element von B und gilt für alle $y \in B$ auch $y \in A$, so gilt $A = B$ und wir sagen es besteht *Gleichheit* zwischen A und B.

Definition. Die *Potenzmenge* $\mathcal{P}(A)$ ist die Menge aller Teilmengen von A. Es gilt

$$\mathcal{P}(A) := \{\, B \mid B \subseteq A \,\}.$$

Wie leicht nachzuprüfen ist, gilt:

$$\begin{array}{ccccc} & & A \in \mathcal{P}(A) & \text{und} & \\ a \in A & \iff & \{a\} \subseteq A & \iff & \{a\} \in \mathcal{P}(A). \end{array}$$

Definition. Die *leere Menge* wird mit dem Symbol $\emptyset$ bezeichnet. Es gilt:

$$\emptyset \subseteq A \quad \text{und} \quad \emptyset \in \mathcal{P}(A).$$

Definition. Wir bezeichnen die Menge $\mathbf{N} := \{1, 2, 3, \ldots\}$ als die Menge der *natürlichen Zahlen*. Die Menge $\mathbf{N}_0$ enthält zusätzlich auch die Null: $\mathbf{N}_0 := \mathbf{N} \cup \{0\}$. Weiter bezeichnet $\mathbf{Z}$ die Menge aller *ganzen Zahlen*, $\mathbf{Q}$ die Menge aller *rationalen Zahlen* (*Brüche*), $\mathbf{R}$ die Menge aller *rationalen und irrationalen Zahlen* und $\mathbf{C}$ die Menge aller *komplexen Zahlen*.

Die „naive Mengenlehre“ von Cantor kann nicht alles erfassen, wie die folgende Antinomie von Russell (ca. 1900) zeigt: Nach Cantor ist $\mathcal{M}$ die „Menge aller Mengen“ möglich. Danach gilt $\mathcal{M} \in \mathcal{M}$. Sei

$$P := \{\, M \in \mathcal{M} \mid M \in M \,\}$$

die Menge aller Mengen die sich selbst enthalten und

$$Q := \{ M \in \mathcal{M} \mid M \notin M \}$$

die Menge aller Mengen die sich nicht selbst enthalten. In welcher Menge aber liegt Q? Die Menge Q liegt nicht in P: $Q \notin P$. Ferner gilt aber auch $Q \notin Q$. Also ist Q weder in P noch in Q! (Außerdem folgt aus $Q \notin Q$, daß $Q \in Q$ gilt!)

0.1.2 Mengenoperationen

Definition. Seien A und B Mengen. Als Mengenoperationen definieren wir uns zunächst

- den Durchschnitt mit $A \cap B := \{ x \mid x \in A \wedge x \in B \}$,
- die Vereinigung mit $A \cup B := \{ x \mid x \in A \vee x \in B \}$ und
- die Differenzmenge mit $A \setminus B := \{ x \mid x \in A, x \notin B \}$.

Die Mengen A, B heißen *disjunkt* (elementfremd), wenn gilt $A \cap B = \emptyset$.

0.1.3 Familien von Mengen

Definition. Sei $I \neq \emptyset$ eine Menge. Jedem $i \in I$ sei eine Menge A_i zugeordnet. Dann heißt $(A_i)_{i \in I}$ eine *Familie von Mengen*, die *durch* I *indiziert wird* und es gilt

$$\bigcap_{i \in I} A_i := \{ x \mid x \in A_i \quad \text{für alle } i \in I \}.$$
$$\bigcup_{i \in I} A_i := \{ x \mid x \in A_i \quad \text{für ein } i \in I \}.$$

Definition. Das *kartesische Produkt* zweier Mengen A und B ist definiert durch

$$A \times B := \{ (a, b) \mid a \in A, b \in B \}.$$

Für n Mengen lautet die Definition

$$A_1 \times \cdots \times A_n := \{ (a_1, \ldots, a_n) \mid a_i \in A_i \quad (1 \leq i \leq n) \}.$$

Allgemein ist das kartesische Produkt der durch I indizierten Mengen A_i für $i \in I$ definiert durch

$$\prod_{i \in I} A_i := \{ (a_i)_{i \in I} \mid a_i \in A_i \quad \text{für alle } i \in I \}.$$

0.1.4 Abbildungen

Definition. Seien X, Y Mengen und $F \subseteq X \times Y$. Dann heißt das Tripel (X, Y, F) *Abbildung* von X in (oder nach) Y, wenn zu jedem $x \in X$ *genau ein* $y \in Y$ existiert, mit $(x, y) \in F$. Die Menge X heißt auch *Definitionsbereich*, die Menge Y *Bildbereich*.
Wir schreiben statt dem Tripel $(X, Y, F) =: f$ auch $f: X \to Y$ und statt $(x, y) \in F$ auch $f(x) = y$.

Definition. Es sei die Abbildung $f: X \to Y$ gegeben. Die Abbildung f heißt genau dann

- *injektiv*, wenn $f(x) = f(x') \Longrightarrow x = x'$ (mit $x, x' \in X$),
- *surjektiv*, wenn $f(X) = Y$, d.h. für alle $y \in Y$ existiert ein $x \in X$ mit $f(x) = y$ und
- *bijektiv*, wenn f injektiv und surjektiv ist.

Manchmal bezeichnen wir eine injektive Abbildung auch als *eineindeutige* oder *Eins-zu-Eins-Abbildung* („*one-to-one*“ Abbildung).

Definition. Unter der *Einschränkung* oder *Restriktion* $f|_A$ der Abbildung $f: X \to Y$ auf die (nichtleere) Teilmenge A von X verstehen wir diejenige Abbildung, die jedem $x \in A$ den Wert $f(x)$ zuordnet:

$$f|_A: A \to Y, \quad x \mapsto f(x)$$

Die Abbildung f heißt dann auch *Fortsetzung von* $f|_A$ *auf* X.

Definition. Es sei $f: X \to Y$, $g: Y \to Z$. Dann definieren wir

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

0.1.5 Mächtigkeit von Mengen

Definition. Seien X und Y Mengen. Dann sind sie *gleichmächtig*, falls es eine bijektive Abbildung $f: X \to Y$ gibt. In Zeichen:

$$|X| = |Y|.$$

Definition. Eine Menge X heißt *endlich*, wenn es ein n gibt, so daß gilt:

$$|X| = \big|\{1, \ldots, n\}\big| := n.$$

Definition. Eine Menge X heißt *abzählbar*, falls X endlich ist oder $|X| = |\mathbf{N}|$ gilt.

Beispiele:

1. Es gilt:
$$|\mathbf{N}| = |\mathbf{Z}| = |\mathbf{Q}| \neq |\mathbf{R}|.$$
Wir bezeichnen die Mächtigkeit der Menge $\mathbf{N}$ mit dem Symbol $\aleph_0$ (*Aleph*):
$$|\mathbf{N}| =: \aleph_0.$$
Da $|\mathbf{N}| = \infty$ ist $\aleph_0$ nicht endlich! Weiter gilt:
$$\aleph_0 < |\mathbf{R}| =: \aleph_1.$$
2. Es gilt $|X| \leq |Y|$ genau dann, wenn es eine injektive Abbildung $f: X \to Y$ gibt.

$\Diamond$

0.1.6 Ordnungen und Relationen

Definition. Sei $X \neq \emptyset$ eine Menge. Dann heißt

$$R \subseteq X \times X = \{ (x,y) \mid x,y \in X \}$$

Relation.

Definition. Die Relation $R \subseteq X \times X$ heißt *Ordnung* auf X (die Menge X wird durch R geordnet) wenn gilt:

$$\text{für alle } x \text{ ist } (x,x) \in R \quad (\textit{reflexiv}) \tag{Re1}$$
$$(x,y) \in R,\ (y,x) \in R \implies y = x \quad (\textit{antisymmetrisch}) \tag{Re2}$$
$$(x,y) \in R,\ (y,z) \in R \implies (x,z) \in R \quad (\textit{transitiv}) \tag{Re3}$$

Wir schreiben $(\leq)$ für die Ordnung R auf X: Es gilt $x \leq y$ genau dann, wenn $(x,y) \in R$.

Damit gilt $x \leq x$ für alle x, aus $x \leq y$ und $y \leq x$ folgt $y = x$ und aus $x \leq y$ und $y \leq z$ folgt $x \leq z$.

Beispiele für Ordnungen:

1. Die Relation $(\leq)$ ist eine Ordnung auf $\mathbf{R}$.
2. Sei $X = \mathcal{P}(A)$ die Potenzmenge von A. Dann ist die Inklusion eine Ordnung:
$$U \leq V \quad :\Longleftrightarrow \quad U \subseteq V \qquad \text{mit } U, V \in X.$$
3. Betrachten wir nun $X = \mathbf{N}$, so ist mit $a, b \in X$
$$a \leq b \quad :\Longleftrightarrow \quad a \backslash b \quad \text{„}a \text{ teilt } b\text{“}$$
eine Ordnung auf X.

$\Diamond$

Definition. Sei X mit $(\leq)$ geordnet.

- Die Menge $K \subseteq X$ heißt *Kette* (oder auch *vollständige Ordnung*), wenn für je zwei Elemente $u, v \in K$
$$u \leq v \quad \text{oder} \quad v \leq u$$
gilt.
- Ein Element $m \in X$ heißt *maximales Element*, falls für alle $x \in X$
$$m \leq x \implies m = x$$
gilt. Analog heißt $m \in X$ *minimales Element*, falls
$$x \leq m \implies m = x$$
für alle $x \in X$ gilt.

- Sei $L \subseteq X$. Dann heißt $s \in X$ *obere Schranke von* L, wenn
$$l \leq s \quad \text{für alle } l \in L$$
gilt. Analog dazu nennen wir $t \in X$ *untere Schranke von* L, wenn
$$t \leq l \quad \text{für alle } l \in L$$
gilt.

Definition. Sei X ein nichtleere Menge. Eine Relation $R \subseteq X \times X$ heißt *Äquivalenzrelation*, wenn

$$(x,x) \in R \quad \text{für alle } x \in X \quad (\textit{reflexiv}) \tag{A1}$$
$$(x,y) \in R \implies (y,x) \in R \quad (\textit{symmetrisch}) \tag{A2}$$
$$(x,y) \in R,\ (y,z) \in R \implies (x,z) \in R \quad (\textit{transitiv}) \tag{A3}$$

Wir schreiben auch $x \sim y$ für $(x,y) \in R \subseteq X \times X$.

Definition. Sei R eine Äquivalenzrelation auf X. Dann heißt

$$[x]_R := \{ y \in X \mid (x,y) \in R \} \text{ oder kurz } [x]$$

Äquivalenzklasse von $x \in X$.

Satz. *Sei $R \subseteq X \times X$ mit $X \neq \emptyset$ eine Äquivalenzrelation. Dann gilt*

(1) $[x] \neq \emptyset$,

(2) $X = \bigcup_{x \in X} [x]$ *und*

(3) $x \in [y] \iff x \sim y \iff [x] = [y] \iff [x] \cap [y] \neq \emptyset$.

Beweis: Wir bemerken, daß X nichtleer ist, denn R ist eine Äquivalenzrelation.

(1) Wegen $x \sim x$ ist $x \in [x] \neq \emptyset$.

(2) Wegen 1. gilt für beliebiges $x \in X$, daß $x \in \bigcup_{x \in X} [x]$ und umgekehrt, d.h. $X = \bigcup_{x \in X} [x]$.

(3) Ist $x \in [y]$, so ist per definitionem $x \sim y$. Sei $u \in [x]$, Dann ist aufgrund der Transitivität $u \sim y$ und damit $[x] \subseteq [y]$, sowie aus Symmetriegründen $[y] \subseteq [x]$, d.h. $[x] = [y]$. Damit ist ein beliebiges $u \in [x]$ auch Element von $[y]$, d.h. $[x] \cap [y] \neq \emptyset$. Daraus folgt wieder, daß es ein v mit $v \in [x]$ und $v \in [y]$ gibt, d.h. $(x,v) \in R, (v,y) \in R$, also $x \sim y$.

□

Definition. Wir definieren die *Quotientenmenge* als

$$X/R := \{ [x]_R \mid x \in X \}.$$

Definition. Jedes Element aus $[x]$ heißt *Repräsentant* oder *Vertreter* der Äquivalenzklasse $[x]$. Ein *Vertretersystem* enthält genau einen Vertreter jeder Klasse.

0.1.7 Das Auswahlaxiom und das Lemma von Zorn

Auswahlaxiom. *Sei I eine nichtleere Menge und $(A_i)_{i\in I}$ eine Familie von nichtleeren Mengen. Dann gilt:*

$$\prod_{i\in I} A_i \neq \emptyset.$$

Lemma von Zorn. *Sei M eine nichtleere, geordnete Menge, in der jede Kette eine obere Schranke besitzt. Dann gibt es in M wenigstens ein maximales Element.*

Bemerkung: Das Lemma von Zorn ist ein Axiom der formalen Mengenlehre und ist zum Auswahlaxiom äquivalent, d.h. erkennt man eines als richtig an, so läßt sich daraus das andere herleiten. Allerdings werden diese Axiome nicht von allen Mathematikern als richtig anerkannt.

0.2 Gruppen

Definition. Wir bezeichnen eine Abbildung $X \times X \to X$ bezüglich einer nichtleeren Menge X als *Verknüpfung* oder *innere Komposition*. Die Anwendung der Verknüpfung auf zwei Elemente $x, y \in X$ schreiben wir z.B. als $x \circ y$.

Definition. Eine Menge H mit einer inneren Komposition $(\circ)$ auf H heißt *Halbgruppe* und wir schreiben dafür $(H, \circ)$, wenn für $x, y, z \in H$ gilt:

$$H \neq \emptyset \quad \text{und} \tag{H1}$$
$$(x \circ y) \circ z = x \circ (y \circ z) \quad (\textit{Assoziativität}). \tag{H2}$$

Definition. Ist $(H, \circ)$ Halbgruppe, so bezeichnen wir als

- *neutrales Element* $e \in H$, wenn $x \circ e = e \circ x = x$ für alle $x \in H$ gilt,
- *Linksinverses* $a \in H$ von $b \in H$, wenn e existiert und $a \circ b = e$,
- *Rechtsinverses* $b \in H$ von $a \in H$, wenn e existiert und $a \circ b = e$,
- *Invertierbarkeit* von $a \in H$, wenn a ein Links- und Rechtsinverses besitzt.

Das neutrale Element bezüglich einer Addition nennen wir *Nullelement* (oder einfach *Null*), bezüglich einer Multiplikation *Einselement* (oder *Eins*).

Bemerkung: Ist $(H, \circ)$ eine Halbgruppe mit neutralem Element, so ist

- das neutrale Element eindeutig bestimmt, denn bei e, e' neutral gilt $e = e \circ e' = e'$ und

- bei $a, b, c \in H$ folgt aus $b \circ a = a \circ c = e$, daß $b = c$ gilt. Denn $b = b \circ (a \circ c) = (b \circ a) \circ c = c$. Damit ist das Linksinverse stets gleich dem Rechtsinversen. Aufgrund dieser Eindeutigkeit sprechen wir allgemein von „dem Inversen“ eines Elements $a \in H$, das wir bei einer Multiplikation mit a^{-1} und bei einer Addition mit $-a$ bezeichnen.

Definition. Die Halbgruppe $(H, \circ)$ heißt *kommutativ oder abelsch*, wenn

$$a \circ b = b \circ a$$

gilt.

Definition. Wir nennen $(G, \circ)$ *Gruppe*, wenn

$$(G, \circ) \text{ Halbgruppe mit neutralem Element } e \text{ ist und} \tag{G1}$$
$$\text{es zu jedem } a \in G \text{ ein } a^{-1} \in G \text{ gibt, so daß } a \circ a^{-1} = e \text{ gilt.} \tag{G2}$$

Ist die Verknüpfung kommutativ, so heißt $(G, \circ)$ *abelsche* oder *kommutative* Gruppe.

0.3 Ringe und Körper

Definition. Es heißt $(R, +, \circ)$ *Ring*, wenn

$$(R, +) \text{ eine additive abelsche Gruppe ist,} \tag{R1}$$
$$(R, \circ) \text{ eine multiplikative Halbgruppe ist und} \tag{R2}$$
$$x \circ (y + z) = x \circ y + x \circ z \quad \text{und} \quad (x + y) \circ z = x \circ z + y \circ z \quad (\textit{Distributivität}) \tag{R3}$$

für alle $x, y, z \in R$ gilt.

Definition. Wir bezeichnen $(K, +, \circ)$ als *Körper*, wenn

$$(K, +, \circ) \text{ ein kommutativer Ring mit Eins ist und} \tag{K1}$$
$$\text{jedes von Null verschiedene Element invertierbar ist.} \tag{K2}$$

Beispiele: Bekannte Beispiele für Körper sind die Mengen $\mathbf{Q}$, $\mathbf{R}$ und $\mathbf{C}$ bezüglich der „normalen“ Addition und Multiplikation.
Die Restklassengruppen $\mathbf{Z}_n$ bilden einen Ring, die Gruppen $\mathbf{Z}_p$, wobei p eine Primzahl ist, bilden einen Körper. Es gilt:

$$\mathbf{Z}_n := \{\bar{1}, \bar{2}, \ldots, \overline{n-1}\},$$
$$\bar{x} + \bar{y} := \bar{r}, \quad \text{falls } x + y = qn + r \text{ und } 0 \leq r \leq n-1,$$
$$\bar{x} \cdot \bar{y} := \bar{s}, \quad \text{falls } xy = q'n + s \text{ und } 0 \leq s \leq n-1.$$

$\Diamond$

Kapitel 1

Matrizen und lineare Gleichungssysteme

1.1 Matrizen

Im folgenden bezeichne K einen Körper.

Definition. Das rechteckige Schema

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \quad \text{mit } \alpha_{ij} \in K \quad (1 \le i \le m,\ 1 \le j \le n)$$

heißt $(m \times n)$-*Matrix über* K (Matrix *vom Typ* $m \times n$). Die Indizes i, j werden *Zeilen-* bzw. *Spaltenindex* genannt.

Wir schreiben auch kurz

$$A = (\alpha_{ij})_{m \times n}.$$

Definition. *Gleichheit zweier Matrizen.* Sind $A = (\alpha_{ij})_{m \times n}$ und $B = (\beta_{kl})_{r \times p}$, so gilt $A = B$ genau dann, wenn sowohl $m = r$ und $n = p$ als auch $\alpha_{ij} = \beta_{ij}$ für $1 \le i \le m$, $1 \le j \le n$ gilt.

Im folgenden seien nun $A = (\alpha_{ij})_{m \times n}$, $B = (\beta_{ij})_{m \times n} \in K^{m \times n}$ und $\lambda \in K$.

Definition.

- *Die Addition und Subtraktion zweier Matrizen* definieren wir durch

$$A + B := (\alpha_{ij} + \beta_{ij})_{m \times n} \text{ und}$$
$$A - B := (\alpha_{ij} - \beta_{ij})_{m \times n}.$$

- *Die Skalarmultiplikation* legen wir fest durch

$$\lambda A := (\lambda \alpha_{ij})_{m \times n}.$$

Bemerkung: Die $(m \times n)$-Matrizen über K bilden einen K-Vektorraum (siehe Kapitel 2).

Definition. Als die *transponierte Matrix* einer Matrix $A = (\alpha_{ij})_{m \times n}$ bezeichnen wir

$$A^T := (\alpha'_{ij})_{n \times m} \quad \text{mit } \alpha'_{ij} = \alpha_{ji}$$

und es gilt bei $\alpha \in K$, $A, B \in K^{m \times n}$ und $C \in K^{n \times p}$

$$(\alpha_1, \alpha_2, \dots, \alpha_n)^T = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}^T = (\beta_1, \beta_2, \dots, \beta_n),$$

$$(A + B)^T = A^T + B^T,$$
$$(\alpha A)^T = \alpha A^T,$$
$$(A^T)^T = A.$$

Definitionen besonderer Matrizen.

- *Die Einheitsmatrix* ist die Matrix, die auf der Diagonalen nur die Werte Eins hat und ansonsten nur Nullen enthält. Es gilt $E_n \in K^{n \times n}$:

$$E_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

  Ist die Anzahl der Zeilen bzw. Spalten der Einheitsmatrix aus dem Kontext einer Gleichung bekannt, so schreiben wir für E_n auch kurz E.

- Wir definieren E_{ij} als die Matrix, die im Schnittpunkt der i-ten Zeile und der j-ten Spalte eine Eins enthält und sonst nur aus Nullen besteht. Also

$$E_{ij} := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & 1 & \vdots \\ 0 & \dots & 0 \end{pmatrix}.$$

- *Die Nullmatrix* definieren wir durch

$$0 := \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix}.$$

- Ferner bezeichnen wir eine Matrix, die in ihrer Hauptdiagonalen ausgezeichnete Werte hat und sonst nur Nullen enthält als *Diagonalmatrix* und schreiben

$$\mathrm{Diag}(\alpha_1, \alpha_2, \dots, \alpha_n) := \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_n \end{pmatrix}.$$

Unter einer Matrix mit *oberer Dreiecksgestalt* verstehen wir eine Matrix der Form:

$$\begin{pmatrix} * & * & \dots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & * \end{pmatrix}.$$

(Dabei bezeichnet '$*$' ein beliebiges Element aus dem Grundkörper.)

Analog definieren wir den Begriff der *unteren Dreiecksgestalt*. Enthält eine Matrix mit Dreiecksgestalt auf der Diagonalen nur Nullen, so hat die Matrix *echte obere (untere) Dreiecksgestalt*.

- Zur Vereinfachung definieren wir $K_n := K^{1\times n}$ als die Menge der Zeilenvektoren der Länge n und analog $K^n := K^{n\times 1}$ als die Menge der Spaltenvektoren der Länge n.

1.2 Matrizenmultiplikation

Sei $A = (\alpha_{ij}) \in K^{m\times n}$ und $B = (\beta_{kl}) \in K^{n\times p}$. Dann gilt

$$\begin{aligned} C = (\gamma_{uv})_{m\times p} &= AB \\ &:= \begin{pmatrix} \alpha_{11}\beta_{11} + \dots + \alpha_{1n}\beta_{n1} & \dots & \alpha_{11}\beta_{1p} + \dots + \alpha_{1n}\beta_{np} \\ \vdots & \ddots & \vdots \\ \alpha_{m1}\beta_{11} + \dots + \alpha_{mn}\beta_{n1} & \dots & \alpha_{m1}\beta_{1p} + \dots + \alpha_{mn}\beta_{np} \end{pmatrix} \in K^{m\times p} \\ &= (Ab^1, Ab^2, \dots, Ab^p) \\ &= \begin{pmatrix} a_1 B \\ \vdots \\ a_m B \end{pmatrix}, \end{aligned}$$

wobei b^i $(1 \le i \le p)$ die i-te Spalte von B und a_j $(1 \le j \le m)$ die j-te Zeile von A ist. Es gilt

$$\gamma_{uv} = \sum_{k=1}^{n} \alpha_{uk}\beta_{kv}.$$

Seien K ein Körper und $A = (\alpha_{ij})_{m\times n} = (a^1, a^2, \dots, a^n)$ mit $\alpha_{ij} \in K$ und $x = (\xi_1, \dots, \xi_n)^t \in K^n$. Dann gilt

$$Ax = \begin{pmatrix} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \dots + \alpha_{1n}\xi_n \\ \vdots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \dots + \alpha_{mn}\xi_n \end{pmatrix} = \sum_{i=1}^{n} \xi_i a^i.$$

Diese Formel wollen wir die *Hauptformel der Matrizenmultiplikation* nennen.
Insbesondere gilt bei $A = (a^1, a^2, \dots, a^n) \in K^{m\times n}$

$$Ae^i = a^i \quad (1 \le i \le n),$$

wobei $e^i \in K^n$ der Spaltenvektor ist, der in der i-ten Zeile den Wert 1 und sonst nur Nullen enthält.

Für $A, A_1, A_2 \in K^{m\times n}$, $\alpha \in K$, $B, B_1, B_2 \in K^{n\times p}$ und $C \in K^{p\times q}$ gilt:

$$\begin{aligned} (A_1 + A_2)B &= A_1B + A_2B, \\ A(B_1 + B_2) &= AB_1 + AB_2, \\ \alpha(AB) &= (\alpha A)B = A(\alpha B), \\ (AB)C &= A(BC), \\ (AC)^T &= C^T A^T. \end{aligned}$$

Bemerkung: Wie wir leicht einsehen ist $(K^{n\times n}, +, \cdot)$ sogar ein Ring mit E_n als Eins.

1.3 Lineare Gleichungssysteme

1.3.1 Aussehen

Ein lineares Gleichungssystem mit m Gleichungen und n Unbekannten hat folgendes Aussehen:

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n &= \beta_2 \\ &\vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n &= \beta_m \end{aligned}$$

Das Gleichungssystem kann auch folgendermaßen verkürzt dargestellt werden:

$$Ax = b,$$

mit $A = (\alpha_{ij})_{m\times n}$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$. Die Matrix A wird *Koeffizientenmatrix*, die Matrix $(A \mid b)$ *erweiterte Matrix* genannt.

Definition. Ein lineares Gleichungssystem $Ax = b$ heißt *homogen*, falls $b = 0$. Ansonsten heißt es *inhomogen*.

Definition. Der Vektor $u \in K^n$ heißt *Lösung des Gleichungssystems* $Ax = b$, wenn $Au = b$ erfüllt ist.

1.3.2 Sonderformen linearer Gleichungssysteme

Gleichungen in Dreiecksform

Ein lineares Gleichungssystem, das von der Gestalt

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{22}x_2 + \dots + \alpha_{2n}x_n &= \beta_2 \\ \ddots \qquad\quad &\vdots \\ \alpha_{nn}x_n &= \beta_n \end{aligned}$$

mit $\alpha_{ii} \neq 0$ ist, hat *Dreiecksform*.
Ein Gleichungssystems in Dreiecksform lösen wir „von unten nach oben“:

$$x_n = \alpha_{nn}^{-1}\beta_n,$$

damit können wir aus der vorherigen Gleichung x_{n-1} berechnen. Dies wiederholen wir solange, bis wir bei x_1 angekommen sind.

Gleichungen in Staffel- bzw. Stufenform

Ein lineares Gleichungssystem hat *Staffel-* bzw. *Stufenform*, wenn es folgendes Aussehen hat:

$$\begin{aligned}
\alpha_{11}x_1 + \cdots + \cdots + \cdots + \cdots + \alpha_{1n}x_n &= \beta_1 \\
\alpha_{2j_2}x_{j_2} + \cdots + \cdots + \alpha_{2n}x_n &= \beta_2 \\
&\vdots \\
\alpha_{rj_r}x_{j_r} + \cdots + \alpha_{rn}x_n &= \beta_r
\end{aligned}$$

mit $\alpha_{ij_i} \neq 0$. Die Koeffizienten α_{ij_i} heißen *Pivots*.
Die Berechnung der Lösung eines solchen Gleichungssystems wird im nächsten Abschnitt behandelt.

1.3.3 Der Gauß-Algorithmus

Gegeben sei ein lineares Gleichungssystem (L). Wir rechnen mit der erweiterten Koeffizientenmatrix

$$\left(\begin{array}{cccc|c}
\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\
\alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m
\end{array}\right).$$

Wir lösen (L) durch

- Vertauschen zweier Gleichungen.
- Multiplizieren einer Gleichung mit einer Konstanten ungleich Null.
- Addieren eines Vielfachen einer Gleichung zu einer anderen.

Der Gaußalgorithmus:

1. Die erweiterte Koeffizientenmatrix wird durch elementare Zeilenumformungen in die *Zeilenstufenform* gebracht.
2. Das Gleichungssystem ist nur dann lösbar, falls in allen Zeilen ohne Pivot der Eintrag in der letzten Spalte Null ist.
3. Ist das Gleichungssystem nach Punkt 2 lösbar, so werden alle *freien Variablen* (ohne Pivot) der Reihe nach mit $\lambda_1, \dots, \lambda_{n-r}$ bezeichnet. (Hier bezeichnet r die Anzahl der Pivots.) Anschließend werden die Terme mit den freien Variablen auf die rechte Seite gebracht und das so erhaltene Gleichungssystem in Dreiecksform (vgl. 1.3.2) kann „von unten nach oben“ aufgelöst werden. Die Lösung ist dann von den Werten $\lambda_1, \dots, \lambda_{n-r}$ abhängig und wird deshalb auch *allgemeine Lösung* genannt. Bei speziellen λ-Werten spricht man von einer *speziellen Lösung*.

1.3.4 Homogene, lineare Gleichungssysteme

Bemerkung: Homogene Gleichungssysteme sind stets mit der *trivialen Lösung* $x = 0$ lösbar.

Satz. *Jedes homogene lineare Gleichungssystem* $Ax = 0$, $A \in K^{m\times n}$, *mit* $m < n$ *besitzt mindestens eine nichttriviale Lösung.*

Beweis: Formen wir die Matrix in die Staffelform um, so erhalten wir r Pivots mit $r \leq m < n$. Nun gibt es $n-r > 0$ freie Variablen. Setzen wir diese gleich 1, so ergibt dies eine Lösung $x \neq 0$.
□

1.4 Invertierbare Matrizen und Elementarmatrizen

1.4.1 Invertierbare Matrizen

Wir wollen nun den Begriff der *invertierbaren Matrizen* einführen.

Definition. Eine Matrix $A \in K^{n\times n}$ heißt *invertierbar*, wenn es $B, C \in K^{n\times n}$ gibt, mit $AB = CA = E$. In diesem Fall ist $B = C =: A^{-1}$ die *inverse Matrix zu* A (vgl. 0.2, „Invertierbarkeit“).

Seien $A, B \in K^{n\times n}$ invertierbar. Dann ist auch A^{-1} invertierbar, da $(A^{-1})^{-1} = A$ gilt, AB ist invertierbar, da $(AB)^{-1} = B^{-1}A^{-1}$ gilt und schließlich gilt $(A^T)^{-1} = (A^{-1})^T$. Dies folgt aus $A^{-1}A = AA^{-1} = E$ und $B^{-1}A^{-1}AB = ABB^{-1}A^{-1} = E$ wegen der Eindeutigkeit der Inversen.

Definition. Die Menge aller invertierbaren Matrizen des $K^{n\times n}$ bezeichnen wir mit

$$GL(n, K) := \left\{ A \in K^{n\times n} \;\middle|\; A \text{ invertierbar} \right\}.$$

Dabei ist $GL(n, K)$ mit dem Matrizenprodukt eine Gruppe; sie heißt *allgemeine lineare Gruppe* (engl. *general linear group*).

Wir können invertierbare Matrizen verwenden, um die Unbekannte $X \in K^{n\times p}$ in folgender Gleichung zu bestimmen:

$$AX = B \quad \text{mit } A \in GL(n, K) \text{ und } B \in K^{n\times p}.$$

Durch Anwenden der Inversen auf beiden Seiten erhalten wir $X = A^{-1}B$.

Satz. *Enthält* A *eine Nullzeile oder Nullspalte, dann ist* A *nicht invertierbar.*

Beweis: Sei $A = (a^1, \dots, a^n)$ eine invertierbare Matrix und $B = (b^1, \dots, b^n)$ die zu ihr inverse Matrix. Es gilt

$$AB = (Ab^1, \dots, Ab^n) = BA = E_n = (e^1, \dots, e^n).$$

Durch Spaltenvergleich erhalten wir

$$Ab^i = Ba^i = e^i \qquad (1 \leq i \leq n),$$

und insbesondere

$$a^i \neq 0.$$

□

1.4.2 Elementarmatrizen

Wir unterscheiden drei verschiedene Typen von elementaren Umformungen. Sei im folgenden $1 \leq i \leq j \leq n$ und

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \in K^{m\times n}.$$

Typ 1. Vertauschung zweier Zeilen.

$$\widetilde{A} = \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = P_{ij}A, \text{ mit } P_{ij} := \begin{pmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ e_i \\ \vdots \\ e_n \end{pmatrix} \in K^{m\times m}.$$

Typ 2. Multiplikation einer Zeile mit $\alpha \neq 0$.

$$\widetilde{A} = \begin{pmatrix} a_1 \\ \vdots \\ \alpha a_i \\ \vdots \\ a_n \end{pmatrix} = D_i(\alpha)A, \text{ mit } D_i(\alpha) := \begin{pmatrix} e_1 \\ \vdots \\ \alpha e_i \\ \vdots \\ e_n \end{pmatrix} \in K^{m\times m}.$$

Typ 3. Addition des α-fachen einer Zeile zu einer anderen.

$$\widetilde{A} = \begin{pmatrix} a_1 \\ \vdots \\ a_i + \alpha a_j \\ \vdots \\ a_n \end{pmatrix} = T_{ij}(\alpha)A, \text{ mit } T_{ij}(\alpha) := \begin{pmatrix} e_1 \\ \vdots \\ e_i + \alpha e_j \\ \vdots \\ e_n \end{pmatrix} \in K^{m\times m}.$$

Analog sehen die elementaren Spaltenumformungen aus.
Wir bezeichnen P_{ij}, $D_i(\alpha)$ und $T_{ij}(\alpha)$ als *Elementarmatrizen*. Manchmal schreiben wir auch S_k und meinen damit allgemein eine Elementarmatrix.

Satz. *Elementarmatrizen sind invertierbar.*

Beweis: Seien $P_{ij}, D_i(\alpha), T_{ij}(\alpha) \in K^{m\times m}$. Dann ist offensichtlich

$$P_{ij}P_{ij} = E,$$
$$D_i(\alpha)D_i(\alpha^{-1}) = D_i(\alpha^{-1})D_i(\alpha) = E \text{ und}$$
$$T_{ij}(\alpha)T_{ij}(-\alpha) = T_{ij}(-\alpha)T_{ij}(\alpha) = E.$$

□

Satz.

(1) *Entstehen* $\widetilde{A}$ *bzw.* $\widetilde{E}$ *aus* $A \in K^{m\times n}$ *bzw.* E_m *durch dieselbe elementare Zeilenumformung, so gilt* $\widetilde{A} = \widetilde{E}A$.

(2) *Entstehen* $\widetilde{A}$ *bzw.* $\widetilde{E}$ *aus* $A \in K^{m\times n}$ *bzw.* E_n *durch dieselbe elementare Spaltenumformung, so gilt* $\widetilde{A} = A\widetilde{E}$.

(3) *Entsteht* M *aus* $A \in K^{m\times n}$ *und* P *aus* E_m *durch dieselben endlich vielen elementaren Zeilenumformungen, dann ist* P *das Produkt von Elementarmatrizen und es gilt* $M = PA$.

(4) *Entsteht* N *aus* $A \in K^{m\times n}$ *und* Q *aus* E_n *durch dieselben endlich vielen elementaren Spaltenumformungen, dann ist* Q *das Produkt von Elementarmatrizen und es gilt* $N = AQ$.

Beweis:

(1) Die Behauptung ergibt sich sofort aus der Definition der elementaren Zeilenumformungen.

(2) Analog zu 1.

(3) Da eine elementare Zeilenumformung der Multiplikation mit einer entsprechenden Elementarmatrix entspricht, gilt bei $A \in K^{m\times n}$ und $E \in K^{m\times m}$:

$$A \mid E$$
$$S_k \ldots S_2S_1A \mid S_k \ldots S_2S_1E$$
$$M \mid P$$

(4) Analog zu 3.

□

Satz. *Das Produkt endlich vieler Elementarmatrizen ist invertierbar.*

Beweis durch Induktion: Sei $A = S_1S_2 \ldots S_n$ ein Produkt von Elementarmatrizen. Den Fall $n = 1$ haben wir bereits im vorletzten Satz bewiesen. Sei also $B = S_1S_2 \ldots S_{n-1}$ nach Induktionsvoraussetzung invertierbar, dann gilt $AS_n^{-1}B^{-1} = E$ und somit ist auch A invertierbar. □

Definition. Als die *Normalform* bzgl. Zeilen- bzw. Spaltenumformungen sehen wir die Zeilen- bzw. Spaltenstufenform an.

Satz über die P-Q Normalform. *Ist* $A \in K^{m\times n}$, *so existiert ein* $P \in GL(m, K)$ *und ein* $Q \in GL(n, K)$, *so daß*

$$PAQ = \begin{pmatrix} \boxed{E_r} & \boxed{0} \\ \boxed{0} & \boxed{0} \end{pmatrix} \qquad (0 \leq r \leq \min(m, n))$$

ist (Blockform).

Beweis: Formen wir die Matrix A durch elementare Zeilenumformungen zur Zeilenstufenform um, d.h. multiplizieren wir geeignete Elementarmatrizen S_k $(1 \leq k \leq n)$, mit $P = S_1S_2 \ldots S_n$, von links an A und formen anschließend A durch elementare Spaltenumformungen zur Spaltenstufenform um, so erhalten wir sofort die Behauptung. □

Definition. Die Menge aller endlichen Produkte der Elementarmatrizen $T_{ij}(\lambda) \in K^{n\times n}$ ($i \neq j$, $\lambda \in K$) nennen wir *spezielle lineare Gruppe* (engl. *special linear group*) und bezeichnen sie mit $SL(n, K)$.

1.4.3 Darstellung invertierbarer Matrizen durch Elementarmatrizen

Im folgenden wollen wir weitere Eigenschaften invertierbarer Matrizen untersuchen und Berechnungsmethoden angeben.
Sei nun $A \in GL(n, K)$. Mit einer elementaren Zeilenumformung ergibt sich:

$$\widetilde{A} = \widetilde{E}A,$$

wobei $\widetilde{E}$ und $\widetilde{A}$ invertierbare Matrizen sind. Durch weitere, endlich viele elementare Zeilenumformungen erhalten wir

$$M = PA,$$

wobei M invertierbar ist. Ziel ist es, die Matrix M in die Zeilenstufenform zu überführen. Anschließend erhalten wir mit weiteren elementaren Zeilenumformungen die Matrix E.

Satz. *Eine Matrix $A \in GL(n, K)$ läßt sich durch elementare Zeilenumformungen in die Einheitsmatrix E umformen.*

Schema: Die erweiterte Matrix $(A \mid E)$ geht durch elementare Zeilenumformungen in die Matrix $(E \mid P)$ über, wobei P ein Produkt von Elementarmatrizen ist und die Eigenschaft $PA = E$ erfüllt, d.h. $P = A^{-1}$.
Dieses Verfahren kann zur Berechnung der inversen Matrix A^{-1} verwendet werden.

Satz. *Die Inverse des endlichen Produkts von Elementarmatrizen ist wieder ein Produkt von Elementarmatrizen.*

Beweis: Sei $P = S_1S_2 \dots S_n$ ein endliches Produkt von Elementarmatrizen. Mit den Rechenregeln für Inverse folgt $P^{-1} = S_n^{-1} \dots S_2^{-1}S_1^{-1}$. Nach den Formeln aus 1.4.2 sind die S_i^{-1} Elementarmatrizen. □

Satz. *Jedes $A \in GL(n, K)$ ist endliches Produkt von $(n \times n)$-Elementarmatrizen.*

Beweis: Entstehen E_n aus $A \in GL(n, K)$ und P aus E_n durch dieselben elementaren Zeilenumformungen, so gilt $E_n = PA$. Dabei ist P das endliche Produkt von Elementarmatrizen und A die Inverse von P. Da nach obigem Satz die Inverse des endlichen Produkts von Elementarmatrizen wieder ein endliches Produkt von Elementarmatrizen ist, ist A ein endliches Produkt von Elementarmatrizen. □

Satz. *Sei $A \in K^{n\times n}$. Genau dann, wenn aus $Ax = 0$ auch $x = 0$ folgt, ist A invertierbar.*

Beweis: Zunächst folgt für invertierbares A aus $0 = Ax$ sofort $0 = A^{-1} \cdot 0 = A^{-1}Ax = x$.
Gibt es andererseits nur die triviale Lösung $x = 0$, so entsteht beim Gaußalgorithmus (durch entsprechende Multiplikation von Elementarmatrizen) eine obere Dreiecksmatrix mit Diagonalelementen ungleich 0. Weitere Zeilenumformungen „von unten nach oben“ führen zu E. Also ist A ein Produkt von Elementarmatrizen und damit invertierbar (vgl. auch 3.1.5). □

Kapitel 2

Vektorräume

2.1 Vektoren und Vektorräume

Definition. Sei K ein Körper. Eine nichtleere Menge V zusammen mit einer Addition $(x, y) \mapsto x + y$, $x, y \in V$, und einer Skalarmultiplikation $(\lambda, x) \mapsto \lambda \cdot x$, $\lambda \in K, x \in V$, heißt K-*Vektorraum* (oder auch *linearer Raum über* K), wenn folgende *Vektorraumaxiome* erfüllt sind:

$$x + y = y + x \quad \text{für alle } x, y \in V \tag{V1}$$
$$x + (y + z) = (x + y) + z \quad \text{für alle } x, y, z \in V \tag{V2}$$
$$\text{es gibt } 0 \in V \text{ mit } x + 0 = x \text{ für alle } x \in V \tag{V3}$$
$$\text{Zu jedem } x \in V \text{ gibt es ein } -x \in V \text{ mit } x + (-x) = 0 \tag{V4}$$
$$1 \cdot x = x \quad \text{für alle } x \in V \text{ (Eins } 1 \in K) \tag{V5}$$
$$\alpha(\beta x) = (\alpha\beta)x \quad \text{für alle } \alpha, \beta \in K, x \in V \tag{V6}$$
$$(\alpha + \beta)x = \alpha x + \beta x \quad \text{für alle } \alpha, \beta \in K, x \in V \tag{V7}$$
$$\alpha(x + y) = \alpha x + \alpha y \quad \text{für alle } \alpha \in K, x, y \in V \tag{V8}$$

Die Elemente aus V heißen *Vektoren*, $x + y$ *Summe der Vektoren* x, y und αx *heißt* α*-faches von* x.

Beispiele für Vektorräume:

1. Die Menge K_n ist bei folgender Addition zweier Vektoren $a, b \in K_n$
$$a + b = (\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n),$$
wobei α_i, β_i $(1 \leq i \leq n)$ die Komponenten von a bzw. b sind, und bei folgender Multiplikation eines Vektors a mit einem Skalar λ
$$\lambda \cdot a = \lambda(\alpha_1, \dots, \alpha_n) = (\lambda\alpha_1, \dots, \lambda\alpha_n),$$
wobei α_i $(1 \leq i \leq n)$ die *Komponenten* von a sind, ein Vektorraum.
Analog bilden auch die Mengen K^n und $K^{m\times n}$ Vektorräume.

2. Sei X eine nichtleere Menge und V ein K-Vektorraum. Dann ist
$$V^X := \mathrm{Abb}(X, V) := \{ f \mid f\colon X \to V \}$$
ebenfalls ein Vektorraum mit der Addition $(f + g)(x) = f(x) + g(x)$ und der Skalarmultiplikation $(\lambda f)(x) = \lambda f(x)$ mit $f, g \in V^X$. Wie wir sehen *erbt* V^X die Vektorraumeigenschaften von V.

◊

2.2 Unterräume, Linearkombinationen, lineare Hülle

2.2.1 Unterräume

Definition. Sei V ein K-Vektorraum. Eine nichtleere Menge $U \subseteq V$ heißt *Untervektorraum* (*Unterraum* oder *linearer Teilraum*) von V, wenn gilt

$$x, y \in U \implies x + y \in U, \tag{U1}$$
$$x \in U \text{ und } \alpha \in K \implies \alpha x \in U. \tag{U2}$$

Somit ist der Unterraum U bezüglich der Addition und Skalarmultiplikation abgeschlossen.

Lemma 2.1: *Jeder Unterraum ist mit gegebener Addition und Skalarmultiplikation selbst ein* K*-Vektorraum.*

Beweis: Beweis des Axioms (V3): Da U nach Voraussetzung nichtleer ist gibt es mindestens ein Element $u \in U$. Nach (U2) ist auch $-u \in U$ und nach (U1) auch deren Summe, $u + (-u) = 0 \in U$. Es gilt: $0_U = 0_V$. Die anderen Axiome sind, da sie sogar in ganz V gelten, auch erfüllt. □

Unterraum-Test. *Seien V ein K-Vektorraum und $U \subseteq V$. Die Menge U ist genau dann ein Unterraum, wenn U nichtleer ist und für alle $\alpha \in K$ und $u, v \in U$ auch $\alpha u + v$ in U ist.*

Lemma 2.2: *Seien V ein K-Vektorraum und $(U_i)_{i \in I}$ eine Familie von Untervektorräumen aus V. Dann ist*

$$U := \bigcap_{i \in I} U_i = \left\{ u \;\middle|\; u \in U_i \quad \text{für alle } i \in I \right\}$$

ein Untervektorraum von V.

Beweis: Da 0_V in allen U_i ($i \in I$) enthalten ist, gilt $0_V \in U$. Seien $v, w \in U$ und $\alpha \in K$ beliebig. Dann gilt $v, w \in U_i$ für alle $i \in I$ und damit auch $v + \alpha w \in U_i$. Daraus folgt $v + \alpha w \in U$. □

Lemma 2.3: *Seien V ein Vektorraum und $U_1, \ldots, U_r \subseteq V$ Untervektorräume. Dann ist der* Summenraum

$$S := U_1 + \cdots + U_r := \left\{ u_1 + \cdots + u_r \;\middle|\; u_i \in U_i \quad (1 \le i \le r) \right\}$$

ein Untervektorraum von V.

Beweis: Seien $v, w \in S$ und $\alpha \in K$. Dann gibt es $u_i, u_i' \in U_i$ ($1 \le i \le r$) derart, daß $v = u_1 + \cdots + u_r$ und $w = u_1' + \cdots + u_r'$. Da $u_i + \alpha u_i' \in U_i$ ($1 \le i \le r$), ist auch $v + \alpha w = u_1 + \alpha u_1' + \cdots + u_r + \alpha u_r' \in S$. □

2.2.2 Linearkombinationen, lineare Hülle

Im folgenden sei V stets ein K-Vektorraum.

Definition. Gegeben sei eine Familie $(v_1, \ldots, v_n)$ der Vektoren $v_i \in V$. Jeder Vektor $v \in V$ der Form

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i$$

mit Koeffizienten $\alpha_i \in K$ heißt *Linearkombination* dieser Familie bzw. kurz „der v_i".

Definition. Die Menge

$$\mathrm{Lin}(v_1, \ldots, v_n) := \left\{ \sum_{i=1}^{n} \alpha_i v_i \;\middle|\; \alpha_i \in K, 1 \le i \le n \right\} = \sum K \cdot v_i$$

heißt *lineare Hülle* der Familie $(v_1, \ldots, v_n)$.

Lemma 2.4: *Die lineare Hülle* $\mathrm{Lin}(v_1, \ldots, v_n)$ *ist ein Untervektorraum von V.*

Beweis: Die Menge $Kv := \{ \alpha v \mid \alpha \in K \}$ mit $v \in V$ ist ein Unterraum von V, denn seien $u, u' \in Kv$ und $\gamma \in K$, so gibt es ein $\alpha \in K$ mit $u = \alpha v$ und ein $\beta \in K$ mit $\gamma u' = \beta v$, also $u + \gamma u' \in Kv$. Betrachten wir also die Unterräume $Kv_i \subseteq V$ ($1 \le i \le n$), so ist nach Lemma 2.2.1 auch $Kv_1 + \cdots + Kv_n$ ein Unterraum von V und damit auch $\mathrm{Lin}(v_1, \ldots, v_n)$. □

Definition. Der K-Vektorraum V heißt *endlich erzeugt* wenn es eine endliche Familie von Vektoren $v_i \in V$ gibt, mit

$$V = \mathrm{Lin}(v_1, \ldots, v_n).$$

In diesem Fall heißt $(v_1, \ldots, v_n)$ ein *Erzeugendensystem*.

Definition. Seien V ein K-Vektorraum und $(v_1, \ldots, v_n)$ eine endliche Familie von Vektoren mit $v_i \in V$.

- Die Familie heißt *linear abhängig*, wenn es $\alpha_1, \ldots, \alpha_n \in K$ gibt, so daß *nicht alle* α_i *Null sind*, d.h. $(\alpha_1, \ldots, \alpha_n) \neq 0$, und
$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

- Die Familie $(v_1, \ldots, v_n)$ heißt *linear unabhängig*, wenn sie nicht linear abhängig sind, d.h.
$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \implies \alpha_1 = \cdots = \alpha_n = 0.$$

Lemma 2.5: *Sei $n > 1$. Die Vektoren $v_1, \ldots, v_n$ sind genau dann* linear abhängig, *wenn es ein i mit $1 \le i \le n$ gibt, so daß sich $v_i \neq 0$ als Linearkombination der übrigen darstellen läßt.*

Beweis: Läßt sich v_i ($1 \le i \le n$) als Linearkombination der übrigen v_k mit $1 \le k \le n$ und $k \neq i$ darstellen, so folgt aus

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} - 1 \cdot v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_n v_n = 0,$$

daß die Vektoren $v_1, \ldots, v_n$ linear abhängig sind.
Sind umgekehrt die Vektoren $v_1, \ldots, v_n$ linear abhängig, so gibt es ein $i \in \{1, 2, \ldots, n\}$ mit $\alpha_i \neq 0$ und

$$\alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_i v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_n v_n = 0.$$

Daraus folgt dann

$$v_i = -\frac{\alpha_1}{\alpha_i} v_1 - \cdots - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \cdots - \frac{\alpha_n}{\alpha_i} v_n.$$

□

Satz. *Sei $V = \mathrm{Lin}(v_1, \ldots, v_n)$ ein endlich erzeugter Vektorraum. Dann ist jede Familie mit $n+1$ Vektoren $w_i \in V$ linear abhängig.*

Beweis: Nach Voraussetzung besitzt jeder Vektor w_j eine Darstellung der Form

$$w_j = \sum_{i=1}^{n} \alpha_{ij} v_i \quad \text{mit } \alpha_{ij} \in K,$$

also

$$\begin{aligned} w_1 &= \alpha_{11} v_1 + \alpha_{21} v_2 + \cdots + \alpha_{n1} v_n, \\ &\vdots \\ w_{n+1} &= \alpha_{1|n+1|} v_1 + \alpha_{2|n+1|} v_2 + \cdots + \alpha_{n|n+1|} v_n. \end{aligned}$$

Dazu definieren wir uns eine Matrix $A := (\alpha_{ij})_{n \times |n+1|}$. Dann gibt es einen Vektor $u := (\eta_1, \ldots, \eta_{n+1})^T \neq 0$, der die Gleichung $Au = 0$ erfüllt, also gilt:

$$\sum_{j=1}^{n+1} \alpha_{ij} \eta_i = 0 \quad (1 \leq i \leq n).$$

Multiplizieren wir die einzelnen Zeilen aus dem oben definierten linearen Gleichungssystem mit η_i, so erhalten wir

$$\sum_{i=1}^{n+1} \eta_i w_i = \underbrace{\left(\sum_{j=1}^{n+1} \alpha_{ij} \eta_j \right)}_{=0} v_1 + \cdots + \underbrace{\left(\sum_{j=1}^{n+1} \alpha_{nj} \eta_j \right)}_{=0} v_n = 0.$$

Da aber nach der Definition von u nicht alle η_j gleich Null sind, folgt nun daß die Vektoren w_j *linear abhängig* sind. □

2.3 Basis und Dimension

2.3.1 Basis

Im folgenden sei V ein K-Vektorraum.

Definition. Eine Familie $B = (v_1, \ldots, v_n)$ von Vektoren $v_i \in V$ heißt *Basis* von V, wenn gilt

$$V = \mathrm{Lin}(v_1, \ldots, v_n) \quad \text{und} \tag{B1}$$
$$(v_1, \ldots, v_n) \text{ linear unabhängig.} \tag{B2}$$

Zur Vermeidung von Fallunterscheidungen definieren wir, daß der Vektorraum $\{0\}$ die Basis $\emptyset$ hat.

Definition. Sei $V = K^n$. Die Basis $E = (e^1, \ldots, e^n)$ heißt dann *natürliche* oder auch *kanonische Basis des* K^n.

Satz über die Charakterisierung einer Basis. *Für $B = (v_1, \ldots, v_n)$ sind folgende Aussagen äquivalent:*

(1) B *ist Basis von* V.

(2) *Jeder Vektor $v \in V$ hat eine eindeutige Darstellung der Form:*

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \quad \textit{mit } \alpha_i \in K.$$

(3) B *ist minimales Erzeugendensystem von* V, *d.h. in* $V = \mathrm{Lin}(v_1, \ldots, v_n)$ *kann kein* v_i *weggelassen werden.*

(4) B *ist maximale linear unabhängige Familie, d.h.* B *ist linear unabhängig und mit jedem $v \in V$ ist $(v_1, \ldots, v_n, v)$ linear abhängig.*

Beweis:

(1)$\Longrightarrow$*(2)*: Ein beliebiger Vektor $v \in V$ kann nach (B1) als Linearkombination der Basisvektoren dargestellt werden. Seien nun

$$v = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i v_i$$

zwei derartige Darstellungen. Dann folgt unmittelbar

$$0_V = \sum_{i=1}^{n} (\alpha_i - \beta_1) v_i.$$

Da die Basisvektoren nach (B2) linear unabhängig sind folgt nun $\alpha_i = \beta_i$, also ist die Darstellung *eindeutig*.

(2)$\Longrightarrow$*(1)*: Sei nun $v = \alpha_1 v_1 + \cdots + \alpha_n v_n = 0$ eine Darstellung des Nullvektors als Linearkombination der Basisvektoren. Offensichtlich hat der Nullvektor auch noch folgende Darstellung: $0 \cdot v_1 + \cdots + 0 \cdot v_n = 0$. Da die Darstellung nach Voraussetzung eindeutig ist, folgt $\alpha_1 = \cdots = a_n = 0$. Also sind die Basisvektoren linear unabhängig.

(*1*)$\Longleftrightarrow$(*3*): Wir verweisen bei diesem Beweis auf Lemma 2.2.2.

(*1*)$\Longrightarrow$(*4*): Sei $v = \sum \alpha_i v_i$ eine Linearkombination (nach (B1)) der nach (B2) linear unabhängigen Basisvektoren $(v_1, \ldots, v_n)$. Nun folgt aus $\alpha_1 v_1 + \cdots + \alpha_n v_n + (-1)v = 0$ die lineare Abhängigkeit der Familie $(v_1, \ldots, v_n, v)$.

(*4*)$\Longrightarrow$(*1*): Sei nun $v \in V$ und nach Voraussetzung $(v_1, \ldots, v_n, v)$ eine linear abhängige Familie von Vektoren. Damit gibt es $\alpha_i, \alpha \in K$ derart, daß $\alpha_1 v_1 + \cdots + \alpha_n v_n + \alpha v = 0$ und nicht alle Koeffizienten α_i, α Null sind. Hierin ist nach Voraussetzung $\alpha \neq 0$. (Andernfalls wäre $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$ und folglich $\alpha_1 = \cdots = \alpha_n = 0$, was der Annahme, daß nicht alle α_i, α Null sind, widerspricht!) Also besitzt der Vektor v folgende Darstellung: $v = -\alpha_1 v_1/\alpha - \cdots - \alpha_n v_n/\alpha$. Damit ist (B1) erfüllt.

Die Bedingung (B2) folgt direkt aus der Voraussetzung.

□

Hauptsatz für endlich erzeugte Vektorräume. *Sei V ein K-Vektorraum mit $V = \mathrm{Lin}(w_1, \ldots, w_r)$. Dann gilt:*

(*1*) Existenz einer Basis:

$$V \textit{ besitzt eine Basis } (v_1, \ldots, v_n) \textit{ mit } v_i \in \{w_1, \ldots, w_r\}.$$

(*2*) Basisergänzung: *Ist $(u_1, \ldots, u_m)$ eine linear unabhängige Familie von Vektoren aus V, dann ist $(u_1, \ldots, u_m)$ eine Basis oder läßt sich zu einer Basis*

$$(u_1, \ldots, u_m, u_{m+1}, \ldots, u_n) \quad \textit{mit } u_{m+1}, \ldots, u_n \in \{w_1, \ldots, w_r\}$$

ergänzen.

(*3*) Invarianz der Basislänge: *Sind $B = (v_1, \ldots, v_n)$ und $C = (u_1, \ldots, u_m)$ Basen von V, dann gilt $m = n$.*

Beweis:

(*1*) Der Beweis erfolgt *konstruktiv* „durch Weglassen“. Betrachten wir die Familie von Vektoren $(w_1, \ldots, w_r)$. Sind diese linear unabhängig, so stellen sie bereits eine Basis dar. Ansonsten kann nach Lemma 2.2.2 ein w_i zur Erzeugung von V weggelassen werden. Nun überprüfen wir wieder, ob die Familie jetzt linear unabhängig ist. Dies wird solange wiederholt, bis die Familie linear unabhängige Vektoren enthält. Diese bilden dann eine Basis.

(*2*) Der Beweis erfolgt *konstruktiv* „durch Hinzufügen“. Betrachten wir den Vektorraum $V = \mathrm{Lin}(u_1, \ldots, u_m, w_1, \ldots, w_r)$. Entweder ist die Familie $(u_1, \ldots, u_m)$ bereits eine Basis oder es existiert ein w_i, sodaß die Familie $(u_1, \ldots, u_m, w_i)$ linear unabhängig ist. Wir setzen nun $u_{m+1} := w_i$ und wiederholen den Vorgang solange, bis wir eine Basis des Vektorraumes V erhalten.

(*3*) Seien $B = (v_1, \ldots, v_n)$ und $C = (u_1, \ldots, u_n)$ Basen des Vektorraumes V. Da B eine Basis darstellt sind (nach Satz aus 2.2.2) je $n+1$ Vektoren linear abhängig. Also muß $m < n+1$ oder auch $m \leq n$ gelten. Aufgrund der Symmetrie in den Voraussetzungen folgt, daß auch $n \leq m$ sein muß. Daraus folgt unmittelbar $m = n$.

□

2.3.2 Dimension

Definition. Sei V ein endlich erzeugter Vektorraum. Die nach dem Hauptsatz eindeutig bestimmte Länge n einer Basis von V heißt *Dimension* von V und wir sagen, V ist *endlichdimensional*. In Zeichen:

$$n = \mathrm{Dim}\, V.$$

Zur Vermeidung von Fallunterscheidungen definieren wir

$$\mathrm{Dim}\{0\} := 0 \quad \text{und}$$
$$\mathrm{Dim}\, V := \infty, \quad \text{falls } V \text{ nicht endlich erzeugt ist.}$$

Beispiele:

1. Es gilt $\mathrm{Dim}\, K^{m\times n} = m \cdot n$, denn die Familie $(E_{11}, E_{12}, \ldots, E_{21}, \ldots, E_{mn})$ ist eine Basis von $K^{m\times n}$.
2. Ist P_n der Vektorraum der Polynome vom Grad n. Dann ist $\mathrm{Dim}\, P_n = n + 1$.

◊

Satz. *Seien V ein endlich erzeugter (endlichdimensionaler) K-Vektorraum und $U \subseteq V$ ein Untervektorraum von V. Dann ist*

(*1*) $\mathrm{Dim}\, U \leq \mathrm{Dim}\, V$.

(*2*) $\mathrm{Dim}\, U = \mathrm{Dim}\, V \Longleftrightarrow U = V$.

Beweis:

(*1*) Ist $U = \{0\}$, dann ist $\mathrm{Dim}\, U = 0$ immer kleiner oder gleich $\mathrm{Dim}\, V$. Ansonsten können wir einen Vektor $u \in U$ mit $u \neq 0$ wählen. Da, nach dem Satz aus 2.2.2, je $(\mathrm{Dim}\, V)+1$ Vektoren aus V linear abhängig sind, hat also jede maximale, linear unabhängige Familie von Vektoren aus U eine Länge $l \leq \mathrm{Dim}\, V$.

(*2*) Sei die Familie $(u_1, \ldots, u_n)$ eine Basis von U. Da nichts ergänzt oder weggelassen werden kann, ist die Familie auch eine Basis von V.

□

Dimensionsformel. *Seien U_1 und U_2 endlichdimensionale Unterräume eines Vektorraumes. Dann sind die Vektorräume $U_1 + U_2$ und $U_1 \cap U_2$ endlichdimensional und es gilt:*

$$\mathrm{Dim}\, U_1 + \mathrm{Dim}\, U_2 = \mathrm{Dim}(U_1 + U_2) + \mathrm{Dim}(U_1 \cap U_2).$$

Beweis: Wir stellen fest, daß die Summe $U_1 + U_2$ zweier endlich erzeugter Vektorräume wieder endlich erzeugt ist.
Sei nun $(u_1, \ldots, u_r)$ eine Basis von $U_1 \cap U_2$, welche wir zu der Basis $(u_1, \ldots, u_r, v_1, \ldots, v_s)$ von U_1 und zu der Basis $(u_1, \ldots, u_r, w_1, \ldots, w_t)$ von U_2 ergänzen.

Nun zeigen wir, daß $B := (u_1, \ldots, u_r, v_1, \ldots, v_s, w_1, \ldots, w_t)$ eine Basis von $U_1 + U_2$ ist. Zunächst stellen wir fest, daß es zu einem beliebigen Vektor $w \in U_1 + U_2$ zwei Vektoren $u \in U_1$ und $v \in U_2$ gibt, mit $w = u + v$. Diese haben eine Darstellung $u = \sum \alpha_i u_i + \sum \beta_j v_j$ und $v = \sum \alpha'_i u_i + \sum \gamma_k w_k$. Daraus folgt unmittelbar, daß B ein Erzeugendensystem von $U_1 + U_2$ ist.
Ist nun $\sum \alpha_i u_i + \sum \beta_j v_j + \sum \gamma_k w_k = 0$, so setzen wir $v := \sum \alpha_i u_i + \sum \beta_j v_j \in U_1$ und erhalten daraus $v = -\sum \gamma_k w_k \in U_2$. Also ist $v \in U_1 \cap U_2$. Da die oben bezeichneten Basen von U_1 und U_2 aber nur die u_i's gemeinsam haben, gibt es $\lambda_i \in K$ mit $v = \lambda_1 u_1 + \cdots + \lambda_r u_r$. Da die Linearkombination von v eindeutig ist, folgt nun $\lambda_1 = \alpha_1, \ldots, \lambda_r = \alpha_r$ und außerdem $\beta_1 = \cdots = \beta_s = 0$. Nun folgt aus der ursprünglichen Annahme und der linearen Unabhängigkeit der Basis von U_2 sofort, daß auch $\alpha_1 = \cdots = \alpha_r = \gamma_1 = \cdots = \gamma_t = 0$ gilt.
Also ist B eine Basis von $U_1 + U_2$. Dies ergibt:

$$\begin{aligned}\operatorname{Dim}(U_1 + U_2) &= r + s + t = \operatorname{Dim}(U_1) + (\operatorname{Dim} U_2) - r \\ &= \operatorname{Dim}(U_1) + \operatorname{Dim}(U_2) - \operatorname{Dim}(U_1 \cap U_2).\end{aligned}$$

□

Definition. Die Summe $U_1 + U_2$ heißt *direkt*, wenn $U_1 \cap U_2 = \{0\}$. In diesem Fall schreiben wir $U_1 \oplus U_2$. Es gilt:

$$\operatorname{Dim}(U_1 \oplus U_2) = \operatorname{Dim} U_1 + \operatorname{Dim} U_2.$$

Satz. *Seien V ein Vektorraum und U_1, U_2 Unterräume von V mit den Basen $(u_1, \ldots, u_r)$ und $(w_1, \ldots, w_s)$. Gilt dann $V = U_1 \oplus U_2$, so ist $(u_1, \ldots, u_r, w_1, \ldots, w_s)$ eine Basis von V.*

Beweis: Offensichtlich ist $\operatorname{Lin}(u_1, \ldots, u_r, w_1, \ldots, w_s)$ ein Erzeugendensystem von V. Würden wir nun o.B.d.A. u_1 weglassen, so wäre $\operatorname{Lin}(u_2, \ldots, u_r, w_1, \ldots, w_s)$ kein Erzeugendensystem, denn aufgrund der Voraussetzung erkennen wir sofort, daß sich u_1 nicht als Linearkombination der Vektoren $(u_2, \ldots, u_r, w_1, \ldots, w_s)$ darstellen läßt, d.h. $\operatorname{Lin}(u_1, \ldots, u_r, w_1, \ldots, w_s)$ ist ein minimales Erzeugendensystem von V und damit eine Basis. □

Definition. Die Summe V mehrerer Untervektorräume V_i ($i \in I$) heißt *direkt*, wenn gilt

$$V = \sum_{i \in I} V_i \quad \text{und} \quad V_i \cap \sum_{j \in I \setminus \{i\}} V_j = \{0\} \quad \text{für alle } i \in I.$$

2.4 Der Vektorraum der Polynome

2.4.1 Definitionen und Eigenschaften

Definition. Eine Abbildung $f\colon K \to K$ heißt *Polynomfunktion*, wenn es $\alpha_0, \ldots, \alpha_k \in K$ gibt mit

$$f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_k x^k \quad \text{für alle } x \in K.$$

Bemerkung: Im allgemeinen sind die Koeffizienten α_i nicht eindeutig bestimmt. Dies ist vom Körper K abhängig. Betrachten wir dazu den Körper $\mathbf{Z}_2 = \{\bar{0}, \bar{1}\}$ und die Polynomfunktion $f(x) = x^2 + x$. Aus $f(\bar{0}) = \bar{0}$ und $f(\bar{1}) = \bar{0}$ ergibt sich $f \equiv 0$, obwohl f und das Nullpolynom verschiedene Koeffizienten α_i haben.
Wir beschränken uns deshalb im folgenden auf unendliche Körper, bei denen dieses Problem nicht eintritt, wie z.B. $\mathbf{Q}$, $\mathbf{R}$ oder $\mathbf{C}$.

Definition. Ein Polynom heißt *normiert*, wenn der führende Koeffizient α_k Eins ist.

Satz über den Koeffizientenvergleich. *Sei K ein unendlicher Körper und $f(x) = \sum_{i=0}^{n} \alpha_i x^i$ eine Polynomfunktion. Dann gilt:*

(1) $f(x) = 0$ *für alle* $x \iff \alpha_0 = \cdots = \alpha_n = 0$.

(2) $\sum_{i=0}^{n} \alpha_i x^i = \sum_{i=0}^{n} \beta_i x^i \iff \alpha_i = \beta_i$ *für alle* i.

Beweis:

(1) Wähle $n+1$ paarweise verschiedene Körperelemente $x_1, \ldots, x_{n+1}$ aus K aus. (Dies können wir nur aufgrund der Voraussetzung $|K| = \infty$ fordern!) Dann folgt aus $f(x) = 0$ für alle x folgendes lineare Gleichungssystem:

$$\begin{aligned}\alpha_0 + \alpha_1 x_1 + \alpha_2 x_1^2 + \cdots + \alpha_n x_1^n &= 0, \\ &\vdots \\ \alpha_0 + \alpha_1 x_{n+1} + \alpha_2 x_{n+1}^2 + \cdots + \alpha_{n+1} x_{n+1}^{n+1} &= 0.\end{aligned}$$

Stellen wir zur Lösung dieses Gleichungssystems die *Determinante* auf (siehe Kapitel 4), so erhalten wir die *Vandermonde-Determinante*, die wegen $x_i \neq x_j$ ungleich 0 ist:

$$\det \begin{pmatrix} 1 & x_1 & \ldots & x_1^n \\ 1 & x_2 & \ldots & x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n+1} & \ldots & x_{n+1}^{n+1} \end{pmatrix} \neq 0.$$

(2) Folgt aus *(1)*.

□

Betrachten wir unendliche Körper, so sind aufgrund des Satzes über den Koeffizientenvergleich Polynomfunktionen und Polynome im wesentlichen dasselbe!

Definition. Die *Menge aller Polynome* über dem Körper K definieren wir wie folgt:

$$K[x] := \left\{ \sum_{i=0}^{n} \alpha_i x^i \;\middle|\; n \geq 0,\ \alpha_1, \ldots, \alpha_n \in K \right\}.$$

Dazu definieren wir die Addition zweier Polynome aus $K[x]$ durch

$$\sum \alpha_i x^i + \sum \beta_i x^i = \sum (\alpha_i + \beta_i) x^i$$

und die Skalarmultiplikation durch

$$\alpha\left(\sum \alpha_i x^i\right) = \sum (\alpha\alpha_i) x^i.$$

Damit ist $K[x]$ ein K-Vektorraum.

Definition. Die Multiplikation zweier Polynome definieren wir folgendermaßen:

$$\left(\sum_{i=0}^{n} \alpha_i x^i\right) \cdot \left(\sum_{j=0}^{m} \beta_j x^j\right) = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} \alpha_i \beta_j\right) x^k.$$

Hiermit wird $K[x]$ ein *kommutativer Ring mit Eins* ($f(x) \equiv 1$). Er heißt *Polynomring* in x über K.

Definition. Sei $f = \sum_{i=0}^{n} \alpha_i x^i$ mit $\alpha_n \neq 0$ ein Polynom. Dann heißt n *Grad des Polynoms* f. Wir schreiben

$$\operatorname{Grad} f = n.$$

Das Nullpolynom besitzt keinen Grad.

Gradsatz. *Seien $f, g \neq 0$ zwei Polynome. Dann gilt $fg \neq 0$ und*

$$\operatorname{Grad}(fg) = \operatorname{Grad} f + \operatorname{Grad} g.$$

Beweis: Der Beweis ist direkt aus der Definition des Polynomprodukts abzulesen. □

2.4.2 Polynomdivision

Satz über die Polynomdivision. *Zu $f, g \in K[x]$ und $g \neq 0$ existieren $h, r \in K[x]$ mit $r = 0$ oder $\operatorname{Grad} r < \operatorname{Grad} g$ mit folgender Eigenschaft:*

$$f(x) = h(x)g(x) + r(x).$$

Anders gelesen bedeutet dies:

$$\frac{f(x)}{g(x)} = h(x) + \frac{r(x)}{g(x)},$$

wobei $r(x)$ der Rest der Division ist.

Beweis: Der Beweis erfolgt *konstruktiv* und stellt den *Divisionsalgorithmus für Polynome* dar:

Fall 1: $f = 0$. Wir setzen $h := 0$ und $r := 0$.

Fall 2: $f \neq 0$, $\operatorname{Grad} f < \operatorname{Grad} g$. Wir setzen $h := 0$ und $r := f$.

Fall 3: $f \neq 0$, $\operatorname{Grad} g \leq \operatorname{Grad} f$. Offenbar gilt:

$$f(x) = \frac{\alpha_n}{\beta_m} x^{n-m} g(x) + f_1(x) \quad \text{mit } f_1(x) := f(x) - \frac{\alpha_n}{\beta_m} x^{n-m} g(x).$$

Dabei ist $f_1 = 0$ oder besitzt einen kleineren Grad als f. Ist $f_1 \neq 0$ und gilt $\operatorname{Grad} g \leq \operatorname{Grad} f$, so wiederholen wir diesen Schritt mit f_1 an Stelle von f, ansonsten sind wir fertig.

□

2.4.3 Ideale

Definition. Sei I eine nichtleere Teilmenge von $K[x]$. Dann heißt I *Ideal*, wenn gilt:

$$f, g \in I \implies f + g \in I \quad \text{und} \tag{I1}$$

$$f \in I,\ p \in K[x] \implies fp \in I. \tag{I2}$$

Definition. Sei ein Polynom $p \in K[x]$ gegeben. Dann nennen wir

$$I := \left\{ f(x) \cdot p(x) \;\middle|\; f \in K[x] \right\} = p(x)K[x]$$

Hauptideal.

Satz. *Jedes Ideal I in $K[x]$ ist ein Hauptideal.*

Beweis: Es ist zu zeigen, daß zu jedem Ideal aus $K[x]$ ein $p(x) \in I$ existiert derart, daß $I = p(x)K[x]$.
Ist $I = \{0\}$, so ist $I = 0 \cdot K[x]$. Sei also im folgenden $I \neq \{0\}$ und $p(x) \in I$ das Polynom mit minimalem Grad in I. (Ein solches Polynom existiert, denn der Grad ist eine natürliche Zahl und in einer Menge natürlicher Zahlen gibt es immer eine kleinste Zahl.) Ferner sei f ein beliebiges Polynom in I. Nun gibt es ein Polynom $h \in I$, so daß $f(x) = h(x) \cdot p(x) + r(x)$, wobei $r = 0$ oder $\operatorname{Grad} r < \operatorname{Grad} p$. Da p aber das Polynom mit kleinstem Grad in I ist, muß $r = 0$ sein, d.h. $I = p(x)K[x]$. □

Definition. Seien $f(x) \in K[x]$ und $\alpha \in K$. Ist $f(\alpha) = 0$, so nennen wir α *Nullstelle* oder *Wurzel* von f.

Allgemeine Charakterisierung von Nullstellen: Genau dann, wenn $\alpha \in K$ eine Nullstelle von f ist, gibt es ein $g(x) \in K[x]$ mit $f(x) = (x - \alpha)g(x)$.

Bemerkung: Gilt $f(x) = (x - \alpha)^k g(x)$ mit $g(\alpha) \neq 0$, so nennen wir α eine k-*fache Nullstelle*. Sind $\alpha_1, \alpha_2, \ldots, \alpha_r$ alle Nullstellen von f und zwar $k_1, k_2, \ldots, k_r$-fache Nullstellen, so gilt

$$f(x) = (x - \alpha_1)^{k_1} \ldots (x - \alpha_r)^{k_r} \cdot h(x) \quad \text{für alle } x,$$

wobei $h(x) \neq 0$. Dann ist $\operatorname{Grad} f = k_1 + k_2 + \cdots + k_r + \operatorname{Grad} h$. Die Anzahl der Nullstellen von f ist damit immer kleiner oder gleich $\operatorname{Grad} f$.

Spezialfälle für $\mathbf{R}$ *und* $\mathbf{C}$: Jedes Polynom f über $\mathbf{C}$ oder $\mathbf{R}$ ist in $\mathbf{C}$ vollständig zerlegbar, d.h. es gibt eine Darstellung der Art $f(x) = c \cdot \prod_i (x - \alpha_i)^{k_i}$ mit $\alpha_i \in \mathbf{C}$.
Polynome in $\mathbf{R}$ sind über $\mathbf{R}$ in Faktoren vom Grad 1 oder 2 zerlegbar.

Beispiele für Ideale:

1. Die Mengen $\{0\}$ und $K[x]$ sind *triviale* Ideale.
2. Seien zwei Polynome $f_1, f_2 \in K[x]$ gegeben. Dann ist ganz offensichtlich
$$I := f_1(x)K[x] + f_2(x)K[x] = \{\, p(x)f_1(x) + q(x)f_2(x) \mid p, q \in K[x] \,\}$$
ein Ideal, erzeugt von f_1 und f_2.
3. Sei $\alpha \in K$ gegeben. Dann ist
$$I := \{\, f(x) \mid f \in K[x], f(\alpha) = 0 \,\}$$
ein Ideal.
4. Ist f ein Endomorphismus (siehe Abschnitt 3.1.1), so ist
$$I := \{\, p(x) \mid p \in K[x], p(f) = 0 \,\}$$
ein Ideal.

$\Diamond$

Kapitel 3

Lineare Abbildungen

3.1 Definitionen und elementare Eigenschaften

3.1.1 Definitionen

Definition. Seien V und W zwei K-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *linear*, falls gilt:
$$f(v_1 + v_2) = f(v_1) + f(v_2) \quad \text{und}$$
$$f(\alpha v) = \alpha f(v)$$
für alle $\alpha \in K$ und für alle $v, v_1, v_2 \in V$.

Definition. Wir definieren:
$$\operatorname{Bild} f := f(V) := \{\, f(v) \mid v \in V \,\},$$
$$\operatorname{Kern} f := \{\, v \in V \mid f(v) = 0 \,\}.$$

Bemerkung: Sei $f\colon V \to W$ eine lineare Abbildung. Dann ist der Kern von f ein Unterraum von V und das Bild von f ein Unterraum von W (siehe auch Lemma 3.1.2).

Lemma 3.1: *Folgende Aussagen für* $f\colon V \to W$ *sind äquivalent:*

(1) f *ist linear.*

(2) $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$ *für alle* $u, v \in V$ *und für alle* $\alpha, \beta \in K$.

(3) $f(\alpha u + v) = \alpha f(u) + f(v)$ *für alle* $u, v \in V$ *und für alle* $\alpha \in K$.

Beweis: Wir wollen hier nur die Äquivalenz der ersten beiden Aussagen zeigen. Der Beweis der Äquivalenz der dritten Gleichung mit den anderen beiden ist analog durchzuführen.

(1)$\Longrightarrow$*(2)*: Da f linear ist, gilt:
$$f(\alpha u + \beta v) = f(\alpha u) + f(\beta v) = \alpha f(u) + \beta f(v).$$

(2)$\Longrightarrow$*(1)*: Für $\alpha = \beta = 1$ bzw. $\beta = 0$ erhalten wir jeweils aus der zweiten Gleichung die beiden Bedingungen für eine lineare Abbildung.

$\square$

3.1.2 Elementare Eigenschaften

Lemma 3.2: *Sei im folgenden* $f\colon V \to W$ *eine lineare Abbildung. Dann gilt:*

(1) $f(0_V) = 0_W$,

(2) $f(-v) = -f(v)$ *für alle* $v \in V$,

(3) $f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i f(v_i) \quad \alpha_i \in K,\ v_i \in V \quad (1 \le i \le n)$.

Beweis:

(1) Die Behauptung ergibt sich sofort aus $f(0_V) = f(0_K \cdot v) = 0_K \cdot f(v) = 0_W$ für alle Vektoren $v \in V$.

(2) Sei $v \in V$, $w \in W$ und $-v$ das Inverse zu v. Dann erhalten wir die Behauptung aus $0_W = f(0_V) = f(v + (-v)) = f(v) + f(-v)$ und aus $0_W + w = w$.

(3) Dies ergibt sich sofort aus Lemma 3.1.1.

□

Lemma 3.3: *Seien* $u, v \in V$ *und sei* f *linear. Dann sind folgende Aussagen äquivalent:*

(1) $f(u) = f(v)$.

(2) $f(u - v) = 0_W$.

(3) $u - v \in \operatorname{Kern} f$.

Beweis: Seien $u, v \in V$. Dann folgt wegen $0_W = f(u) - f(v) = f(u) + f(-v) = f(u - v)$ aus *(1)* Gleichung *(2)* und umgekehrt. Aufgrund der Definition des Kerns folgt aus $(u - v) \in \operatorname{Kern} f$, daß $f(u - v) = 0_W$. Somit sind auch die Gleichungen *(2)* und *(3)* äquivalent. □

Lemma 3.4: *Ist* f *linear, so ist* f *genau dann injektiv, wenn* $\operatorname{Kern} f = \{0_V\}$.

Beweis: Sei f injektiv und $u \in \operatorname{Kern} f$. Dann ist $f(u) = 0_W = f(0_V)$, also $u = 0_V$. Damit ist $\operatorname{Kern} f = \{0_V\}$.
Nun seien $\operatorname{Kern} f = \{0\}$ und $f(u) = f(v)$. Dann gilt $(u - v) \in \operatorname{Kern} f$ und somit $u - v = 0$ bzw. $u = v$. Also ist f injektiv. □

Lemma 3.5: *Ist* U *ein Untervektorraum von* V *und* $f\colon V \to W$ *eine lineare Abbildung, so ist* $f(U)$ *ein Untervektorraum von* W. *Insbesondere ist das Bild von* f *ein Untervektorraum von* W.

Beweis: Seien $u, v \in U$ und $\alpha \in K$. Dann ist $f(u), f(v) \in f(U)$ und $\alpha f(u) + f(v) = f(\alpha u + v) \in f(U)$, also ist $f(U)$ ein Untervektorraum von W. Ganz analog folgt die zweite Behauptung. □

Satz über lineare Abbildungen des K^n in den K^m. *Jede lineare Abbildung* $f\colon K^n \to K^m$ *hat die Form*

$$f(x) = Fx$$

mit $F = \left(f(e^1), f(e^2), \ldots, f(e^n)\right) \in K^{m \times n}$.

Bemerkung: Die Spalten von F sind also die Bilder der natürlichen Basis.
Beweis: Sei $x = \sum_{i=1}^{n} \lambda_i e^i \in K^n$. Dann ist $f(x) = \sum_{i=1}^{n} \lambda_i f(e^i) = Fx$ (vgl. *Hauptformel der Matrizenmultiplikation*). □

Definition. Sei im folgenden $f\colon V \to W$ eine lineare Abbildung.

Bezeichnung von f	Bedingung	Nachweis
Homomorphismus	f ist linear	$f(\alpha u + v) = \alpha f(u) + f(v)$
Endomorphismus	$V = W$	$u \in V \Longleftrightarrow u \in W$
Monomorphismus	f ist injektiv	$\operatorname{Kern} f = \{0\}$
Epimorphismus	f ist surjektiv	$\operatorname{Bild} f = W$
Isomorphismus	f ist bijektiv	$\operatorname{Kern} f = \{0\}$ und $\operatorname{Bild} f = W$
Automorphismus	f ist bijektiver Endomorphismus	$\operatorname{Kern} f = \{0\}$ und $\operatorname{Bild} f = V$

Definition. Seien V und W zwei K-Vektorräume, f, g zwei Homomorphismen und $\alpha \in K$. Wir definieren dann die Addition und Skalarmultiplikation für lineare Abbildungen wie folgt:

$$(f + g)(v) = f(v) + g(v),$$
$$(\alpha f)(v) = \alpha f(v).$$

Damit ist die Menge aller Homomorphismen von V nach W ein K-Vektorraum, den wir mit

$$\operatorname{Hom}(V, W) := \left\{ f\colon V \to W \;\middle|\; f \text{ linear} \right\}$$

bezeichnen. Damit ist auch die Menge aller Endomorphismen über V, bezeichnet durch

$$\operatorname{End} V := \operatorname{Hom}(V, V)$$

ein K-Vektorraum.

Definition. Im folgenden sei f ein Endomorphismus.

- Die Abbildung f heißt *nilpotent*, falls es ein $k \in \mathbf{N}$ gibt, mit
$$f^k = 0.$$
- Gilt
$$f^2 = f,$$
so nennen wir f *idempotent* oder auch *Projektion*.
- Ferner heißt f *Involution*, falls gilt:
$$f^2 = \mathrm{Id}.$$

Beispiele für lineare Abbildungen: Im folgenden seien $u, v \in V$ und $\alpha \in K$.

1. Die *Nullabbildung*: $0\colon V \to W$, $v \mapsto 0_W$, denn $0(u + v) = 0_W = 0_W + 0_W = 0(u) + 0(v)$ und $\alpha \cdot 0(v) = \alpha \cdot 0_W = 0_W = 0(\alpha v)$.
2. Die *Identität*: $\mathrm{Id}\colon V \to V$, $v \mapsto v$, denn $\mathrm{Id}(u + v) = u + v = \mathrm{Id}(u) + \mathrm{Id}(v)$ und $\mathrm{Id}(\alpha v) = \alpha v = \alpha\,\mathrm{Id}(v)$.

3. Sei p eine Projektion mit

$$p\colon K^n \to K^n, \quad \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto \begin{pmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Seien $u, v \in K^n$ und $\lambda \in K$. Dann ist aufgrund der Rechenregeln für Vektoren $p(u+v) = p(u) + p(v)$ und $p(\lambda u) = \lambda p(u)$.

4. Sei $A \in K^{m\times n}$. Dann ist $f_A\colon K^n \to K^m, x \mapsto Ax$ eine lineare Abbildung, denn aufgrund der Rechenregeln für Matrizen gilt bei $x, y, \in K^n$ und $\alpha, \beta \in K$, daß $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$ (vgl. auch mit dem Satz über lineare Abbildungen des K^n in den K^m).

5. Die *Drehung* der Ebene um den Nullpunkt mit dem Winkel φ ist linear. Dies folgt, da das Bild einer Linearkombination gleich der Linearkombination der Bilder ist. Also gilt

$$d_\varphi\colon \mathbf{R}^2 \to \mathbf{R}^2, \quad d_\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = D(\varphi) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

mit der *Drehmatrix*

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

6. Die *Koordinatenabbildung* ist linear. Sei $B = (v_1, \ldots, v_n)$ eine Basis von V. Dann hat jeder Vektor $v \in V$ eine eindeutige Darstellung $v = \sum_{i=1}^n \alpha_i v_i$. Wir nennen dann

$$_B[v] := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Koordinatenvektor von v bezüglich B. Die Abbildung

$$V \to K^n, \quad v \mapsto {}_B[v]$$

heißt *Koordinatenabbildung*; sie ist linear und bijektiv, also ein Isomorphismus.

◊

3.1.3 Lineare Fortsetzung, Konstruktionssatz

Seien V und W zwei Vektorräume, $B = (v_1, \ldots, v_n)$ eine Basis von V, $f\colon V \to W$ eine lineare Abbildung und $v = \sum_{i=1}^n \alpha_i v_i$ ein Vektor aus V. Dann ist

$$f(v) = \sum_{i=1}^n \alpha_i f(v_i).$$

Also ist die Abbildung f durch die Bilder der Basisvektoren eindeutig bestimmt.

Konstruktionssatz. *Seien V und W zwei Vektorräume, $(w_1, \ldots, w_n)$ ein beliebiges n-Tupel mit $w_i \in W$. Dann wird durch*

$$g(v) := \sum_{i=1}^n \alpha_i w_i \quad \textit{für } v = \sum \alpha_i v_i$$

eine lineare Abbildung $g\colon V \to W$ definiert, die wir lineare Fortsetzung *von $v_i \mapsto w_i$ nennen.*

Konstruktion einer Basis für $\mathrm{Hom}(V,W)$. *Seien V und W zwei Vektorräume mit Basen $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$. Dann definieren wir eine lineare Fortsetzung von V nach W durch die linearen Abbildungen $e_{ij}\colon V \to W$ für $1 \le i \le n$ und $1 \le j, k \le m$ mit*

$$e_{ij}(v_k) := \begin{cases} w_i & \textit{für } k = j, \\ 0_W & \textit{sonst.} \end{cases}$$

Explizit gilt dann für $v = \sum_{k=1}^n \alpha_k v_k$

$$e_{ij}(v) = e_{ij}\Big(\sum_{k=1}^n \alpha_k v_k\Big) = \sum_{k=1}^n \alpha_k e_{ij}(v_k) = \alpha_j w_i.$$

Damit bildet $(e_{11}, e_{12}, \ldots, e_{1m}, \ldots, e_{nm})$ eine Basis von $\mathrm{Hom}(V,W)$. Insbesondere gilt

$$\mathrm{Dim}\big(\mathrm{Hom}(V,W)\big) = (\mathrm{Dim}\, V)(\mathrm{Dim} W) = m \cdot n.$$

Beweis: Der Beweis sei dem Leser als Übung empfohlen. □

3.1.4 Der Rang einer Abbildung und Rangformeln

Definition. Sei V ein endlichdimensionaler Vektorraum und $f\colon V \to W$ eine lineare Abbildung. Dann heißt

$$\mathrm{Rang}\, f := \mathrm{Dim}(\mathrm{Bild}\, f)$$

der *Rang* der linearen Abbildung f.

Rangsatz. *Sei V wieder endlichdimensional und $f\colon V \to W$ linear. Dann gilt:*

$$\mathrm{Dim}\, V = \mathrm{Dim}(\mathrm{Kern}\, f) + \mathrm{Rang}\, f,$$
$$\mathrm{Rang}\, f = \mathrm{Dim}\, V - \mathrm{Dim}(\mathrm{Kern}\, f).$$

Beweis: Sei $(v_1, \ldots, v_k)$ eine Basis vom $\mathrm{Kern}\, f$. Dann ergänzen wir diese Basis zu einer Basis $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ von V. Nun bleibt noch zu zeigen, daß die Familie von Vektoren $\big(f(v_{k+1}), \ldots, f(v_n)\big)$ eine Basis von $f(V)$ darstellt:

1. Sie stellen ein Erzeugendensystem dar:

$$f\Big(\sum_{i=1}^n \alpha_i v_i\Big) = \sum_{i=1}^n \alpha_i f(v_i) = \sum_{i=k+1}^n \alpha_i f(v_i),$$

da $f(v_i) = 0$ für $1 \le i \le k$.

2. Sie sind linear unabhängig:

$$\begin{aligned}
&\alpha_{k+1} f(v_{k+1}) + \cdots + \alpha_n f(v_n) = 0 \\
&\quad \Longrightarrow f(\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n) = 0, \\
&\quad \Longrightarrow \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n \in \operatorname{Kern} f, \\
&\quad \Longrightarrow \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n = \alpha_1 v_1 + \cdots + \alpha_k v_k, \\
&\quad \Longrightarrow \alpha_1 = \cdots = \alpha_k = \alpha_{k+1} = \cdots = \alpha_n = 0.
\end{aligned}$$

Damit gilt $\operatorname{Rang} f = \operatorname{Dim} f(U) = n - k$. □

Verallgemeinerung des Rangsatzes. *Seien U, V, W endlichdimensional und f, g lineare Abbildungen mit $U \xrightarrow{g} V \xrightarrow{f} W$. Dann gilt: $f \circ g\colon U \to W$, $(f \circ g)(u) = f\big(g(u)\big)$ ist linear und es ist*

$$\operatorname{Rang}(f \circ g) = \operatorname{Rang} g - \operatorname{Dim}(\operatorname{Kern} f \cap \operatorname{Bild} g).$$

Beweis: Wende Rangsatz auf die folgende Einschränkung an:

$$\begin{aligned}
&f|_{g|U|}\colon\ g(U) \to W, \\
&f|_{g|U|}\big(g(u)\big) = f\big(g(u)\big) = (f \circ g)(u).
\end{aligned}$$

□

Folgerungen:

1. Für den Fall $f\colon V \to W$ linear und $\operatorname{Dim} V = \operatorname{Dim} W < \infty$ sind folgende Aussagen äquivalent:

   (a) f ist injektiv.

   (b) f ist surjektiv.

   (c) f ist bijektiv.

   Beweis:

$$\begin{aligned}
f \text{ injektiv} &\Longleftrightarrow \operatorname{Kern} f = \{0\} \Longleftrightarrow \operatorname{Dim}(\operatorname{Kern} f) = 0 \\
&\Longleftrightarrow \operatorname{Dim} W = \operatorname{Dim} V = \operatorname{Rang} f = \operatorname{Dim}(\operatorname{Bild} f) \\
&\Longleftrightarrow W = \operatorname{Bild} f.
\end{aligned}$$

   □

   Dies kann folgendermaßen angewandt werden: Ist $f\colon V \to W$ linear und $\operatorname{Dim} V = \operatorname{Dim} W$ so ist f bijektiv, falls aus $f(x) = 0$ folgt, daß $x = 0$ gilt.

2. (a) Ist g surjektiv, so gilt $\operatorname{Rang}(f \circ g) = \operatorname{Rang} f$.

   (b) Ist f injektiv, so gilt $\operatorname{Rang}(f \circ g) = \operatorname{Rang} g$.

   (c) Sind g, h bijektiv, so gilt $\operatorname{Rang}(h \circ f \circ g) = \operatorname{Rang} f$.

3.1.5 Der Rang einer Matrix

Definition. Sei $A \in K^{m \times n}$ gegeben, dann definieren wir eine Abbildung $f_A\colon K^n \to K^m$ mit $f_A(x) = Ax$. Damit können wir den Kern, das Bild und den Rang der Matrix A wiefolgt definieren:

$$\operatorname{Kern} A := \operatorname{Kern} f_A, \qquad \operatorname{Bild} A := \operatorname{Bild} f_A, \qquad \operatorname{Rang} A := \operatorname{Rang} f_A.$$

Rangsatz für Matrizen. *Sei $A \in K^{m \times n}$. Dann gilt*

$$n = \operatorname{Rang} A + \operatorname{Dim}(\operatorname{Kern} A)$$

oder auch

$$\operatorname{Dim} \text{„}\textit{Lösungsraum}\text{“} = n - \operatorname{Rang} A.$$

(Mit Lösungsraum *ist der Lösungsraum der Gleichung $Ax = 0$ gemeint.) Diese Gleichung ist auch aus den expliziten Lösungsverfahren ablesbar.*

Wir wollen uns nun mit der Rangbestimmung befassen. Sei dazu wieder $A \in K^{m \times n}$. Dann folgt aus Folgerung 2c die Gleichung

$$\operatorname{Rang}(PAQ) = \operatorname{Rang} A \quad \text{für alle } P \in \mathrm{GL}(m, K) \text{ und } Q \in \mathrm{GL}(n, K).$$

(Die Abbildungen f_P und f_Q sind genau dann bijektiv, wenn P und Q invertierbare Matrizen sind.) Das heißt, der Rang ändert sich bei elementaren Zeilen- und Spaltenumformungen nicht. Damit kann der Rang einer Matrix A einfach bestimmt werden: Wir formen die Matrix durch elementare Zeilen- und Spaltenumformungen in eine Matrix $\widetilde{A}$ um, bis wir den Rang von $\widetilde{A}$ erkennen kann. Dann gilt:

$$\operatorname{Rang} A = \operatorname{Rang} \widetilde{A}.$$

Nun wollen wir den Fall für invertierbare Matrizen näher beleuchten.

Satz. *Für $A \in K^{n \times n}$ und $f_A(x) = Ax$ sind folgende Aussagen äquivalent:*

(1) $A \in \mathrm{GL}(n, K)$.

(2) Es gibt ein $A^{-1} \in K^{n \times n}$ derart, daß $AA^{-1} = A^{-1}A = E$.

(3) Die Zeilen von A sind linear unabhängig.

(4) Die Spalten von A sind linear unabhängig.

(5) $\operatorname{Rang} A = n$.

(6) Aus $Ax = 0$ folgt $x = 0$.

Beweis:

(1) $\Longleftrightarrow$ *(2)*: Folgt sofort aus der Definition von $\mathrm{GL}(n, K)$.

(1) $\Longleftrightarrow$ *(3)*: Ist $A \in \mathrm{GL}(n, K)$, so ist auch $A^T \in \mathrm{GL}(n, K)$. Damit gilt *(6)* auch für A^T und somit auch *(4)* für die Transponierte, d.h. *(3)* ist auch zu *(1)* äquivalent.

(*1*)$\Longleftrightarrow$(*6*): Ergibt sich aus Folgerung 1 in 3.1.4.

(*4*)$\Longleftrightarrow$(*6*): Sei $A = (a^1, \ldots, a^n)$. Sind die Spalten linear unabhängig, so gilt $\sum_{i=1}^n \xi_i a^i = 0$ genau dann, wenn alle $\xi_i = 0$. Fassen wir die ξ_i als die Komponenten von x auf, so folgt nach der Hauptformel aus $Ax = 0$ auch $x = 0$. Ebenso folgt aus (*6*) mittels der Hauptformel, daß die Spalten von A linear unabhängig sind.

(*5*)$\Longleftrightarrow$(*6*): Dies ergibt sich sofort aufgrund von Abschnitt 3.1.5, denn $\operatorname{Dim}\operatorname{Kern} f_A = 0$.

□

3.2 Darstellung von linearen Abbildungen durch Matrizen

3.2.1 Darstellungssatz

Satz. *Die Koordinatenabbildung* $f: v \mapsto {}_B[v]$ *für ein Element* v *aus einem Vektorraum* V *mit Basis* $B = (v_1, \ldots, v_n)$ *ist ein Isomorphismus von* V *auf den* K^n.

Beweis: Die Linearität haben wir bereits in Beispiel 6 von Kapitel 3.1.2 gesehen.
Da $v = \sum_{i=1}^n \alpha_i v_i \in V$ dann und nur dann Null ist, wenn alle $\alpha_i = 0$ $(1 \le i \le n)$, gilt $\operatorname{Kern} f = \{0\}$.
Sei $x \in K^n$ beliebig mit den Komponenten β_i $(1 \le i \le n)$. Dann ist $\sum_{j=1}^n \beta_j v_j \in V$, d.h. $\operatorname{Bild} f = K^n$. Damit ist die Koordinatenabbildung ein Isomorphismus. □

Definition. Wir bezeichnen

$$ {}_C[f]_B := \Big({}_C\big[f(v_1)\big], \ldots, {}_C\big[f(v_n)\big] \Big) \in K^{m\times n} $$

als *Darstellungs-* oder *Abbildungsmatrix* von f bezüglich B und C.

Darstellungssatz. *Seien* V *und* W *Vektorräume mit den Basen* B *und* C. *Sei ferner* $f: V \to W$ *linear. Dann gilt*

$$ {}_C\big[f(v)\big] = {}_C[f]_B \, {}_B[v] \quad \textit{für alle } v \in V. $$

Beweis: Es sei $B = (v_1, \ldots, v_n)$. Jeder Vektor $v \in V$ hat eine eindeutige Darstellung $v = \sum_{i=1}^n \alpha_i v_i$ und es gilt $f(v) = \sum_{i=1}^n \alpha_i f(v_i)$. Da die Koordinatenabbildung linear ist, erhalten wir unter Verwendung der Hauptformel der Matrizenmultiplikation

$$ {}_C\big[f(v)\big] = \sum_{i=1}^n \alpha_i \, {}_C\big[f(v_i)\big] = {}_C[f]_B \, {}_B[v]\,. $$

□

Bemerkung: Sei $A \in K^{m\times n}$ und $f_A: K^n \to K^m$ derart, daß $f_A(x) = Ax$. Dann gilt bezüglich der natürlichen Basen des K^n bzw. K^m:

$$ {}_{E_m}[f_A]_{E_n} = (Ae^1, \ldots, Ae^n) = (a^1, \ldots, a^n) = A. $$

Damit ist A die Darstellungsmatrix von f_A bezüglich der natürlichen Basen.

3.2.2 Basiswechsel

Definition. Sei V ein Vektorraum mit den Basen $B = (v_1, \ldots, v_n)$ und $C = (w_1, \ldots, w_n)$. Dann nennen wir die Darstellungsmatrix von $\mathrm{Id}: V \to V$ *Übergangsmatrix* von B nach C. Sie hat die Form

$$ {}_C[\mathrm{Id}]_B = \big({}_C[v_1], \ldots, {}_C[v_n] \big) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix} $$

mit $v_j = \sum_{i=1}^n \alpha_{ij} w_i$ für $1 \le j \le n$.

Satz vom Basiswechsel. *Seien* $f: V \to W$ *ein Isomorphismus,* A, B *zwei Basen von* V *und* C, D *zwei Basen von* W. *Dann gilt*

$$ {}_D[f]_A = T \, {}_C[f]_B \, S $$

mit den Übergangsmatrizen $S = {}_B[\mathrm{Id}]_A$ *und* $T = {}_D[\mathrm{Id}]_C$.

Beweis: Die Multiplikationsformel zeigt

$$ {}_D[f]_A = {}_D[\mathrm{Id}]_C \, {}_C[f]_B \, {}_B[\mathrm{Id}]_A. $$

□

Sei $f: V \to V$ ein Endomorphismus und V ein Vektorraum mit zwei Basen A und B, dann gilt

$$ {}_B[f]_B = {}_B[\mathrm{Id}]_A \, {}_A[f]_A \, {}_A[\mathrm{Id}]_B = S^{-1} \, {}_A[f]_A S \quad \text{mit } S = {}_A[\mathrm{Id}]_B. $$

Definition. Zwei Matrizen $A, B \in K^{n\times n}$ heißen *ähnlich*, wenn es eine Matrix $C \in \mathrm{GL}(n, K)$ mit $B = C^{-1}AC$ gibt.

Bemerkung: Die Ähnlichkeit von Matrizen ist eine Äquivalenzrelation.
Beweis: Seien A und B ähnliche Matrizen, ebenso B und C. Dann ist die Ähnlichkeit reflexiv, da $A = E^{-1}AE$ gilt. Aufgrund von $A = S^{-1}BS \Longrightarrow B = TAT^{-1}$ mit $T = S$ ist sie symmetrisch und schließlich wegen $A = S^{-1}BS, B = U^{-1}CU \Longrightarrow A = W^{-1}CW$ mit $W = US$ transitiv.

□

Wir stellen fest, daß die Darstellungsmatrizen von f bezüglich A und B ähnlich sind.

Satz. *Sei $A \in K^{n\times n}$ und sei $f_A\colon K^n \to K^n$, $x \mapsto Ax$, eine lineare Abbildung. Ferner sei $B := (b^1, \ldots, b^n)$ eine Basis des K^n. Dann gilt*

$$ {}_B[f_A]_B = B^{-1}AB. $$

Beweis: Aufgrund des Darstellungssatzes erhalten wir

$$ {}_B[Ab^i] = {}_B[\mathrm{Id}]_E \, {}_E[Ab^i] = B^{-1}Ab^i \qquad (1 \le i \le n). $$

Damit ist

$$ {}_B[f_A]_B = \left({}_B[Ab^1], \ldots, {}_B[Ab^n] \right) = \left(B^{-1}Ab^1, \ldots, B^{-1}Ab^n \right) = B^{-1}AB. $$

□

Beispiele:

1. Seien $B := (b^1, \ldots, b^n)$ und $C := (c^1, \ldots, c^n)$ Basen des K^n. Dann ist

   $$ {}_B[\mathrm{Id}]_B = \left({}_B[b^1], \ldots, {}_B[b^n] \right) = (e^1, \ldots, e^n) = E $$

   und ferner ist

   $$ {}_C[\mathrm{Id}]_B = \left({}_C[b^1], \ldots, {}_C[b^n] \right) \in GL(n, K). $$

2. Sei der Vektorraum V gegeben durch

   $$ V := \{\, p = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 \mid \alpha_i \in \mathbf{R} \quad (i = 1, 2, 3) \,\}. $$

   Ferner sei $d\colon V \to V$, $p \mapsto p'$ gegeben, sowie die Basen $B := (1, x, x^2, x^3)$ und $C := \left(1, x-1, (x-1)^2, (x-1)^3\right)$ von V. Dann ist

   $$ {}_B[d]_B = \left({}_B\left[d(1)\right], {}_B\left[d(x)\right], {}_B\left[d(x^2)\right], {}_B\left[d(x^3)\right] \right) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} $$

   und

   $$ {}_C[\mathrm{Id}]_B = \left({}_C\left[\mathrm{Id}(1)\right], {}_C\left[\mathrm{Id}(x)\right], {}_C\left[\mathrm{Id}(x^2)\right], {}_C\left[\mathrm{Id}(x^3)\right] \right) = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}. $$

   Außerdem ist

   $$ {}_B[d]_C = \left({}_B\left[d(1)\right], {}_B\left[d(x-1)\right], {}_B\left[d\left((x-1)^2\right)\right], {}_B\left[d\left((x-1)^3\right)\right] \right) = \begin{pmatrix} 0 & 1 & -2 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}. $$

3. Offenbar sind

   $$ A = \begin{pmatrix} 4 & 5 & 1 \\ 1 & 7 & 2 \\ 1 & 5 & 4 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & -5 & 6 \\ 1 & 3 & -1 \\ 1 & -5 & -1 \end{pmatrix} \text{ zwei Basen des } K^3. $$

   Damit ist

   $$ {}_B[f_A]_B = \left({}_B[Ab^1], {}_B[Ab^2], {}_B[Ab^3] \right) = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = B^{-1}AB =: C $$

   (siehe Satz oben). Daraus erhalten wir $A = BCB^{-1}$.

   Wollen wir etwa A^k berechnen, so können wir die Beziehung $A^k = BC^kB^{-1}$ benutzen.

◊

3.2.3 Invariante Unterräume

Definition. Sei f ein Endomorphismus über dem Vektorraum V. Ein Unterraum $U \subseteq V$ heißt *f-invariant*, falls $f(U) \subseteq U$ ist.

Ist $U \subseteq V$ ein f-invarianter Unterraum, so können wir eine Basis $B' = (u_1, \ldots, u_k)$ von U zu einer Basis $B = (u_1, \ldots, u_n)$ von V ergänzen und erhalten folgende Darstellung von f:

$$ {}_B[f]_B = \begin{pmatrix} \boxed{{}_{B'}[f]_{B'}} & \boxed{*} \\ \boxed{0} & \boxed{*} \end{pmatrix}. $$

Betrachten wir nun den Spezialfall $U = Kv$ mit $v \in V \setminus \{0\}$. Sei $B = (v_1, \ldots, v_n)$ ein Basis von V. Der Untervektorraum U ist genau dann f-invariant, wenn es ein $\lambda \in K$ mit der Eigenschaft $f(v) = \lambda v$ gibt. Dazu wählen wir $v = v_1$. Es gilt dann

$$ {}_B[f]_B = \begin{pmatrix} \lambda & * & \ldots & * \\ 0 & & & \\ \vdots & & \boxed{*} & \\ 0 & & & \end{pmatrix}. $$

Definition. Sei nun $U \subseteq V$ ein f-invarianter Untervektorraum. Man nennt $U' \subseteq V$ ein *f-invariantes Komplement*, wenn gilt:

(1) U' ist f-invariant.

(2) $V = U \oplus U'$.

Das heißt, sind $B' = (u_1, \ldots, u_k)$ und $B'' = (u_{k+1}, \ldots, u_n)$ Basen von U und U', so ist $B = (u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ eine Basis von V und es gilt:

$$ {}_B[f]_B = \begin{pmatrix} \boxed{{}_{B'}[f|_U]_{B'}} & \boxed{0} \\ \boxed{0} & \boxed{{}_{B''}[f|_{U'}]_{B''}} \end{pmatrix}. $$

3.3 Der Dualraum und duale Abbildungen

Im folgenden sei V ein K-Vektorraum.

Definition. Eine lineare Abbildung $f\colon V \to K$ in den eindimensionalen K-Vektorraum K bezeichnen wir als *Linearform*. (Eine Abbildung in den Grundkörper wird meistens mit der Endung „-form“ gekennzeichnet.)

Die Matrix ${}_{(1)}[f]_B$ einer Linearform ist ein Zeilenvektor, der Form $\big(f(v_1), \dots, f(v_n)\big)$.

Beispiele für Linearformen:

1. Seien $B = (v_1, \dots, v_n)$ eine Basis und $v = \sum \alpha_i v_i$. Dann ist die Abbildung $f(v) := \alpha_1$ eine Linearform.
2. Sei $A = (\alpha_{ij})$ eine $n \times n$ Matrix. Man nennt
$$\operatorname{Sp} A := \alpha_{11} + \alpha_{22} + \cdots + \alpha_{nn}$$
die *Spur* von A. Diese Abbildung $\operatorname{Sp}\colon K^{n\times n} \to K$, $A \mapsto \operatorname{Sp} A$ ist eine Linearform.

$\Diamond$

Definition. Der Vektorraum V^* aller Linearformen eines Vektorraumes V ist definiert durch
$$V^* := \operatorname{Hom}(V, K)$$
und wird *Dualraum von* V genannt.

Satz von der dualen Basis. *Seien* $B = (v_1, \dots, v_n)$ *eine Basis von* V *und* $\lambda_i \in V^*$ *für alle* $1 \le i \le n$ *definiert durch*
$$\lambda_i(v) := \alpha_i \quad \textit{mit } v = \sum_{k=1}^{n} \alpha_k v_k.$$
Dann ist $(\lambda_1, \dots, \lambda_n)$ *eine Basis von* V^*. *Sie wird auch die zu* B *duale Basis genannt.*

Beweis: Nach Voraussetzung ist $v = \sum \alpha_k v_k = \sum \lambda_k(v) v_k$. Außerdem ist $\operatorname{Dim} V^* = \operatorname{Dim} V$. Sei $\beta_1\lambda_1 + \cdots + \beta_n\lambda_n = 0$, also $\sum \beta_k \lambda_k(v) = 0$ für alle v. Insbesondere folgt damit für $v = v_i$, daß $\beta_i = 0$ ist! Daraus folgt die Behauptung. □

Definition. Zu einem Endomorphismus f über dem Vektorraum V definieren wir mit $\lambda \in V^*$ die *duale Abbildung* f^*:
$$f^*(\lambda) = \lambda \circ f.$$

Satz. *Die duale Abbildung* f^* *ist ein Endomorphismus über dem Vektorraum* V^*.

Beweis: Die Behauptung folgt mit
$$f^*(\lambda+\mu)(v) = (\lambda+\mu)\big(f(v)\big) = \lambda\big(f(v)\big) + \mu\big(f(v)\big) = (f^*\lambda)(v) + (f^*\mu)(v).$$
□

Lemma 3.6: *Zu jedem Vektorraum* V *gibt es einen Isomorphismus*
$$\psi\colon V \to (V^*)^*,$$
d.h. wir können V *mit* $(V^*)^*$ *identifizieren.*

Beweis: Sei $u \in V$ fest und $\varphi_u\colon V^* \to K, \lambda \mapsto \lambda(u)$. Offenbar ist φ_u linear und damit $\varphi_u \in (V^*)^*$.
Betrachten wir nun die lineare Abbildung $\psi\colon V \to (V^*)^*, u \mapsto \varphi_u$: Sie ist injektiv, denn aus $\varphi_u = 0$ folgt $\lambda(u) = 0$ für alle $\lambda \in V^*$ und somit ist auch $u = 0$. Da zusätzlich $\operatorname{Dim} V = \operatorname{Dim}(V^*)^*$ ist das sogar eine Isometrie. Über diesen Isomorphismus ψ wird V mit $(V^*)^*$ identifiziert. □

Definition. Nun definieren wir das *Kronecker Symbol* δ_{ij} durch
$$\delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Satz. *Sei* V *ein* K*-Vektorraum mit dazugehörigem Dualraum* V^*, U *ein Unterraum von* V *und* $U^\perp := \big\{ \lambda \in V^* \,\big|\, \lambda(u) = 0, u \in U \big\}$. *Dann gilt*
$$\operatorname{Dim} V = \operatorname{Dim} U + \operatorname{Dim} U^\perp.$$

Beweis: Sei $(u_1, \dots, u_r)$ eine Basis von U. Dann erhalten wir durch Basisergänzung
$$(u_1, \dots, u_r, u_{r+1}, \dots, u_n)$$
eine Basis von V. Betrachten wir dazu die Dualbasis $(\lambda_1, \dots, \lambda_r, \lambda_{r+1}, \dots, \lambda_n)$, so ist mit $\lambda := \sum_{i=1}^{n} \alpha_i\lambda_i \in V^*$ nach Konstruktion $\lambda(u_j) = 0$ für $j \in \{1, \dots, r\}$. Damit ist $\sum_{i=1}^{n} \alpha_i\lambda_i(u_j) = 0$ mit $1 \le j \le r$ und $\sum_{i=1}^{r} \alpha_i\lambda_i(u_j) = \sum_{i=1}^{r} \alpha_i\delta_{ij} = \alpha_j$. Also ist $\lambda = \sum_{i=r+1}^{n} \alpha_i\lambda_i$, d.h. $U^\perp = \operatorname{Lin}(\lambda_{r+1}, \dots, \lambda_n)$. Daraus folgt unmittelbar die Behauptung. □

Satz. *Sei* V *wieder ein* K*-Vektorraum und* f *ein Endomorphismus über* V. *Dann gilt:*

(1) $\operatorname{Kern} f^* = (\operatorname{Bild} f)^\perp$,

(2) $\operatorname{Kern} f = (\operatorname{Bild} f^*)^\perp$,

(3) $\operatorname{Rang} f^* = \operatorname{Rang} f$.

Beweis:

(1) Sei $\lambda \in \operatorname{Kern} f^*$. Genau dann ist $f^*(\lambda) = \lambda \circ f = 0$ und damit $\lambda \in (\operatorname{Bild} f)^\perp$.

(2) Erhalten wir sofort aus 1 und $(V^*)^* = V$, sowie $(f^*)^* = f$.

(3) Es gilt
$$\begin{aligned} \operatorname{Rang} f^* &= \operatorname{Dim} V^* = \operatorname{Dim}(\operatorname{Kern} f^*) = \operatorname{Dim} V - \operatorname{Dim}(\operatorname{Bild} f) \\ &= \operatorname{Dim} V - \big(\operatorname{Dim} V - \operatorname{Dim}(\operatorname{Bild} f)\big) = \operatorname{Rang} f. \end{aligned}$$
□

3.4 Ergänzungen

3.4.1 Isomorphie

Sei $f\colon V \to W$ ein *Isomorphismus*. Dann existiert $f^{-1}\colon W \to V$, wobei auch f^{-1} ein Isomorphismus ist. Sei weiter auch g ein Isomorphismus mit $U \xrightarrow{g} V \xrightarrow{f} W$. Dann ist auch $f \circ g$ ein Isomorphismus und es gilt $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

Beispiele für Isomorphismen:

1. Die Koordinatenabbildung $V \to K^n$, $v \mapsto {}_B[v]$, ist ein Isomorphismus falls $B = (v_1, \ldots, v_n)$ eine Basis ist.
2. Seien nun V ein Vektorraum mit Basis B und f ein Endomorphismus über V. Dann ist $f \mapsto {}_B[f]_B$ ein Isomorphismus von $\operatorname{End} V$ auf $K^{n\times n}$. Zusätzlich gilt mit $g \in \operatorname{End} V$:
$$f \circ g \mapsto {}_B[f \circ g]_B = {}_B[f]_B \, {}_B[g]_B,$$
d.h. die Abbildung ist ein *Ringisomorphismus*.

◊

Definition. Wir sagen V *ist isomorph zu* W, falls es einen Isomorphismus $f\colon V \to W$ gibt. In Zeichen:
$$V \cong W.$$

Es gilt:
$$V \cong W \implies W \cong V, \quad V \cong V \qquad \text{(Symmetrie), (Reflexivität),}$$
$$V \cong W, \quad W \cong U \implies V \cong U \qquad \text{(Transitivität).}$$

Somit handelt es sich um eine Äquivalenzrelation.

Satz. *Sei $f\colon V \to W$ ein Isomorphismus. Dann gilt:*

(1) Sind die Vektoren $(v_1, \ldots, v_k)$ linear unabhängig, so sind auch $\big(f(v_1), \ldots, f(v_k)\big)$ linear unabhängig.

(2) Ist $(v_1, \ldots, v_n)$ eine Basis von V, so ist $\big(f(v_1), \ldots, f(v_n)\big)$ eine Basis von W.

Beweis:

(1) Wir benötigen nur die Injektivität von f:
$$0_W = \sum \alpha_i f(v_i) = f\Big(\sum \alpha_i v_i\Big) \implies \sum \alpha_i v_i = 0,$$
Aufgrund der linearen Unabhängigkeit der v_i folgt $\alpha_1 = \cdots = \alpha_k = 0$.

(2) Die lineare Unabhängigkeit folgt nach (1). Da f surjektiv ist gibt es zu jedem $w \in W$ ein v mit $f(v) = w$. Aus $v = \sum \alpha_i v_i$ ergibt sich $w = \sum \alpha_i f(v_i)$. Damit bildet die Familie $\big(f(v_1), \ldots, f(v_n)\big)$ ein Erzeugendensystem und damit eine Basis.

□

Satz. *Seien V und W endlichdimensionale Vektorräume mit den Basen B, C und $f\colon V \to W$ eine lineare Abbildung. Dann gilt:*

(1) $\operatorname{Dim} V = n \iff V \cong K^n$.

(2) $\operatorname{Dim} V = \operatorname{Dim} W \iff V \cong W$.

(3) $\operatorname{Rang} f = \operatorname{Rang} {}_C[f]_B$.

Beweis: *Die Beweise zu (1) und (2) seien dem Leser als Übung empfohlen.*
Es gilt (3), denn
$$\begin{aligned}\operatorname{Rang} f &= \operatorname{Dim} \operatorname{Bild} f = \operatorname{Dim} \operatorname{Lin}\big(f(v_1), \ldots, f(v_n)\big)\\ &= \operatorname{Dim} \operatorname{Lin}\Big({}_C[f(v_1)], \ldots, {}_C[f(v_n)]\Big) = \operatorname{Rang} {}_C[f]_B.\end{aligned}$$

□

3.4.2 Faktorräume, Quotientenräume

Seien im folgenden V ein K-Vektorraum und $U \subseteq V$ ein Unterraum.

Definition. Die Menge
$$v + U := \{\, v + u \mid u \in U \,\} \quad \text{für ein } v \in V$$
heißt *Nebenklasse modulo* U. Die *Menge aller Nebenklassen* wird durch
$$V/U := \{\, v + U \mid v \in V \,\}$$
bezeichnet.

Satz. *Es gilt:*

(1) $v + U = w + U \iff v - w \in U$.

(2) $v + U = U \iff v \in U$.

Beweis:

(1) Es gilt $v \in v + U = w + U$. Also gibt es ein $u \in U$ mit $v = w + u$ bzw. $v - w = u \in U$.
Es gibt ein $u' \in U$ mit $v - w = u'$ bzw. $v = w + u'$. Daraus folgt:
$$v + U = \{\, v + u \mid u \in U \,\} = \{\, w + (u' + u) \mid u \in U \,\} = w + U.$$

(2) Dies ist ein Sonderfall von (1) mit $w = 0$.

□

Definition. Wir definieren eine Addition und eine Skalarmultiplikation wie folgt:

$$(v + U) + (w + U) := (v + w) + U,$$

$$\alpha(v + U) := \alpha v + U.$$

Den dadurch definierten K-Vektorraum V/U nennen wir *Faktorraum* V mod U bzw. *Faktorraum* V *nach* U oder auch *Quotientenraum*. Er enthält das Nullelement U.

Diese Definition ist wohldefiniert, denn aus $v+U = v'+U$ und $w+U = w'+U$ folgt $(v+w)+U = (v'+w')+U$ und $\alpha v + U = \alpha v' + U$.

Beachte: Die Abbildung $\pi\colon V \to V/U$, $\pi(v) := v + U$, ist ein Epimorphismus. Sie heißt *kanonischer* bzw. *natürlicher Epimorphismus*.
Sei nun V endlichdimensional. Dann gilt für die Abbildung π:

$$\operatorname{Kern}\pi = U,$$

$$\operatorname{Dim}(V/U) = \operatorname{Dim} V - \operatorname{Dim} U.$$

Bemerkung: Sei $(u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ eine Basis von V derart, daß $(u_1, \ldots, u_k)$ eine Basis von U ist. Dann ist die Familie $(u_{k+1} + U, \ldots, u_n + U)$ eine Basis von V/U. (Vergleiche mit Beweis des Rangsatzes; hier für $\pi\colon V \to V/U$.)

3.4.3 Der Homomorphiesatz

Homomorphiesatz. *Sei* $f\colon V \to W$ *eine lineare Abbildung. Dann gilt:*

$$V/\operatorname{Kern} f \cong \operatorname{Bild} f.$$

Beweis: Wir müssen zeigen, daß es einen Isomorphismus

$$\Phi\colon V/\operatorname{Kern} f \to \operatorname{Bild} f, \quad \Phi(v + \operatorname{Kern} f) := f(v)$$

gibt.
Offenbar ist Φ wohldefiniert, d.h. für $v, v' \in V$ folgt aus $v + \operatorname{Kern} f = v' + \operatorname{Kern} f$, daß $f(v) = \Phi(v + \operatorname{Kern} f) = \Phi(v' + \operatorname{Kern} f) = f(v')$. Denn es ist $v - v' \in \operatorname{Kern} f$ und damit $f(v) = f(v')$.
Jetzt zeigen wir die Linearität von Φ: Seien dazu $v, v' \in V$. Es folgt dann einerseits

$$\begin{aligned}\varphi\big((v + \operatorname{Kern} f) + (v' + \operatorname{Kern} f)\big) &= \varphi(v + v' + \operatorname{Kern} f)\\ &= f(v + v')\\ &= \varphi(v + \operatorname{Kern} f) + \varphi(v' + \operatorname{Kern} f)\end{aligned}$$

und andererseits

$$\begin{aligned}\varphi\big(\alpha(v + \operatorname{Kern} f)\big) &= \varphi(\alpha v + \operatorname{Kern} f)\\ &= f(\alpha v)\\ &= \alpha \cdot f(v)\\ &= \alpha \cdot \varphi(v + \operatorname{Kern} f).\end{aligned}$$

Als letztes beweisen wir, daß Φ bijektiv ist: Die Surjektivität von Φ folgt sofort aus der Definition von Φ. Außerdem ist Φ injektiv, denn wären $v + \operatorname{Kern} f$ und $v' + \operatorname{Kern} f$ zwei verschiedene Nebenklassen mit demselben Bild, so würde $f(v) = f(v')$ sein. Also $v - v' \in \operatorname{Kern} f$, was ein Widerspruch zur Annahme ist. □

Wir können den Homomorphiesatz sehr wirkungsvoll bei einigen Beweisen einsetzen. Wie wir leicht sehen folgt etwa der Rangsatz sofort aus dem Homomorphiesatz.

Kapitel 4

Determinanten

4.1 Definitionen und elementare Eigenschaften

4.1.1 Rekursive Definition der Determinante

Definition. Sei $A = (\alpha) \in K^{1\times 1}$. Dann definieren wir

$$\det A := \alpha.$$

Für $A = (\alpha_{ij}) \in K^{n\times n}$ und $n > 1$ definieren wir

$$\det A := \alpha_{11} \det A_{11} - \alpha_{21} \det A_{21} + \cdots + (-1)^{n+1}\alpha_{n1} \det A_{n1},$$

wobei $A_{i1} \in K^{(n-1)\times(n-1)}$ aus A durch Entfernen der ersten Spalte und i-ten Zeile entsteht (*Entwicklung der Determinante nach der ersten Spalte*).
Dabei heißt $\det A$ *Determinante von* A.

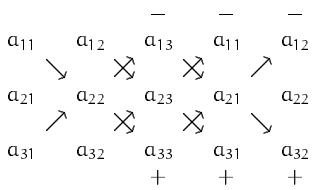
4.1.2 Determinantenberechnung

Regel von Sarrus:

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

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Diese Formel können wir uns leicht mit folgendem Schema merken:



Bemerkung: Die Regel von Sarrus gilt nur für Determinanten von Matrizen aus dem $K^{3\times 3}$.

4.1.3 Rechenregeln

Satz.

(*1*) *Die Determinante ist in jeder Zeile linear, d.h. sind* $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in K_n$ *fest, so ist die Abbildung*

$$x \mapsto \det \begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ x \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$$

linear.

(*2*) *Die Determinante ist bzgl. der Zeilen alternierend, d.h. entsteht* $\tilde{A}$ *aus* A *durch Vertauschung zweier Zeilen, dann gilt* $\det \tilde{A} = -\det A$.

Beweis: Die Behauptung ergibt sich aus der Definition leicht mit vollständiger Induktion. □

Bemerkung: Insbesondere gilt $\det A = 0$, falls 2 Zeilen gleich sind.

Satz. *Es gilt:*

$$\det \begin{pmatrix} \alpha_{11} & * & \ldots & * \\ 0 & \alpha_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \alpha_{nn} \end{pmatrix} = \alpha_{11}\alpha_{22}\ldots\alpha_{nn}.$$

Beweis mit vollständiger Induktion: Für $A = (\alpha) \in K^{1\times 1}$ gilt

$$\det A = \alpha \quad \text{(Induktionsanfang)}$$

und unter der Induktionsvoraussetzung, daß die Behauptung für $A \in K^{(n-1)\times(n-1)}$ gilt, erhalten wir für $A \in K^{n\times n}$

$$\det A = \alpha_{11} \cdot \det A_{11} = \alpha_{11}\alpha_{22}\ldots\alpha_{nn}.$$

□

Bemerkung: Damit erkennen wir auch sofort, daß

$$\det E = 1.$$

Änderung der Determinante bei elementaren Zeilen- bzw. Spaltenumformungen.

Typ 1. $\det\big(D_i(\alpha)A\big) = \alpha \det A = \det D_i(\alpha) \det A.$

Typ 2. $\det(P_{ij}A) = -\det A = \det P_{ij} \det A.$

Typ 3. $\det\big(T_{ij}(\alpha)A\big) = \det A = \det T_{ij} \det A.$

Beweis: Die linken Gleichungen ergeben sich wie folgt:

Typ 1. Der Faktor α der i-ten Zeile kann vorgezogen werden. (Die Determinante ist in der i-ten Zeile linear.)

Typ 2. Dies ist Rechenregel (*2*) aus obigem Satz.

Typ 3. Wird die i-te Zeile von A durch $a_i + a_j$ ersetzt, kann man die Determinante davon wegen der Linearität in 2 Summanden zerlegen. Der eine ist $\det A$ und der andere ist $\det A_1 = 0$, da in A_1 der Vektor a_j sowohl in der i-ten, als auch in der j-ten Zeile steht.

Mit $A = E$ und $\det E = 1$ erhalten wir insbesondere

$$\det D_i(\alpha) = \alpha, \qquad \det P_{ij} = -1, \qquad \det T_{ij}(\alpha) = 1$$

und damit dann die rechten Gleichungen. □

Determinantentest für Invertierbarkeit. *Sei $A = (\alpha_{ij})_{n\times n}$. Dann gilt*

$$A \in \mathrm{GL}(n, K) \iff \det A \neq 0.$$

Dazu ist folgende Aussage äquivalent:

$$\operatorname{Rang} A < n \iff \det A = 0.$$

Beweis: Eine Matrix A ist genau dann invertierbar, wenn sie mit elementaren Zeilenumformungen in eine Dreiecksform gebracht werden kann, in der alle Elemente der Hauptdiagonalen ungleich Null sind. Dies entspricht der Bedingung $\det A \neq 0$ (siehe oben Berechnung der Determinanten einer Matrix in Dreiecksform und beachte, daß sich die Determinante bei Zeilenumformungen nur um einen Faktor ungleich Null ändert). □

Multiplikationssatz. *Für $A, B \in K^{n\times n}$ gilt:*

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

Beweis:

1. Sei A nicht invertierbar. Dann folgt nach dem vorhergehenden Satz $\det A = 0$. Man sieht sofort, daß auch AB nicht invertierbar sein kann. Dies ergibt $\det(AB) = 0$.

2. Sei $A \in \mathrm{GL}(n, K)$ eine invertierbare Matrix. Dann ist A ein Produkt von Elementarmatrizen: $A = S_k \dots S_1$. Damit gilt für $\det(AB)$ nach den oben bewiesenen Formeln für Elementarmatrizen:

   $$\det(AB) = \det\big((S_k \dots S_1)B\big) = \det S_k \det\big((S_{k-1} \dots S_1)B\big) = \dots = \det S_k \dots \det S_1 \det B.$$

   Setzen wir nun $B = E$, so erhalten wir $\det A = \det S_k \dots \det S_1$.

□

Folgerungen:

1. $\det(AB) = \det(BA)$ (trivial nach Multiplikationssatz).

2. $\det(A^k) = (\det A)^k$, $k \in \mathbf{N}$ (Beweis durch Induktion und Multiplikationssatz).

3. $\det(A^{-1}) = (\det A)^{-1}$. Denn es gilt:

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

   Damit gilt 2. auch für negative Exponenten, falls $\det A \neq 0$.

4. $\det(C^{-1}AC) = \det A$ (Diese Beziehung ist für den Basiswechsel wichtig!)

Satz über die Symmetrie der Determinanten in Zeilen und Spalten. *Es gilt:*

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

Beweis:

1. Sei A nicht invertierbar. Dann ist auch A^T nicht invertierbar. Also sind beide Seiten der Gleichung Null.

2. Sei $A \in \mathrm{GL}(n, K)$ eine invertierbare Matrix. Dann ist $A = S_k \dots S_1$ ein Produkt von Elementarmatrizen. Weiter gilt $A^T = S_1^T \dots S_k^T$, sowie $\det S_i^T = \det S_i$ $(1 \le i \le k)$ und es folgt nach dem Multiplikationssatz

   $$\det A^T = \prod_{i=1}^{k} \det S_i^T = \prod_{i=1}^{k} \det S_i = \det A.$$

□

Kästchensatz. *Sei*

$$A = \begin{pmatrix} \boxed{A_1} & * & \dots & * \\ 0 & \boxed{A_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \boxed{A_r} \end{pmatrix} \in K^{n\times n}$$

mit $A_i \in K^{n_i\times n_i}$ $(1 \le i \le r)$ *und* $\sum n_i = n$. *Dann gilt*

$$\det A = \prod_{i=1}^{r} \det A_i.$$

Beweis: Beweis mit Entwicklung nach erster Spalte und vollständiger Induktion. Die benötigten A_i haben ebenfalls Kastenform. □

4.1.4 Der Entwicklungssatz nach Laplace

Entwicklungssatz nach Laplace. *Sei* $A \in K^{n\times n}$.

- *Dann ist einerseits*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det A_{ij} \quad (1 \le j \le n),$$

  wobei A_{ij} *aus* A *durch Entfernen der* i*-ten Zeile und* j*-ten Spalte entsteht (Entwicklung nach der* j*-ten Spalte).*

- *Andererseits ist*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det A_{ij} \quad (1 \le i \le n)$$

  (Entwicklung nach der i*-ten Zeile).*

Beweis: Die Matrix $\widetilde{A} = (a^j, a^1, a^2, \dots, a^{j-1}, a^{j+1}, \dots, a^n)$ $(1 \le j \le n)$ entsteht aus $A = (a^1, \dots, a^n)$ durch $j-1$ sukzessive Vertauschungen benachbarter Spalten. Also gilt $\det \widetilde{A} = (-1)^{j-1} \det A$. Entwickelt man andererseits erneut $\det \widetilde{A}$ nach der ersten Spalte, so ergibt sich die erste Behauptung.
Durch Verwendung von A^T erhalten wir das analoge Ergebnis für die Entwicklung nach der i-ten Zeile. □

4.2 Die vollständige Entwicklung der Determinante

4.2.1 Permutationen

Definition. Sei X eine nichtleere Menge. Dann heißt

$$S(X) := \{ f\colon X \to X \mid f \text{ bijektiv} \}$$

die *symmetrische Gruppe* von X und $f \in S(X)$ eine Permutation von X.
Ferner ist

$$S_n := S(\{1, 2, 3, \dots, n\}).$$

Die Permutationen aus S_n werden dargestellt in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

Definition. Es sei $\sigma \in S_n$ und $P(\sigma) = (e^{\sigma(1)}, \dots, e^{\sigma(n)})$. Dann nennen wir

$$\varepsilon(\sigma) := \det P(\sigma) = (-1)^k$$

die *Signatur* von σ. Dabei ist k die Anzahl der paarweisen Vertauschungen der Spalten von $P(\sigma)$, die erforderlich sind, um $P(\sigma)$ nach E_n umzuformen, das ist gleich der Anzahl der Vertauschungen mit denen $(\sigma(1), \dots, \sigma(n))$ in die natürliche Reihenfolge $(1, 2, \dots, n)$ gebracht wird.

4.2.2 Der Entwicklungssatz von Leibniz

Vollständiger Entwicklungssatz. *Ist* $A \in K^{n\times n}$ *und* $a^j = \sum_{k=1}^{n} \alpha_{kj} e^k$, *so gilt*

$$\det A = \sum_{\sigma\in S_n} \varepsilon(\sigma) \alpha_{\sigma(1)1} \dots \alpha_{\sigma(n)n}.$$

Beweis: Es gilt

$$\begin{aligned} \det A = \det\left(\sum_{i_1=1}^{n} \alpha_{i_1 1} e^{i_1}, a^2, \dots, a^n \right) &= \sum_{i_1=1}^{n} \alpha_{i_1 1} \det(e^{i_1}, a^2, \dots) \\ &= \sum_{i_1,i_2=1}^{n} \alpha_{i_1 1} \alpha_{i_2 2} \det(e^{i_1}, e^{i_2}, a^3, \dots, a^n) \\ &= \sum_{i_1,i_2,\dots,i_n=1}^{n} \alpha_{i_1 1} \alpha_{i_2 2} \dots \alpha_{i_n n} \det(e^{i_1}, e^{i_2}, \dots, e^{i_n}) \\ &= \sum_{\sigma\in S_n} \varepsilon(\sigma) \alpha_{\sigma(1)1} \dots \alpha_{\sigma(n)n}, \end{aligned}$$

da $\det(e^{i_1}, \dots, e^{i_n})$ nur dann ungleich Null ist, wenn die Indizes paarweise verschieden sind, also zu einer Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

gehören. □

4.3 Anwendungen von Determinanten

1. *Die Cramer-Regel.* Sei $A = (a^1, \ldots, a^n) \in \mathrm{GL}(n, K)$ und $b \in K^n$. Dann erhalten wir für die Komponenten x_i, $1 \le i \le n$, des Lösungsvektors x von $Ax = b$

$$x_i = \frac{\det(a^1, \ldots, a^{i-1}, b, a^{i+1}, \ldots, a^n)}{\det A}.$$

Beweis: Es gilt nach der Hauptformel $b = \sum_j x_j a^j$. Damit ist

$$\begin{aligned}\det(a^1, \ldots, a^{i-1}, b, a^{i+1}, \ldots, a^n) &= \det\left(a^1, \ldots, a^{i-1}, \sum_j x_j a^j, a^{i+1}, \ldots, a^n\right)\\ &= \sum_j x_j \det(a^1, \ldots, a^{i-1}, a^j, a^{i+1}, \ldots, a^j)\\ &= x_i \det A,\end{aligned}$$

da für $i \neq j$ die Determinantenfaktoren Null sind (2 gleiche Spalten). □

2. *Eine Formel für A^{-1}.* Sei $A = (\alpha_{ij}) \in \mathrm{GL}(n, K)$. Dann nennen wir

$$\operatorname{adj} A := (\alpha^*_{ij})_{n\times n} \quad \text{mit } \alpha^*_{ij} = (-1)^{i+j} \det A_{ji}$$

die *Adjunkte* zu A. Es gilt

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

Beweis: Der (i, j)-Koeffizient von A^{-1} ist die i-te Komponente von $A^{-1}e^j$. Berechne dieses nach der Cramer-Regel als Lösung von $Ax = e^j$.

3. *Berechnung von Volumina*

   a. in der Ebene: Seien $a = \begin{pmatrix}\alpha_1\\ \alpha_2\\ 0\end{pmatrix}$ und $b = \begin{pmatrix}\beta_1\\ \beta_2\\ 0\end{pmatrix}$ zwei Vektoren in der Ebene. Dann gilt für die Fläche A des aufgespannten Parallelogramms:

$$F = |a \times b| = \det\begin{pmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{pmatrix}.$$

   Entsprechend hat das aufgespannte Dreieck nur die halbe Fläche.

   b. im $\mathbf{R}^3$: Für das Volumen V eines von drei Vektoren a, b, c aufgespannten Spats gilt:

$$\begin{aligned}V &= |(a, b, c)| = |\langle a, b \times c\rangle| \qquad \text{(Spatprodukt)}\\ &= |\alpha_1(\beta_2\gamma_3 - \beta_3\gamma_2) - \alpha_2(\beta_3\gamma_1 - \beta_1\gamma_3) + \alpha_3(\beta_1\gamma_2 - \beta_2\gamma_1)|\\ &= \left|\det\begin{pmatrix}\alpha_1 & \beta_1 & \gamma_1\\ \alpha_2 & \beta_2 & \gamma_2\\ \alpha_3 & \beta_3 & \gamma_3\end{pmatrix}\right|.\end{aligned}$$

   Entsprechend gilt für das Volumen V eines Tetraeders mit den Kanten a, b, c:

$$V = \frac{1}{6} \cdot |\det(a, b, c)|.$$

   c. im $\mathbf{R}^n$: Der durch Vektoren $a^1, \ldots, a^n$ aufgespannte n-Spat S ist definiert als

$$S := \left\{ \sum_{i=1}^n \alpha_i a^i \;\middle|\; 0 \le \alpha_i \le 1,\ 1 \le i \le n \right\}.$$

   Dann definieren wir sein „Volumen“ als

$$V := |\det A|, \qquad \text{mit } A := (a^1, \ldots, a^n).$$

   Betrachten wir nun die zu einer Matrix $B \in K^{n\times n}$ gehörige lineare Abbildung $x \mapsto Bx$ (für $x \in K^n$), die den n-Spat S auf den Spat mit den Kanten $(Ba^1, \ldots, Ba^n)$ abbildet. Für das Volumen V' des neuen Spats gilt dann:

$$V' = |\det BA| = |\det B| \cdot |\det A| = |\det B| \cdot V.$$

   Dabei ist $|\det B|$ der *Volumenverzerrungsfaktor*.

4. *Kegelschnitte.* Nun wollen wir mit Hilfe von Determinanten die Gleichung eines Kreises in der Ebene berechnen, wenn dieser durch 3 Punkte $A_i = (x_i, y_i)$ (für $i = 1, 2, 3$) bestimmt ist, d.h. die Punkte dürfen nicht auf einer Geraden liegen.

   Die allgemeine Kreisgleichung in der Ebene lautet:

$$a(x^2 + y^2) + bx + cy + d = 0,$$

   wobei wir hier ohne Einschränkung annehmen können, daß $a \neq 0$ gilt. Nun erhalten wir durch Einsetzen der drei Punkte in diese Kreisgleichung ein lineares homogenes Gleichungssystem für die Koeffizienten a, b, c, d mit einer nichttrivialen Lösung:

$$\begin{aligned}a(x_1^2 + y_1^2) + bx_1 + cy_1 + d &= 0,\\ a(x_2^2 + y_2^2) + bx_2 + cy_2 + d &= 0,\\ a(x_3^2 + y_3^2) + bx_3 + cy_3 + d &= 0.\end{aligned}$$

   Dieses Gleichungssystem besitzt genau dann eine nichttriviale Lösung wenn die Koeffizientenmatrix nicht invertierbar ist. Also muß die Determinante der Koeffizientenmatrix Null sein:

$$\det\begin{pmatrix}x^2 + y^2 & x & y & 1\\ x_1^2 + y_1^2 & x_1 & y_1 & 1\\ x_2^2 + y_2^2 & x_2 & y_2 & 1\\ x_3^2 + y_3^2 & x_3 & y_3 & 1\end{pmatrix} = 0.$$

   Durch Berechnen dieser Determinante erhalten wir sofort die gesuchte Kreisgleichung.

4.4 Determinante eines Endomorphismus

Definition. Sei V ein endlichdimensionaler Vektorraum und f ein Endomorphismus über V. Ferner sei $B = (v_1, \ldots, v_n)$ irgendeine Basis von V. Wir definieren dann

$$\det f := \det {}_B[f]_B.$$

Bemerkung: Die Determinante eines Endomorphismus ist wohldefiniert, das bedeutet sie ist unabhängig von der gewählten Basis. Denn sind C und B Basen, so gilt

$$\det {}_C[f]_C = \det\bigl(T^{-1} \cdot {}_B[f]_B \cdot T\bigr) = \det {}_B[f]_B.$$

Kapitel 5

Eigenwerte

5.1 Eigenwerte und Eigenvektoren

5.1.1 Definitionen

Definition. Seien K ein Körper, V ein K-Vektorraum und f ein Endomorphismus über V. Existieren ein $v \in V$ und $\lambda \in K$ mit

$$v \neq 0 \quad \text{und} \quad f(v) = \lambda v,$$

so nennen wir λ *Eigenwert* von f und v *Eigenvektor* zu λ.
Ferner bezeichnen wir

$$V(f,\lambda) := \{\, v \in V \mid f(v) = \lambda v \,\}$$

als Eigenraum bezüglich λ. Wir schreiben auch kurz $V(\lambda)$, wenn klar ist, um welche Abbildung es sich handelt.

Bemerkung: Wie wir sofort sehen, gilt

$$\begin{aligned} V(f,\lambda) &= \{\, v \in V \mid f(v) = \lambda v \,\} \\ &= \{\, v \in V \mid (\lambda \mathrm{Id} - f)(v) = 0 \,\} \\ &= \mathrm{Kern}(\lambda \mathrm{Id} - f). \end{aligned}$$

Wir werden die Definitionen nun auf Matrizen übertragen:
Definition. Seien K ein Körper und $A \in K^{n\times n}$.
Gibt es ein $u \in K^n$ mit $u \neq 0$ und $\lambda \in K$ derart, daß

$$Au = \lambda u$$

gilt, so nennen wir λ *Eigenwert* von A, u *Eigenvektor* bezüglich λ und

$$V(A,\lambda) := \{\, u \in K^n \mid Au = \lambda u \,\} = \mathrm{Kern}(\lambda E - A)$$

Eigenraum zum Eigenwert λ. Auch hier schreiben wir manchmal kurz $V(\lambda)$.

Bemerkung: Da für $B \in K^{n\times n}$ bekanntlich $\mathrm{Kern}\, B \neq \{0\} \Longleftrightarrow \det B = 0$ gilt, haben wir

$$V(A,\lambda) \neq \{0\} \Longleftrightarrow \det(\lambda E - A) = 0.$$

Beispiele:

1. Sei

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

Dann hat A den Eigenwert 7 mit dem Eigenvektor e^2. Zur Bestimmung sämtlicher Eigenvektoren zum Eigenwert 7 lösen wir die Gleichung $(7E - A)\cdot x = 0$.

2. Sei $V = C^\infty(\mathbf{R})$ und

$$\frac{d}{dx}\colon V \to V, \quad \frac{d}{dx} f(x) = f'(x).$$

Für festes $\lambda \in \mathbf{R}$ und $g(x) = e^{\lambda x}$ erhalten wir $g'(x) = \lambda e^{\lambda x}$. Damit ist λ ein Eigenwert mit Eigenvektor $g(x)$ von d/dx.

◊

5.1.2 Elementare Eigenschaften

Im folgenden sei K ein Körper, V ein K-Vektorraum und $f\colon V \to V$ eine lineare Abbildung mit einem Eigenwert λ bzgl. dem Eigenvektor v. Wir wählen zu V eine Basis $B = (v_1, v_2, \ldots, v_n)$ mit $v_1 := v$. Dann gilt:

$${}_B[f]_B = \begin{pmatrix} \lambda & * & \ldots & * \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}.$$

Diese Vereinfachung ergibt sich aus dem Eigenwert λ und dem Eigenvektor v von f. Diese Darstellung von f kann weiter vereinfacht werden, falls mehrere (oder am besten alle) Eigenwerte und -vektoren von f bekannt sind.

Satz. *Seien $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Eigenwerte von f, d.h. $\lambda_i \neq \lambda_j$ für $i \neq j$, und $v_1, \ldots, v_k$ die dazugehörigen Eigenvektoren ($f(v_i) = \lambda_i v_i$). Dann ist $(v_1, \ldots, v_k)$ linear unabhängig.*

Beweis: Wenden wir f auf den Ansatz $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$ an, so erhalten wir $\lambda_1\alpha_1 v_1 + \cdots + \lambda_k\alpha_k v_k = 0$. Multiplizieren wir nun den Ansatz mit λ_1 und subtrahieren ihn von der zweiten Gleichung, so erhalten wir

$$(\lambda_2 - \lambda_1)\alpha_2 v_2 + \cdots + (\lambda_k - \lambda_1)\alpha_k v_k = 0.$$

Weiter ergibt sich über vollständige Induktion:

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

also

$$\alpha_2 = \cdots = \alpha_k = 0 \quad (\text{beachte } \lambda_i \neq \lambda_1 \text{ für } i \geq 2)$$

und damit auch $\alpha_1 = 0$. □

Folgerung: Besitzt f genau n verschiedene Eigenwerte, dann existiert eine Basis $B = (v_1, \ldots, v_n)$ aus Eigenvektoren.

5.1.3 Bemerkungen

Sei f ein Endomorphismus über V mit dem Eigenwert λ und dem Eigenvektor v. Dann gilt

$$f(v) = \lambda v, \quad f^2(v) = f(\lambda v) = \lambda^2 v, \quad \text{etc.}$$

und allgemein

$$f^k(v) = \lambda^k v.$$

Somit hat f^k den Eigenwert λ^k bzgl. des Eigenvektors v.

Sei ferner $p(x) = \alpha_0 + \alpha_1 x + \cdots + a_s x^s$ ein Polynom mit Koeffizienten aus K. Dann sind die α_i eindeutig bestimmt, also ist $p(f) := \alpha_0 \mathrm{Id} + \alpha_1 f + \cdots + a_s f^s$ ein *Polynom in* f. Dann gilt:

$$p(f)v = \alpha_0 v + \alpha_1 f(v) + \cdots + \alpha_s f^s(v) = p(\lambda)v,$$

also hat $p(f)$ den Eigenwert $p(\lambda)$ – ebenfalls mit dem Eigenvektor v.

Sonderfälle:

a. Sei f *nilpotent*. Dann gibt es eine natürliche Zahl k, so daß $f^k = 0$ ist. Sei nun λ ein Eigenwert von f mit dem Eigenvektor v. Dann gilt:

$$0 = f^k(v) = \lambda^k.$$

Also ist $\lambda = 0$. Somit haben nilpotente Abbildungen ausschließlich den Eigenwert Null.

b. Sei f eine *Projektion*, d.h. $f^2 = f$. Dann gilt für den Eigenwert λ:

$$f(v) = \lambda v = f^2(v) = \lambda^2 v.$$

Somit haben Projektionen nur die Eigenwerte Null und Eins.

Ist dabei f ein Endomorphismus über V, so gilt das bekannte Ergebnis

$$V = \operatorname{Kern} f \oplus \operatorname{Bild} f.$$

Damit gilt für Projektionen $V(0) = \operatorname{Kern} f$ und $V(1) = \operatorname{Bild} f$.

5.2 Die Bestimmung von Eigenwerten

Definition. Seien V ein endlich erzeugter K-Vektorraum, f ein Endomorphismus über V und $A \in K^{n\times n}$. Dann nennen wir

$$p_f(x) := \det(x\,\mathrm{Id} - f) \text{ charakteristisches Polynom von } f \text{ und}$$

$$p_A(x) := \det(xE - A) \text{ charakteristisches Polynom von } A.$$

Satz. *Seien $A, B \in K^{n\times n}$ zueinander ähnlich, d.h. es gibt ein $C \in \mathrm{GL}(n, K)$ derart, daß $B = C^{-1}AC$. Dann gilt*

$$p_A(x) = p_B(x).$$

Beweis: Wir erhalten nach der vierten Folgerung des Multiplikationssatzes für Determinanten und den Rechenregeln für Matrizenmultiplikation

$$p_{C^{-1}AC}(x) = \det(xE - C^{-1}AC) = \det(xC^{-1}C - C^{-1}AC) = \det(xE - A) = p_A(x).$$

□

Beispiele:

1. Sei

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

Dann gilt für das charakteristische Polynom $p_A(x)$:

$$p_A(x) = \det \begin{pmatrix} x-1 & -2 & 0 \\ 0 & x-1 & -3 \\ -1 & 0 & x-2 \end{pmatrix} = (x-1)^2(x-2) - 6 = x^3 - 4x^2 + 5x - 8.$$

2. Sei

$$A := \begin{pmatrix} 0 & 0 & \ldots & 0 & -\alpha_0 \\ 1 & 0 & \ddots & \vdots & -\alpha_1 \\ 0 & 1 & \ddots & 0 & -\alpha_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -\alpha_{n-1} \end{pmatrix} \in K^{n\times n}.$$

Dann gilt für das charakteristische Polynom $p_A(x)$:

$$p_A(x) = \det \begin{pmatrix} x & 0 & \ldots & 0 & \alpha_0 \\ -1 & x & \ddots & \vdots & \alpha_1 \\ 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & x & \alpha_{n-2} \\ 0 & \ldots & 0 & -1 & x + \alpha_{n-1} \end{pmatrix} = x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0.$$

◊

Satz. *Ist f ein Endomorphismus bzw. $A \in K^{n\times n}$, so ist $\lambda \in K$ genau dann ein Eigenwert von f bzw. A, wenn $p_f(\lambda) = 0$ bzw. $p_A(\lambda) = 0$ ist.*

Beweis: Die Behauptung folgt sofort aus der Definition von Eigenwert und der Definition des charakteristischen Polynoms. □

Vorgehensweise zur Berechnung von Eigenwerten und Eigenvektoren:

Im folgenden sei f ein Endomorphismus über dem K-Vektorraum V. Die Vorgehensweise bei Matrizen ist ganz analog.

1. Zunächst bestimmen wir das charakteristische Polynom:

$$p_f(x) = \det(x\,\mathrm{Id} - f) = x^n - (\mathrm{Sp}\, f)x^{n-1} + \cdots + (-1)^n \det f$$

2. Dann faktorisieren wir $p_f(x)$:

$$p_f(x) = (x-\lambda_1)^{k_1} \ldots (x-\lambda_r)^{k_r} q(x),$$

derart, daß $q(x)$ in K keine Nullstelle besitzt und erkennen daraus die Nullstellen von $p_f(x)$.

3. Jetzt berechnen wir zu jedem Eigenwert λ_i $(1 \le i \le r)$ den Eigenraum

$$V(\lambda_i) = \left\{ v \in V \;\middle|\; f(v) = \lambda_i v \right\} = \mathrm{Kern}(\lambda_i\,\mathrm{Id} - f)$$

und davon eine Basis. Nehmen wir nun von jedem Eigenraum eine Basis und setzen diese zusammen, so erhalten wir eine Basis B eines Unterraums von V.

4. Für eine solche Basis B gilt dann

$${}_B[f]_B = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_1 \end{matrix}} & & \boxed{0} & \boxed{*} \\ & \ddots & & \vdots \\ \boxed{0} & & \boxed{\begin{matrix} \lambda_r & 0 & \ldots & 0 \\ 0 & \lambda_r & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_r \end{matrix}} & \boxed{*} \\ \boxed{0} & \cdots & \boxed{0} & \boxed{*} \end{pmatrix}.$$

Definition. Sei λ ein Eigenwert des Endomorphismus f. Dann heißt die Vielfachheit der Nullstelle λ in $p_f(x)$ *algebraische Vielfachheit des Eigenwertes* λ.

Definition. Sei $V(\lambda)$ der Eigenraum zum Eigenwert λ. Die Dimension des Eigenraumes $V(\lambda)$ heißt dann *geometrische Vielfachheit des Eigenwertes* λ.

Satz. *Für die geometrische Vielfachheit* $\mathrm{Dim}\, V(\lambda)$ *eines Eigenwertes* λ *mit algebraischer Vielfachheit* k *gilt:*

$$1 \le \mathrm{Dim}\, V(\lambda) \le k.$$

Beweis: Mit einer Basis, die möglichst viele Eigenvektoren enthält, gilt dann (vgl. auch mit Punkt 4 oben)

$$p_f(x) = \left(\prod_i (x-\lambda_i)^{\mathrm{Dim}\, V(\lambda_i)} \right) h(x),$$

d.h. $x - \lambda_i$ kommt in $p_f(x)$ mindestens $\mathrm{Dim}\, V(\lambda_i)$-mal vor, also $\mathrm{Dim}\, V(\lambda_i) \le k_i$. □

Beispiel: Sei

$$A := \begin{pmatrix} 0 & 1 & \ldots & 1 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \ldots & 0 & 0 \end{pmatrix}$$

mit $A^n = 0$, d.h. A ist nilpotent. Dann gilt für das charakteristische Polynom $p_f(x) = x^n$. Somit hat A den einzigen Eigenwert Null mit der algebraischen Vielfachheit n. Die geometrische Vielfachheit ist Eins, da

$$V(0) = \left\{ x \;\middle|\; Ax = 0 \right\} = \left\{ \alpha e^1 \;\middle|\; \alpha \in K \right\} = Ke^1.$$

◊

5.3 Trigonalisierung

Definition. Eine Matrix $A \in K^{n\times n}$ heißt *trigonalisierbar* oder *triangulierbar*, wenn es eine zu A ähnliche Matrix gibt, die obere Dreiecksgestalt hat. In anderen Worten, wenn es ein $B \in \mathrm{GL}(n, K)$ gibt, so daß

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

gilt.
Ein Endomorphismus f auf V heißt genau dann trigonalisierbar, wenn seine Darstellungsmatrix trigonalisierbar ist:

$${}_B[f]_B = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}.$$

Eine Matrix, die wir auf obere Dreiecksgestalt gebracht haben bzw. ein Endomorphismus, dessen Darstellungsmatrix obere Dreiecksgestalt besitzt, hat *Schur-Normalform.*

Satz. *Ist ein Endomorphismus f bzw. eine Matrix A trigonalisierbar, so sind die Koeffizienten in der Hauptdiagonalen von ${}_B[f]_B$ bzw. $B^{-1}AB$ die (nicht notwendigerweise verschiedenen) Eigenwerte von f bzw. A.*

Beweis: Es gilt

$$\det(x\,\mathrm{Id} - f) = \det(xE - {}_B[f]_B) = \prod_{i=1}^{n}(x - \lambda_i),$$

denn $(xE - {}_B[f]_B)$ hat obere Dreiecksgestalt und in der Hauptdiagonalen steht $x-\lambda_1, \ldots, x-\lambda_n$. Da das charakteristische Polynom seine Nullstellen genau bei den Eigenwerten hat, müssen die λ_i $(1 \le i \le n)$ die Eigenwerte von f sein. □

Satz. *Ein Endomorphismus f bzw. eine Matrix A eines endlichdimensionalen K-Vektorraumes V ist genau dann trigonalisierbar, wenn alle Eigenwerte von f bzw. A in K liegen, d.h. $p_f(x) = \prod(x-\lambda_i)^{k_i}$ bzw. $p_A(x) = \prod(x-\lambda_i)^{k_i}$ gilt.*

Beweis: Wir beweisen den Satz hier nur für Matrizen. Der Beweis für Endomorphismen verläuft analog.
Ist A trigonalisierbar, so erhalten wir für das charakteristische Polynom

$$p_A(x) = p_{B^{-1}AB} = \prod_{i=1}^{n}(x - \lambda_i).$$

Sind andererseits alle Eigenwerte in K, gilt also

$$p_A(x) = \prod_{i=1}^{n}(x - \lambda_i),$$

so gibt es zum Eigenwert λ_1 einen Eigenvektor v_1. Wir ergänzen zur Basis $C = (v_1, w_2, \ldots, w_n)$ und erhalten

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

Wegen $p_A(x) = (x-\lambda_1)p_{A_1}(x)$ hat auch A_1 alle Eigenwerte in K und somit gibt es nach Induktionsvoraussetzung eine Matrix $D \in GL(n-1, K)$, so daß gilt:

$$D^{-1}A_1D = \begin{pmatrix} \lambda_2 & * & \ldots & * \\ 0 & \lambda_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}.$$

Für

$$B := C\begin{pmatrix} 1 & * & \ldots & * \\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{pmatrix}$$

folgt nun die Behauptung

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}.$$

□

Satz von Schur.

(1) Jeder Endomorphismus eines endlich erzeugten $\mathbf{C}$-Vektorraums ist trigonalisierbar.

(2) Jede Matrix $A \in \mathbf{C}^{n\times n}$ ist trigonalisierbar.

Beweis: Über dem Körper $\mathbf{C}$ zerfällt jedes Polynom in Linearfaktoren, d.h. alle Eigenwerte von f bzw. A liegen im Körper. Somit ist f bzw. A trigonalisierbar. □

5.4 Diagonalisierung

Definition. Ein Endomorphismus f über V heißt genau dann *diagonalisierbar*, wenn es eine Basis B gibt, so daß die Darstellungsmatrix von f bezüglich der Basis B Diagonalgestalt hat:

$${}_B[f]_B = \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Satz.

(1) Falls $B = (v_1, \ldots, v_n)$ eine Basis aus Eigenvektoren mit $f(v_i) = \lambda_i v_i$ ist (die $\lambda_i \in K$ müssen nicht notwendigerweise verschieden sein), so gilt

$${}_B[f]_B = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$$

und f ist daher diagonalisierbar.

(2) Seien $A \in K^{n\times n}$ und B eine Basis aus Eigenvektoren. Dann hat $B^{-1}AB$ Diagonalgestalt.

Beweis: Wir beweisen hier nur *(2)*, denn der Beweis von *(1)* verläuft ganz analog.
Es gilt

$$\begin{aligned} AB &= (Ab^1, \ldots, Ab^n) = (\lambda_1 b^1, \ldots, \lambda_n b^n) \\ &= (\lambda_1 Be^1, \ldots, \lambda_n Be^n) = B(\lambda_1 e^1, \ldots, \lambda_n e^n) \\ &= B \cdot \mathrm{Diag}(\lambda_1, \ldots, \lambda_n). \end{aligned}$$

Daraus folgt:

$$B^{-1}AB = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n).$$

□

Satz. *Ein Endomorphismus bzw. eine Matrix ist genau dann diagonalisierbar, wenn*

- *das charakteristische Polynom in Linearfaktoren zerfällt und*
- *die geometrische Vielfachheit aller Eigenwerte gleich der algebraischen Vielfachheit ist.*

Beweis: Betrachten wir einen n-dimensionalen Vektorraum und bezeichnen die algebraische Vielfachheit der sämtlichen Eigenwerte $\lambda_1, \dots, \lambda_r$ mit $\mu_1, \dots, \mu_r$ respektive, so gilt $\sum_{i=1}^{r} \mu_i = n$. Damit ist nach Voraussetzung die Dimension aller Eigenräume zusammen n, d.h. es existiert eine Basis aus Eigenvektoren. Daraus folgt mit dem vorherigen Satz die Diagonalisierbarkeit. □

5.5 Der Satz von Cayley und Hamilton

Sei f ein Endomorphismus über dem endlich erzeugten Vektorraum V und U ein f-invarianter Unterraum von V. Dann betrachten wir die Abbildungen

- $f|_U \colon U \to U$, $f|_U(u) = f(u)$ für alle $u \in U$ (Restriktion auf U) und
- $\bar{f} \colon V/U \to V/U$, $\bar{f}(v + U) = f(v) + U$.

Nun wählen wir uns eine Basis $C := (u_1, u_2, \dots, u_r)$ von U und ergänzen diese zu einer Basis $B := (u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_n)$ von V. Dabei ist $D := (u_{r+1} + U, u_{r+2} + U, \dots, u_n + U)$ eine Basis von V/U. Für die Abbildungsmatrix von f erhalten wir somit

$$ {}_B[f]_B = \begin{pmatrix} \boxed{\ {}_C[f|_U]_C\ } & \boxed{\ *\ } \\ \boxed{\ 0\ } & \boxed{\ {}_D[\bar{f}]_D\ } \end{pmatrix}. $$

Satz. *Sei f ein Endomorphismus über dem endlich erzeugten Vektorraum V und U ein f-invarianter Unterraum von V. Dann gilt*

$$ p_f(x) = p_{\bar{f}}(x) \cdot p_{f|_U}(x). $$

Beweis: Die Behauptung sehen wir aufgrund der Vorbetrachtung sofort ein, denn nach dem Kästchensatz gilt

$$ p_f(x) = \det\bigl(xE - {}_B[f]_B\bigr) = \det\bigl(xE - {}_C[f|_U]_C\bigr) \cdot \det\bigl(xE - {}_D[\bar{f}]_D\bigr) = p_{f|_U}(x) \cdot p_{\bar{f}}(x). $$

□

Sei ein beliebiger Vektor $v \in V$ gegeben. Dann wählen wir k so, daß

$$ B := \bigl(v, f(v), f^2(v), \dots, f^k(v)\bigr) $$

linear unabhängig und zugleich $\bigl(v, f(v), f^2(v), \dots, f^k(v), f^{k+1}(v)\bigr)$ linear abhängig ist, d.h. B ist eine maximale linear unabhängige Familie und damit eine Basis. Dann ist offensichtlich $U := \mathrm{Lin}\bigl(v, f(v), f^2(v), \dots, f^k(v)\bigr)$ ein f-invarianter Unterraum und für die Darstellungsmatrix von $f|_U$ gilt bei $\alpha_0 v + \alpha_1 f(v) + \cdots + \alpha_k f^k(v) + f^{k+1}(v) = 0$:

$$ {}_B[f|_U]_B = \begin{pmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \ddots & \vdots & -\alpha_1 \\ 0 & 1 & \ddots & 0 & -\alpha_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & -\alpha_k \end{pmatrix}. $$

Daraus folgt

$$ p_{f|_U}(x) = x^{k+1} + \alpha_k x^k + \cdots + a_1 x + \alpha_0 $$

und weiter

$$ p_{f|_U}(f) \equiv 0. $$

Satz von Cayley und Hamilton.
Seien V ein endlich erzeugter Vektorraum und f ein Endomorphismus über V bzw. $A \in K^{n\times n}$. Dann ist

$$ p_f(f) = 0 \quad \textit{bzw.} \quad p_A(A) = 0. $$

Beweis: Seien $v \in V \setminus \{0\}$ und ein f-invarianter Unterraum U mit Basis $\bigl(v, f(v), \dots, f^k(v)\bigr)$ gegeben. Dann gilt nach obiger Betrachtung $p_{f|_U}(f)(v) = 0$ für alle $v \in V$ und damit

$$ p_f(f)(v) = p_{\bar{f}}(f) \cdot p_{f|_U}(f)(v) = 0 \quad \text{für alle } v \in V. $$

Das bedeutet $p_f(f) \equiv 0$.
Der Beweis für Matrizen verläuft analog. □

Nach dem Satz von Cayley und Hamilton liegt das charakteristische Polynom $p_f(x)$ im Ideal

$$ I = \bigl\{\, q(x) \in K[x] \;\bigm|\; q(f) = 0 \,\bigr\}. $$

Definition. Das *normierte Polynom kleinsten Grades* in I heißt *Minimalpolynom* von f. Wir bezeichnen es mit $m_f(x)$.

Wir finden das Minimalpolynom $m_f(x)$, indem wir ein minimales k bestimmen, so daß

$$ (\mathrm{Id}, f, f^2, \dots, f^{k-1}) $$

linear unabhängig ist, jedoch $(\mathrm{Id}, f, f^2, \dots, f^k)$ linear abhängig. Dann gilt $f^k + \alpha_{k-1} f^{k-1} + \cdots + \alpha_1 f + \alpha_0 \mathrm{Id} = 0$ und $m_f(x) = x^k + \alpha_{k-1} x^{k-1} + \cdots + \alpha_1 x + \alpha_0$.

Nach der Theorie aus Kapitel 2.4.3 gilt $I = m_f(x)K[x]$, insbesondere gibt es $h(x) \in K[x]$, so daß

$$p_f(x) = m_f(x)h(x) \quad (\text{„}m_f \text{ teilt } p_f\text{“}).$$

Satz. *Die Polynome $m_f(x)$ und $p_f(x)$ haben dieselben Nullstellen (im allgemeinen aber in verschiedener Vielfachheit).*

Beweis: Aus $m_f(x) = 0$ folgt $p_f(x) = 0$, denn m_f teilt ja p_f.
Ist nun $p_f(\lambda) = 0$, also λ Eigenwert mit Eigenvektor $u \neq 0$, so gilt $f(u) = \lambda u$ und $q(f)(u) = q(\lambda) \cdot u$ für alle $q(x) \in K[x]$. Insbesondere ist dann für $q = m_f$

$$0 = m_f(f)(u) = m_f(\lambda) \cdot u,$$

also $m_f(\lambda) = 0$. □

Kapitel 6

Euklidische und unitäre Vektorräume

In diesem Kapitel beschränken wir uns auf Vektorräume über den Grundkörpern $\mathbf{R}$ und $\mathbf{C}$.

6.1 Das Skalarprodukt

Definition. Für $z = x + iy$ $(x, y \in \mathbf{R})$ setzen wir

$$\bar{z} := x - iy.$$

Die Abbildung $z \mapsto \bar{z}$ heißt *Konjugation*.

Sie besitzt folgende Eigenschaften:

$$\begin{aligned} \alpha = \bar{\alpha} &\iff \alpha \in \mathbf{R}, \\ \bar{\bar{z}} &= z, \\ \overline{z + w} &= \bar{z} + \bar{w}, \\ \overline{zw} &= \bar{z} \cdot \bar{w}. \end{aligned}$$

Definition. Seien V ein K-Vektorraum und $\varphi\colon V \times V \to K$ eine Abbildung mit folgenden Eigenschaften:

$$\varphi(\lambda u + v, w) = \lambda\varphi(u, w) + \varphi(v, w) \qquad \text{für alle } \lambda \in K \text{ und alle } u, v, w \in V, \tag{S1}$$

$$\varphi(u, v) = \overline{\varphi(v, u)}. \tag{S2}$$

Für den Fall $K = \mathbf{R}$ nennen wir φ *symmetrische Bilinearform*, für den Fall $K = \mathbf{C}$ *hermitsche Form* (nach dem französischen Mathematiker *Charles Hermite*).

Bemerkung:

1. Die Abbildung φ ist *im ersten Argument linear*, d.h. bei festem $w \in V$ ist $v \mapsto \varphi(v, w)$ eine *Linearform*.

2. Über dem Körper $\mathbf{R}$ ist φ symmetrisch und auch *im zweiten Argument linear*:

$$\varphi(u, v) = \varphi(v, u) \qquad \varphi(u, \lambda v + w) = \lambda\varphi(u, v) + \varphi(u, w).$$

3. Über dem Körper $\mathbf{C}$ gilt:

$$\varphi(u, \lambda v) = \overline{\varphi(\lambda v, u)} = \bar{\lambda} \cdot \overline{\varphi(v, u)} = \bar{\lambda}\varphi(u, v).$$

Speziell gilt dann

$$\varphi(u, iv) = -i\varphi(u, v).$$

Definition. Sei φ eine Abbildung mit den oben beschriebenen Eigenschaften. Dann heißt φ *positiv definit*, wenn gilt:

$$\varphi(v, v) > 0 \quad \text{für alle } v \neq 0.$$

Weiter heißt φ *positiv semidefinit*, wenn gilt:

$$\varphi(v, v) \geq 0 \quad \text{für alle } v \in V.$$

Gibt es $u, v \in V$ derart, daß

$$\varphi(u, u) < 0 \quad \text{und} \quad \varphi(v, v) > 0$$

so heißt φ *indefinit*.
Analog werden die Begriffe *negativ definit* und *negativ semidefinit* definiert.

Definition. Ein Vektorraum über $\mathbf{R}$ mit einer positiv definiten, symmetrischen Bilinearform φ heißt *euklidisch*, ein Vektorraum über $\mathbf{C}$ mit einer positiv definiten, hermitschen Form φ *unitär*.
Die Abbildung φ heißt in diesen Fällen *inneres Produkt* oder *Skalarprodukt auf* V.
Der Vektorraum V und die Abbildung φ bilden zusammen den *Innenproduktraum* (V, φ).

Wir können bereits jetzt eine sehr wichtige Aussage über das Skalarprodukt machen:
Lemma 6.1: *Ist φ ein Skalarprodukt über dem Vektorraum V und gilt $\varphi(u, v) = 0$ für alle $v \in V$, so folgt $u = 0$.*

Beweis: Da auch $\varphi(u, u) = 0$, folgt die Behauptung sofort aus der Definition. □

Beispiele für Skalarprodukte:

1. Das auf dem Vektorraum $\mathbf{R}^n$ definierte Skalarprodukt φ,

$$\varphi(x, y) := x^T y,$$

heißt *kanonisches* oder *natürliches Skalarprodukt* des $\mathbf{R}^n$. Oft schreiben wir statt $\varphi(x, y)$ auch $\langle x, y \rangle$.

2. Die auf $\mathbf{C}^n$ definierte Abbildung φ mit

$$\varphi(x, y) := x^T \bar{y}$$

ist *hermitsch* und *positiv definit*. Sie ist das *kanonische (natürliche) Skalarprodukt* des $\mathbf{C}^n$.

3. Für $A \in \mathrm{GL}(n, \mathbf{R})$ ist $\varphi_A(x, y) := \langle Ax, Ay \rangle$ ein Skalarprodukt.

4. Sei $V = C([0, 1], K)$ der Vektorraum aller über $[0, 1]$ integrierbarer Funktionen. Dann ist

$$\varphi(f, g) := \int_0^1 f(t)\overline{g(t)}dt$$

ein Skalarprodukt.

◊

6.2 Orthogonalisierung und Orthonormalisierung

6.2.1 Die Norm

Sei im folgenden V ein euklidischer oder ein unitärer Vektorraum mit Skalarprodukt φ.

Definition. Wir definieren die *Norm* bzw. den *Betrag eines Vektors* $v \in V$ folgendermaßen:

$$\|v\| := \sqrt{\varphi(v, v)}.$$

Definition. Zwei Vektoren $u, v \in V$ heißen *orthogonal*, wenn $\varphi(u, v) = 0$ ist. In Zeichen:

$$u \perp v \iff \varphi(u, v) = 0.$$

Es gilt offenbar $u \perp v \iff v \perp u$.

Satz. *Für beliebige Vektoren $u, v \in V$ gilt:*

(1) $u \perp v \implies \|u + v\|^2 = \|u\|^2 + \|v\|^2$ *(Satz von Pythagoras).*

(2) $\left|\varphi(u, v)\right| \leq \|u\| \cdot \|v\|$ *(Ungleichung von Cauchy-Schwarz).*

Beweis:

(1) Für orthogonale Vektoren u, v gilt:

$$\|u + v\|^2 = \varphi(u + v, u + v) = \varphi(u, u) + \varphi(v, u) + \varphi(u, v) + \varphi(v, v) = \|u\|^2 + \|v\|^2.$$

(2) Ist $u = 0$, so ist auch die linke und die rechte Seite der Ungleichung Null.
Sei also im folgenden $u \neq 0$. Dann gilt

$$u \perp v - \frac{\varphi(v, u)}{\|u\|^2}u.$$

Daraus folgt mit *(1)*

$$\|v\|^2 = \left\| \frac{\varphi(v, u)}{\|u\|^2}u + v - \frac{\varphi(v, u)}{\|u\|^2}u \right\|^2 = \frac{\left|\varphi(v, u)\right|^2}{\|u\|^2} + \left\| v - \frac{\varphi(v, u)}{\|u\|^2}u \right\|^2 \geq \frac{\left|\varphi(v, u)\right|^2}{\|u\|^2}$$

und damit die Behauptung.

□

Polarisierung. *Sei eine Norm $\|\,\|$ gegeben. Dann können wir daraus das der Norm zugeordnete Skalarprodukt φ wie folgt rekonstruieren:*

Körper $K = \mathbf{R}$: *Es gilt*

$$\|x+y\|^2 = \varphi(x+y,x+y) = \varphi(x,x) + 2\varphi(x,y) + \varphi(y,y) = \|x\|^2 + 2\varphi(x,y) + \|y\|^2,$$

also gilt für das Skalarprodukt

$$\varphi(x,y) = \frac{1}{2}\big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\big).$$

Körper $K = \mathbf{C}$: *Hier verifizieren wir durch Nachrechnen*

$$\varphi(x,y) = \frac{1}{4}\big(\|x+y\|^2 + \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big).$$

6.2.2 Das Schmidtsche Orthogonalisierungsverfahren

Definition. Sei (V,φ) wieder ein Vektorraum mit innerem Produkt φ. Ferner seien die Vektoren $v_1,\dots,v_n$ aus V. Dann nennen wir

$$(v_1,\dots,v_n) \text{ ein } \textit{Orthogonalsystem}, \text{ falls } \varphi(v_i,v_j) = 0 \quad \text{für } i \neq j,$$
$$(v_1,\dots,v_n) \text{ ein } \textit{Orthonormalsystem}, \text{ falls } \varphi(v_i,v_j) = \delta_{ij}.$$

Analog werden auch die Begriffe *Orthogonal-* und *Orthonormalbasis* definiert.

Satz. *Ist $(v_1,\dots,v_r)$ ein Orthogonalsystem, so sind die $v_1,\dots,v_r$ linear unabhängig.*

Beweis: Gilt $0 = \sum_{i=1}^{r} \alpha_i v_i$, so gilt auch

$$0 = \varphi\Big(\sum \alpha_i v_i, v_k\Big) = \sum \alpha_i \varphi(v_i,v_k) = c_k\alpha_k \quad \text{mit } c_k \neq 0$$

für alle k mit $1 \le k \le r$. □

Das Schmidtsche Orthogonalisierungsverfahren. *Sei (V,φ) euklidisch oder unitär und $(v_1,\dots,v_n)$ eine Basis von V. Ferner definieren wir:*

$$\begin{aligned} w_1 &:= v_1, \\ w_2 &:= v_2 - \frac{\varphi(v_2,w_1)}{\|w_1\|^2}\cdot w_1, \\ &\vdots \\ w_n &:= v_n - \sum_{k=1}^{n-1} \frac{\varphi(v_n,w_k)}{\|w_k\|^2}\cdot w_k. \end{aligned}$$

Dann sind

(1) *alle w_i ungleich Null,*

(2) *die Vektoren $w_1,\dots,w_n$ orthogonal und*

(3) *es gilt $(w_1,\dots,w_n) \in \mathrm{Lin}(v_1,\dots,v_n)$.*

Beweis:

(1) Offensichtlich, denn die v_i sind linear unabhängig.

(2) Folgt sofort nach Konstruktion.

(3) Trivial, denn jedes w_k wird durch die $v_1,\dots,v_k$ konstruiert.

□

Im folgenden sei V ein endlich erzeugter euklidischer oder unitärer Vektorraum mit Innenprodukt φ.

Satz. *Jeder derartige Vektorraum V besitzt eine Orthonormalbasis $(u_1,\dots,u_n)$.*

Beweis: Zuerst bestimmen wir eine beliebige Basis $(v_1,\dots,v_n)$ von V und bilden dann mit der Schmidtschen Orthogonalisierung eine Orthogonalbasis $(w_1,\dots,w_n)$. Schließlich normieren wir mit $u_i = w_i/\|w_i\|$ und erhalten so die Orthonormalbasis $(u_1,\dots,u_n)$. □

Satz. *Jede Orthonormalbasis eines Untervektorraumes U von V läßt sich zu einer Orthonormalbasis von V ergänzen.*

Beweis: Zuerst ergänzen wir die Orthonormalbasis von U zu einer Basis von V. Nun wenden wir das Schmidtsche Orthogonalisierungsverfahren auf die neue Basis an. Dabei ändern sich die Vektoren aus der Basis von U nicht. Daraus folgt die Behauptung. □

Satz. *Sei $(v_1,\dots,v_n)$ eine Orthonormalbasis von V. Dann gilt für alle u:*

$$u = \sum_{i=1}^{n} \varphi(u,v_i)v_i,$$
$${}_B[u] = \begin{pmatrix} \varphi(u,v_1) \\ \vdots \\ \varphi(u,v_n) \end{pmatrix}.$$

Beweis: Sei $u = \sum_{i=1}^{n} \alpha_i v_i$. Dann ist $\varphi(u,v_k) = \alpha_k$. Daraus ergeben sich beide Behauptungen. □

Nun führen wir noch die Orthogonalität von Vektorräumen zueinander ein:

Definition. Zwei Untervektorräume V, W von U heißen *orthogonal*, wenn $v \perp w$ für alle $v \in V$ und $w \in W$ gilt. In Zeichen:

$$V \perp W.$$

Definition. Wir definieren das *orthogonale Komplement* eines Unterraumes U von V folgendermaßen:

$$U^\perp := \{\, v \in V \mid \varphi(v,u) = 0 \quad \text{für alle } u \in U \,\}.$$

Definition. Sind $U_1, \dots, U_n$ Unterräume von V und gilt

$$V = U_1 + \cdots + U_n,$$
$$U_i \perp U_j \qquad \text{für } 1 \le i, j \le n \text{ und } i \ne j,$$

so nennen wir V die *orthogonale Summe* von $U_1, \dots, U_n$ und schreiben dafür

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_n.$$

Satz. *Sei $U \subseteq V$ ein Unterraum von V und $U^\perp$ sein orthogonales Komplement. Dann gilt:*

$$V = U \oplus U^\perp, \quad \left(U^\perp\right)^\perp = U, \quad \operatorname{Dim} V = \operatorname{Dim} U + \operatorname{Dim} U^\perp.$$

Beweis: Wir wollen hier nur die erste Aussage beweisen. Die beiden anderen Aussagen sind dem Leser als Übung empfohlen.
Sei $(u_1, \dots, u_r)$ eine orthonormale Basis von U. Wir ergänzen diese Basis zu einer orthonormalen Basis $(u_1, \dots, u_r, \dots, u_n)$ von V. Sei nun $v \in U^\perp$. Dann gilt

$$\varphi(v, u_i) = 0, \qquad 1 \le i \le r$$

und

$$v = \sum_{i=r+1}^{n} \varphi(v, u_i) u_i.$$

Somit ist $\operatorname{Lin}(u_{r+1}, \dots, u_n) = U^\perp$. Daraus folgt unmittelbar die Behauptung. □

Lemma 6.2: *Sei K ein Körper und $A, B \in K^{n\times n}$. Gilt dann*

$$v^T A w = v^T B w \qquad \textit{für alle } v, w \in K^n,$$

so folgt $A = B$.

Beweis: Sei $A = (\alpha_{ij})$ und $B = (\beta_{ij})$. Durch Einsetzen der kanonischen Basisvektoren des K^n erhalten wir

$$\alpha_{ij} = e_i^T A e_j = e_i^T B e_j = \beta_{ij} \qquad \text{für } 1 \le i, j \le n.$$

□

6.3 Adjungierte Abbildungen

Im folgenden bezeichnen wir mit V stets einen euklidischen oder unitären Vektorraum mit Skalarprodukt φ.

Definition. Sei f ein Endomorphismus über V. Dann heißt $f^* \in \operatorname{End} V$ die zu f *adjungierte Abbildung*, wenn gilt:

$$\varphi\big(f(u), v\big) = \varphi\big(u, f^*(v)\big) \quad \text{für alle } u, v \in V.$$

Beispiel für eine adjungierte Abbildung: Im Körper $\mathbf{C}$ ist $f^*(x) := \bar{A}^T x$ die Adjungierte zu $f(x) := Ax$:

$$\langle Ax, y\rangle = (Ax)^T \bar{y} = x^T A^T \bar{y} = \langle x, \bar{A}^T y\rangle.$$

◊

Satz. *Zu jedem Endomorphismus f gibt es eine eindeutig bestimmte adjungierte Abbildung f^* mit*

$$\varphi\big(f(u), v\big) = \varphi\big(u, f^*(v)\big) \quad \textit{für } u, v \in V.$$

Beweis: Zunächst beweisen wir, daß es eine adjungierte Abbildung gibt. Wählen wir eine Orthonormalbasis $(u_1, \dots, u_n)$, so ist mit $u = \sum \varphi(u, u_i) u_i$

$$f(u) = \sum \varphi(u, u_i) f(u_i).$$

Damit gilt:

$$\varphi\big(f(u), v\big) = \sum \varphi\big(f(u_i), v\big)\varphi(u, u_i) = \varphi\Big(u, \sum \overline{\varphi\big(f(u_i), v\big)} u_i\Big) = \varphi\big(u, f^*(v)\big).$$

Jetzt zeigen wir noch, daß die adjungierte Abbildung eindeutig bestimmt ist, d.h. wir überzeugen uns davon, daß $g = f^*$ aus $\varphi\big(f(u), v\big) = \varphi\big(u, f^*(v)\big) = \varphi\big(u, g(v)\big)$ folgt. Dies sehen wir leicht ein, denn

$$\varphi\big(u, (f^* - g)(v)\big) = 0 \quad \text{für alle } u \text{ und somit } f^*(v) = g(v) \text{ für alle } v.$$

□

Nun wollen wir die Matrixdarstellung von f^* bzgl. einer Orthonormalbasis $B = (u_1, \dots, u_n)$ berechnen. Für f erhalten wir

$${}_B[f]_B = \begin{pmatrix} \varphi\big(f(u_1), u_1\big) & \dots & \varphi\big(f(u_n), u_1\big) \\ \vdots & \ddots & \vdots \\ \varphi\big(f(u_1), u_n\big) & \dots & \varphi\big(f(u_n), u_n\big) \end{pmatrix}.$$

Somit gilt:

$$F := {}_B[f]_B = \Big(\varphi\big(f(u_j), u_i\big)\Big).$$

Analog gilt für $f^*(v) := \sum \overline{\varphi\big(f(u_i), v\big)} u_i$ und $f^*(u_j) = \sum \overline{\varphi\big(f(u_k), u_j\big)} u_k$:

$$F^* := {}_B[f^*]_B = \Big(\varphi\big(f^*(u_j), u_i\big)\Big) = \Big(\overline{\varphi\big(f(u_i), u_j\big)}\Big).$$

Damit folgt

$$F^* = \bar{F}^T$$

und ferner für die Spur und die Determinante der adjungierten Abbildung

$$\operatorname{Sp} f^* = \overline{\operatorname{Sp} f}, \qquad \det f^* = \overline{\det f}.$$

Wir wollen nun einfache Rechenregeln für die Adjungierte angeben:

$$(f + g)^* = f^* + g^*,$$
$$(\alpha f)^* = \bar{\alpha} f^*,$$
$$(f \circ g)^* = g^* \circ f^*,$$
$$f^{**} = f.$$

Beweis: Durch Einsetzen in das Skalarprodukt φ erhalten wir

$$\begin{aligned}
\varphi\big((f+g)(u),v\big) &= \varphi\big(f(u),v\big) + \varphi\big(g(u),v\big) = \varphi\big(u,f^*(v)\big) + \varphi\big(u,g^*(v)\big),\\
\varphi\big(\alpha f(u),v\big) &= \alpha\varphi\big(f(u),v\big) = \alpha\varphi\big(u,f^*(v)\big) = \varphi\big(u,\bar\alpha f^*(v)\big),\\
\varphi\big((f\circ g)(u),v\big) &= \varphi\big(g(u),f^*(v)\big) = \varphi\big(u,(g^*\circ f^*)(v)\big),\\
\varphi\big(f(u),v\big) &= \varphi\big(u,f^*(v)\big) = \overline{\varphi\big(f^*(v),u\big)} = \overline{\varphi\big(v,f^{**}(u)\big)} = \varphi\big(f^{**}(u),v\big).
\end{aligned}$$

Die Gleichheit von f^{**} und f sehen wir ein, wenn wir uns klar machen, daß $\varphi\big(f(u),v\big) = \varphi\big(f^{**}(u),v\big)$ bedeutet, daß $\varphi\big((f-f^{**})(u),v\big) = 0$ für alle v gilt, d.h. $f(u) = f^{**}(u)$ für alle u ist. □

In den folgenden Abschnitten wollen wir genauer auf besondere Endomorphismen und ihre adjungierten Abbildungen eingehen.

6.4 Orthogonale und unitäre Endomorphismen

6.4.1 Definitionen und Eigenschaften

Im folgenden sei (V,φ) ein Innenproduktraum über dem Körper K. Dabei sei K stets $\mathbf{R}$ bzw. $\mathbf{C}$.

Definition. Ein Endomorphismus f über V heißt *orthogonal* bzw. *unitär*, wenn gilt:

$$\varphi(u,v) = \varphi\big(f(u),f(v)\big) \quad\Longleftrightarrow\quad f^*\circ f = f\circ f^* = \mathrm{Id}\,.$$

Definition. Eine invertierbare Matrix A heißt *orthogonal* bzw. *unitär*, wenn gilt:

$$A\bar A^T = E \quad\Longleftrightarrow\quad A^{-1} = \bar A^T.$$

Eine orthogonale Matrix heißt *eigentlich orthogonal*, wenn zusätzlich $\det A = 1$ gilt.

Daß diese Definitionen wohldefiniert sind, zeigt der folgende Satz, der auch gleich eine weitere praktische Eigenschaft von orthogonalen Endomorphismen beschreibt:

Satz. *Sei f ein Endomorphismus über V und A eine darstellende Matrix von f bezüglich einer Orthonormalbasis von V. Dann sind folgende Aussagen äquivalent:*

(*1*) *Die Abbildung f ist orthogonal bzw. unitär.*

(*2*) *Für alle $v\in V$ gilt $\|f(v)\| = \|v\|$.*

(*3*) *Die Matrix A ist orthogonal bzw. unitär.*

(*4*) *Die Spaltenvektoren von A bilden eine Orthonormalbasis des K^n.*

(*5*) *Die Zeilenvektoren von A bilden eine Orthonormalbasis des K_n.*

Beweis: Da $\|f(v)\|^2 = \varphi\big(f(v),f(v)\big)$ und $\|v\|^2 = \varphi(v,v)$, folgt die Äquivalenz von (*1*) und (*2*) mittels Polarisierung aus der Definition.
Nun zeigen wir, daß (*3*) mit (*4*) und (*5*) äquivalent ist: Sei $A = (u^1,\dots,u^n)\in \mathrm{GL}(n,K)$ orthogonal bzw. unitär. Dann folgt aus $AA^T = E$ sofort $u_iu_j^T = \delta_{ij}$, was genau dann gilt, wenn die Zeilen $(u_1,\dots,u_n)$ bzw. Spalten der Matrix A eine Orthonormalbasis des K_n bzw. K^n sind. □

Damit können wir also einfach von orthogonalen bzw. unitären Endomorphismen sprechen, ohne daß wir uns dabei auf die Abbildung oder die darstellende Matrix festlegen.

Definition. Die Mengen

$$\begin{aligned}
O(n) &:= \big\{\, A\in \mathrm{GL}(n,\mathbf{R}) \;\big|\; A^TA = E \,\big\},\\
U(n) &:= \big\{\, A\in \mathrm{GL}(n,\mathbf{C}) \;\big|\; \bar A^TA = E \,\big\} \quad\text{und}\\
SO(n) &:= \big\{\, A\in O(n) \;\big|\; \det A = 1 \,\big\}
\end{aligned}$$

der *orthogonalen*, *unitären* bzw. *eigentlich orthogonalen* Matrizen bilden zusammen mit der Matrizenmultiplikation jeweils eine Gruppe.
Man nennt diese Gruppen auch die *orthogonale*, die *unitäre* bzw. die *spezielle orthogonale* Gruppe. Letztere wird manchmal auch als *Drehgruppe* bezeichnet.

Bemerkung:

1. Für $A\in O(n)$ gilt:

$$\langle Ax,Ay\rangle = (Ax)^T(Ay) = x^TA^TAy = x^Ty = \langle x,y\rangle.$$

   Unter $x\mapsto Ax$ bleiben Längen und Winkel gleich groß. Deshalb heißt die Abbildung dann auch *längen-* und *winkeltreu*.

2. Für $A\in O(n)$ gilt:

$$(\det A)^2 = \det A\det A^T = \det(A^TA) = \det E = 1.$$

   Also ist die Determinante einer orthogonalen Matrix immer gleich ±1.

Satz. *Ist f ein orthogonaler bzw. unitärer Endomorphismus über einem euklidischen bzw. unitären Vektorraum V, so gilt:*

(*1*) *Die Eigenwerte haben immer den Betrag Eins.*

(*2*) *Für alle $v,w\in V$ gilt $v\perp w$ genau dann, wenn $f(v)\perp f(w)$.*

(*3*) *f ist injektiv.*

(*4*) *Bei endlich erzeugten Vektorräumen ist f sogar ein Automorphismus.*

Beweis:

(*1*) Sei $f(v) = \lambda v$ für $v\neq 0$. Dann gilt

$$\varphi(v,v) = \varphi\big(f(v),f(v)\big) = \varphi(\lambda v,\lambda v) = \lambda\bar\lambda\varphi(v,v).$$

   Aus $v\neq 0$ und damit $\varphi(v,v)>0$ folgt die Behauptung $\lambda\bar\lambda = 1$.

(*2*) Die Behauptung erhalten wir sofort aus $\varphi(v,w) = \varphi\big(f(v),f(w)\big)$.

(*3*) Aus dem vorherigen Satz folgt $\mathrm{Kern}\,f = \{0\}$, d.h. f ist injektiv.

(*4*) Folgt mit (*3*) und der ersten Folgerung in 3.1.4.

□

6.4.2 Anwendungen in der Geometrie

Wir wollen nun ein paar geometrische Anwendungen betrachten. Dazu fassen wir den $\mathbf{R}^n$ als Vektorraum mit dem natürlichen Skalarprodukt auf.

Translationen

Definition. Sei a ein konstanter Vektor. Dann heißt die Abbbildung $f\colon \mathbf{R}^n \to \mathbf{R}^n$ mit $f(x) = x + a$ *Translation.*

Bewegungen und Isometrien

Definition. Eine bijektive Abbildung $f\colon \mathbf{R}^n \to \mathbf{R}^n$ mit $\langle f(x), f(y)\rangle = \langle x, y\rangle$ für alle x und y heißt *Bewegung* oder *Isometrie.*

Spiegelungen

Wir wollen hier nur Spiegelungen an *Hyperebenen* (siehe unten) untersuchen.

Definition. Die Menge

$$H_n := \left\{ x \in \mathbf{R}^n \;\middle|\; \langle a, x\rangle = 0 \right\}$$

für einen festen Vektor $a \neq 0$ heißt *Hyperebene.*

Für $n = 2$ stellt die Hyperebene die Gerade durch Null dar, die senkrecht zum Vektor a steht:

$$H_2\colon \quad a_1x_1 + a_2x_2 = 0.$$

Für $n = 3$ ist die Hyperebene die Ebene durch Null, die senkrecht zum Vektor a steht:

$$H_3\colon \quad a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

Allgemein gilt für die Hyperebene:

$$\operatorname{Dim} H_n = n - 1.$$

Definition. Die lineare Abbildung $s\colon \mathbf{R}^n \to \mathbf{R}^n$ mit

$$s(x) = x - 2\langle x, a\rangle a$$

für $\|a\| = 1$ heißt *Spiegelung.*

Wie man leicht nachrechnen kann hat die Darstellungsmatrix S einer Spiegelung folgende Form:

$$S = E - 2aa^T.$$

Damit gilt für die Determinante einer Spiegelung:

$$\det S = -1.$$

Drehungen

Definition. Die Elemente aus der speziellen orthogonale Gruppe SO heißen *Drehungen.*

Definitionsgemäß gilt $\det D = 1$ für jede Drehung D.
Da die Drehachse a einer Drehung D nicht verändert wird, kann man die Drehachse bestimmen, indem man den Eigenvektor zum Eigenwert 1 berechnet:

$$Da = a.$$

Ferner kann der Drehwinkel einer Drehung im $\mathbf{R}^3$ aus der Normalform (siehe Spektralsatz für komplexifizierte Endomorphismen, Abschnitt 6.10.4) berechnet werden:

$$B^TDB = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Dabei gilt:

$$1 + 2\cos\varphi = \operatorname{Sp}(B^TDB) = \operatorname{Sp} D.$$

Also braucht die Normalform der Drehmatrix nicht berechnet werden! Es gilt:

$$\cos\varphi = \frac{\operatorname{Sp} D - 1}{2}.$$

Orthogonale Matrizen im $\mathbf{R}^2$:

$$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \qquad \text{Drehung,}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{Spiegelung.}$$

6.5 Selbstadjungierte Endomorphismen

Definition.

- Wir nennen einen Endomorphismus *selbstadjungiert*, wenn

  $$f = f^*.$$

- Eine Matrix $A \in \mathrm{GL}(n, \mathbf{R})$ nennen wir *symmetrisch*, wenn

  $$A = A^T$$

  gilt. Eine Matrix $A \in \mathrm{GL}(n, \mathbf{C})$ heißt *hermitsch*, wenn

  $$A = \bar{A}^T.$$

Satz. *Ein Endomorphismus ist genau dann selbstadjungiert, wenn seine Darstellungsmatrix bzgl. einer Orthonormalbasis symmetrisch bzw. hermitsch ist.*

Satz. *Die Eigenwerte eines selbstadjungierten Endomorphismus über einem euklidischen oder unitären Vektorraum sind immer reell.*

Beweis: Sei f ein Endomorphismus und $f(u) = \lambda u$ mit $u \neq 0$. Dann gilt:

$$\lambda\varphi(u,u) = \varphi(\lambda u, u) = \varphi\big(f(u), u\big) = \varphi\big(u, f^*(u)\big) = \varphi\big(u, f(u)\big) = \varphi(u, \lambda u) = \bar{\lambda}\varphi(u,u).$$

Da $\varphi(u,u) \neq 0$ für $u \neq 0$ gilt, folgt daraus $\lambda = \bar{\lambda}$ und damit die Behauptung. □

Korollar. *Das charakteristische Polynom eines selbstadjungierten Endomorphismus zerfällt stets in Linearfaktoren.*

Satz. *Die Eigenräume zu paarweise verschiedenen Eigenwerten eines selbstadjungierten Endomorphismus sind zueinander orthogonal.*

Beweis: Seien $\lambda_1, \lambda_2, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte des Endomorphismus f und $V(\lambda_1), V(\lambda_2), \ldots, V(\lambda_r)$ die dazugehörigen Eigenräume. Wir zeigen nun, daß $V(\lambda_i)$ zu $V(\lambda_j)$ mit $1 \leq i, j \leq r$ und $i \neq j$ orthogonal ist: Seien dazu $v \in V(\lambda_i)$ und $w \in V(\lambda_j)$ beliebig. Damit gilt

$$\lambda_j\varphi(v,w) = \varphi(v, \lambda_j w) = \varphi\big(v, f(w)\big) = \varphi\big(f(v), w\big) = \varphi(\lambda_i v, w) = \lambda_i\varphi(v,w).$$

Daraus folgt $(\lambda_j - \lambda_i)\varphi(v,w) = 0$, d.h. $\varphi(v,w) = 0$, denn es gilt ja nach Voraussetzung $\lambda_i \neq \lambda_j$. Damit folgt die Behauptung. □

Satz von der Hauptachsentransformation, orthogonale Diagonalisierung. *Zu einer symmetrischen Matrix $A \in K^{n\times n}$ existiert eine Orthonormalbasis B des $\mathbf{R}^n$ aus Eigenvektoren mit*

$$B^T A B = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n).$$

Dabei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A.

Wir werden uns später in einem eigenen weiteren Abschnitt mit der Hauptachsentransformation beschäftigen. Dort werden dann auch Verfahren zur Bestimmung der Basen gegeben.

6.6 Schiefadjungierte Endomorphismen

Definition.

- Ein Endomorphismus f heißt *schiefadjungiert*, wenn
$$f^* = -f.$$
- Eine Matrix $A \in K^{n\times n}$ heißt *schiefsymmetrisch*, wenn
$$A = -A^T.$$

Satz. *Ein Endomorphismus ist genau dann schiefadjungiert, wenn seine Darstellungsmatrix schiefsymmetrisch ist.*

Satz. *Die Eigenwerte eines schiefadjungierten Endomorphismus sind stets Null oder rein imaginär.*

Beweis: Ist $f(u) = \lambda u$ mit $u \neq 0$, so gilt

$$\lambda\varphi(u,u) = \varphi(\lambda u, u) = \varphi\big(f(u), u\big) = \varphi\big(u, -f(u)\big) = -\bar{\lambda}\varphi(u,u)$$

und damit folgt aus $\lambda = -\bar{\lambda}$ die Behauptung. □

6.7 Normale Endomorphismen

Definition.

- Einen Endomorphismus f nennen wir *normal*, wenn
$$f \circ f^* = f^* \circ f.$$
- Analog heißt eine Matrix $A \in K^{n\times n}$ *normal*, wenn
$$\bar{A}^T A = A\bar{A}^T.$$

Bemerkung: Wie wir sofort einsehen, ist jeder orthogonale, unitäre, selbstadjungierte oder schiefadjungierte Endomorphismus normal.

Lemma 6.3: *Ist V ein euklidischer oder unitärer Vektorraum mit Skalarprodukt φ und der Endomorphismus f über V normal, so gilt für alle $u, v \in V \setminus \{0\}$*

(1) $f(u) = \lambda u \implies f^*(u) = \bar{\lambda} u$ *für alle* $\lambda \in K$,

(2) $f(u) = \lambda u$, $f(v) = \mu v$, $\lambda \neq \mu \implies u \perp v$ *für alle* $\lambda, \mu \in K$.

Beweis:

(1) Aus $\varphi\big(f(u), f(u)\big) = \varphi\big((f^* \circ f)(u), u\big) = \varphi\big((f \circ f^*)(u), u\big) = \varphi\big(f^*(u), f^*(u)\big)$ folgt $\mathrm{Kern}\, f = \mathrm{Kern}\, f^*$. Da f normal ist, ist auch $f - \lambda\,\mathrm{Id}$ normal und damit erhalten wir aus $\mathrm{Kern}(f - \lambda\,\mathrm{Id}) = \mathrm{Kern}(f^* - \bar{\lambda}\,\mathrm{Id})$ die Behauptung.

(2) Unter Verwendung des soeben erzielten Ergebnisses erhalten wir aus

$$\lambda\varphi(u,v) = \varphi(\lambda u, v) = \varphi\big(f(u), v\big) = \varphi\big(u, f^*(v)\big) = \varphi(u, \bar{\mu} v) = \mu\varphi(u,v)$$

sofort die Behauptung.

□

6.8 Orthogonale Trigonalisierung

Satz von Schur, orthogonale Trigonalisierung. *Sei (V, φ) ein Innenproduktraum über K.*

- *Seien f ein Endomorphismus über V und alle Eigenwerte von f in K. Dann existiert eine Orthonormalbasis $B = (u_1, \ldots, u_n)$ von V derart, daß gilt:*

$$ {}_B[f]_B = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}. $$

- *Seien $A \in K^{n\times n}$ und wiederum alle Eigenwerte in K. Dann gibt es eine orthogonale bzw. unitäre Matrix $B = (b^1, \ldots, b^n) \in GL(n, K)$ mit*

$$ \bar{B}^T A B = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}. $$

Dabei sind $\lambda_1, \ldots, \lambda_n$ jeweils die Eigenwerte von f bzw. A.

Beweis durch vollständige Induktion: Für $n = 1$ ist nichts zu beweisen. Nehmen wir nun an, die Behauptung gilt für $n-1$ (Induktionsvoraussetzung). Da alle Eigenwerte in K sind, gibt es einen Eigenwert λ_1 und einen dazugehörigen Eigenvektor v_1. Wir dürfen annehmen, daß v_1 normiert ist, d.h. $\varphi(v_1, v_1) = 1$. Nun ergänzen wir zu einer Orthonormalbasis $C := (v_1, \ldots, v_n)$. Sei $W := (Kv_1)^\perp = \mathrm{Lin}(v_2, \ldots, v_n)$. Dann gilt $V = Kv_1 \oplus W$ und für $w \in W$ ist $f(w) = \lambda_1 v_1 + g(w)$ mit $g(w) \in W$ und $g \in \mathrm{End}\, W$. Für die Darstellungsmatrix von f bezüglich C gilt dann:

$$ {}_C[f]_C = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & F_1 & \\ 0 & & & \end{pmatrix}. $$

Dabei ist $F_1 = {}_D[g]_D$ mit $D := (v_2, \ldots, v_n)$. Offensichtlich sind auch alle Eigenwerte von g in K. Nach Induktionsvoraussetzung existiert nun eine Orthonormalbasis $D' := (u_2, \ldots, u_n)$ von W, so daß

$$ {}_{D'}[g]_{D'} = \begin{pmatrix} \lambda_2 & * & \ldots & * \\ 0 & \lambda_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}. $$

Nun erkennen wir, daß $B := (v_1, u_2, \ldots, u_n)$ die gesuchte Basis ist. □

Folgerung: Sei (V, φ) ein Innenproduktraum über K und der Endomorphismus f orthogonal trigonalisierbar. Dann gilt

$$ {}_B[f^*]_B = {}_B\overline{[f]}_B{}^T = \begin{pmatrix} \bar{\lambda}_1 & 0 & \ldots & 0 \\ * & \bar{\lambda}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \ldots & * & \bar{\lambda}_n \end{pmatrix}. $$

Das bedeutet: Hat f alle Eigenwerte λ_i in K, so hat f^* die Eigenwerte $\bar{\lambda}_i$.

6.9 Orthogonale Diagonalisierung

Spektralsatz, orthogonale Diagonalisierung.

- *Sei V ein endlichdimensionaler K-Vektorraum und f ein normaler Endomorphismus. Ferner seien alle Eigenwerte von f in K. Dann existiert eine Orthonormalbasis B von V aus Eigenvektoren, d.h.:*

$$ {}_B[f]_B = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n). $$

- *Ist $A \in K^{n\times n}$ normal und liegen alle Eigenwerte von A in K, so existiert eine Orthonormalbasis B von V aus Eigenvektoren, d.h.*

$$ \bar{B}^T A B = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n). $$

Dabei sind $\lambda_1, \ldots, \lambda_n$ jeweils die Eigenwerte von f bzw. A.

Beweis: Wir beweisen hier nur den ersten Punkt, der Beweis des zweiten verläuft ganz analog. Zunächst erhalten wir aus einer beliebigen Basis von V durch die Schmidtsche Orthogonalisierung eine Orthogonalbasis. Durch anschließende Normierung ergibt sich eine Orthonormalbasis. Aufgrund des Satzes über die orthogonale Trigonalisierung wissen wir, daß die Abbildungsmatrix F von f bezüglich einer Orthonormalbasis obere Dreiecksgestalt hat, d.h. wir müssen nur noch zeigen, daß die Elemente oberhalb der Hauptdiagonalen alle Null sind. Aus

$$ \bar{F}^T F = F \bar{F}^T $$

folgt dies sofort. □

Vorgehensweise zur orthogonalen bzw. unitären Diagonalisierung:

1. Bestimmung aller Eigenwerte.
2. Bestimmung der dazugehörigen Eigenräume.
3. Zusammensetzen der Basen der Eigenräume zu einer Orthonormalbasis.

6.10 Komplexifizierung

Schon im 5. Kapitel haben wir gesehen, daß die Trigonalisierung eines Endomorphismus, dessen Eigenwerte nicht ausschließlich reell sind, über $\mathbf{R}$ nicht möglich ist. Jedoch wissen wir, daß bei Betrachtung über $\mathbf{C}$ *jeder* Endomorphismus trigonalisierbar ist. Im folgenden wollen wir nun zeigen, daß die Fortsetzung eines euklidischen Vektorraumes mit Skalarprodukt, sowie die Fortsetzung eines Endomorphismus zu einem unitären Innenproduktraum unter Beibehaltung der Eigenschaften möglich ist.

6.10.1 Fortsetzung des Vektorraumes

Sei V ein euklidischer Vektorraum, den wir nun über $\mathbf{C}$ betrachten wollen (Komplexifizierung). Dazu definieren wir einen $\mathbf{C}$-Vektorraum

$$\widetilde{V} := \{ (u,v) \mid u,v \in V \}$$

mit der additiven Verknüpfung

$$\widetilde{V} \times \widetilde{V} \to \widetilde{V}, \quad (u,v)+(u',v') \mapsto (u+u', v+v') \qquad \text{für alle } (u,v),(u',v') \in \widetilde{V}$$

und der Skalarmultiplikation

$$\mathbf{C} \times \widetilde{V} \to \widetilde{V}, \quad (\alpha+i\beta)(u,v) \mapsto (\alpha u - \beta v, \beta u + \alpha v).$$

Dabei beachten wir, daß $(u,v) = (u,0)+(0,v) = (u,0)+i(v,0)$ gilt, und identifizieren u mit $(u,0)$, sowie (u,v) mit $u+iv$.

Lemma 6.4: *Ist* $B = (v_1,\dots,v_n)$ *eine reelle Basis von* V*, so ist* B *auch eine* $\mathbf{C}$*-Basis von* $\widetilde{V}$*. Also gilt*

$$\mathrm{Dim}_{\mathbf{C}} \widetilde{V} = \mathrm{Dim}_{\mathbf{R}} V.$$

Beweis: Sei $u = \sum \alpha_k v_k$ und $v = \sum \beta_k v_k$. Dann ist $u+iv = \sum(\alpha_k + i\beta_k)v_k$, also ein Erzeugendensystem von $\widetilde{V}$.
Ist $\sum(\alpha_k+i\beta_k)v_k = 0 = \sum \alpha_k v_k + i\sum \beta_k v_k$, so gilt $\sum \alpha_k v_k = 0 = \sum \beta_k v_k$, also $\alpha_k + i\beta_k = 0$. □

6.10.2 Fortsetzung des Skalarproduktes

Wir definieren

$$\widetilde{\varphi}(u+iv, u'+iv') := \varphi(u,u') + \varphi(v,v') + i\big(\varphi(v,u') - \varphi(u,v')\big).$$

Nun wollen wir zeigen, daß wir auf einen komplexifizierten Innenproduktraum wirklich die Theorie bezüglich unitärer Vektorräume anwenden dürfen.

Satz. *Sei* V *ein euklidischer Vektorraum mit Skalarprodukt* φ*. Die Fortsetzung* $(\widetilde{V},\widetilde{\varphi})$ *dieses Innenproduktraums ist dann unitär. Insbesondere ist also* $\widetilde{\varphi}$ *ein komplexes inneres Produkt auf* $\widetilde{V}$*.*

Beweis: Wir verifizieren:

1. $\widetilde{\varphi}$ ist im ersten Argument $\mathbf{C}$-linear.
2. $\widetilde{\varphi}(u+iv, u'+iv') = \overline{\widetilde{\varphi}(u'+iv', u+iv)}$
3. φ ist positiv definit, d.h. $\widetilde{\varphi}(u+iv, u+iv) > 0$ falls $u+iv \neq 0$. Denn

$$\widetilde{\varphi}(u+iv, u+iv) = \varphi(u,u) + \varphi(v,v) > 0 \qquad \text{für } u \neq 0 \text{ oder } v \neq 0.$$

□

6.10.3 Fortsetzung des Endomorphismus

Sei f ein Endomorphismus über V. Dann betrachten wir über $\widetilde{V}$ den Endomorphismus $\widetilde{f}$ mit

$$\widetilde{f}(u+iv) = f(u) + if(v).$$

Satz. *Für die Komplexifizierung des Endomorphismus* f *mit Minimalpolynom* $m_f(\tau)$ *über* $\mathbf{R}$ *und reeller Basis* B *von* V *gilt:*

$$\begin{gathered} \widetilde{f} \text{ ist } \mathbf{C}\text{-linear} \\ {}_B[f]_B = {}_B[\widetilde{f}]_B, \\ m_f(\tau) = m_{\widetilde{f}}(\tau). \end{gathered}$$

Beweis: Wir verifizieren die Behauptung leicht anhand der Definition; etwa gilt für die erste Aussage:

$$\begin{aligned} \widetilde{f}\big((u+iv)+(u'+iv')\big) &= \widetilde{f}\big(u+u'+i(v+v')\big) = f(u+u') + if(v+v') \\ &= f(u) + f(u') + i\big(f(v)+f(v')\big) = f(u) + if(v) + f(u') + if(v') \\ &= \widetilde{f}(u+iv) + \widetilde{f}(u'+iv'). \end{aligned}$$

Ebenso folgt $f\big((\alpha+i\beta)(u+iv)\big) = (\alpha+i\beta)\widetilde{f}(u+iv)$. □

Folgerung: Das Minimalpolynom von $\widetilde{f}$ hat reelle Koeffizienten; insbesondere treten die Eigenwerte $\lambda \in \mathbf{C} \setminus \mathbf{R}$ von $\widetilde{f}$ paarweise $(\bar{\lambda}, \lambda)$ auf. Die reellen Eigenwerte von f und $\widetilde{f}$ stimmen überein.

Satz. *Sei* (V,φ) *ein euklidischer Vektorraum und* $f \in \mathrm{End}_{\mathbf{R}} V$*. Betrachten wir nun den* $\mathbf{C}$*-Vektorraum*

$$\widetilde{V} = \{ u+iv \mid u,v \in V \}$$

mit Skalarprodukt

$$\widetilde{\varphi}(u+iv, u'+iv') := \varphi(u+u') + i\varphi(v+v') + i\big(\varphi(v,u') - \varphi(u,v')\big) \quad \textit{und}$$
$$\widetilde{f}(u+iv) := f(u) + if(v),$$

so gilt

(1) $\widetilde{f}^* = \widetilde{f}^*$ *(* Adj. bzgl.* $\varphi, \widetilde{\varphi}$*).*

(*2*) f *ist genau dann normal, wenn* $\widetilde{f}$ *normal ist.*

(*3*) f *ist genau dann selbstadjungiert, wenn* $\widetilde{f}$ *selbstadjungiert ist.*

(*4*) f *ist genau dann schiefadjungiert, wenn* $\widetilde{f}$ *schiefadjungiert ist.*

(*5*) f *ist genau dann orthogonal, wenn* $\widetilde{f}$ *unitär ist.*

Beweis:

(*1*) Es gilt:

$$\begin{aligned}\widetilde{\varphi}\big(\widetilde{f}(u+iv),u'+iv'\big) &= \widetilde{\varphi}\big(f(u)+if(v),u'+iv'\big)\\ &= \varphi\big(f(u),u'\big)+\varphi\big(f(v),v'\big)+i\Big(\varphi\big(f(v),u'\big)-\varphi\big(f(u),v'\big)\Big)\\ &= \varphi\big(u,f^*(u')\big)+\varphi\big(v,f^*(v')\big)+i\Big(\varphi\big(v,f^*(u')\big)-\varphi\big(u,f^*(v')\big)\Big)\\ &= \widetilde{\varphi}\big(u+iv,f^*(u')+if^*(v')\big)\\ &= \widetilde{\varphi}\big(u+iv,\widetilde{f}^*(u'+iv')\big).\end{aligned}$$

Daraus folgt $\widetilde{f^*}=\widetilde{f}^*$.

(*2*)–(*6*) folgen aus (*1*), den Definitionen und der Formel $\widetilde{f\circ g}=\widetilde{f}\circ\widetilde{g}$.

□

6.10.4 Spektralsatz für komplexifizierte Endomorphismen

Um den Spektralsatz (orthogonale Diagonalisierung) auf normale Endomorphismen mit nicht ausschließlich reellen Eingenwerten über einem euklidischen Vektorraum anzuwenden, müssen wir komplexifizieren, denn eine Voraussetzung des Spektralsatzes ist, daß alle Eigenwerte in K liegen.
Wir formulieren nun eine neue Version des Spektralsatzes für normale Endomorphismen eines euklidischen Vektorraumes. Falls alle Eigenwerte reell sind, wird das Ergebnis natürlich mit dem aus Kapitel 6.9 übereinstimmen.

Satz. *Sei also* (V,φ) *euklidisch und* f *ein normaler Endomorphismus über* V. *Dann existiert stets eine Orthonormalbasis* B *von* V, *so daß*

$$ {}_B[f]_B = \begin{pmatrix} \lambda_1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & & & \\ 0 & & \lambda_r & & & \\ 0 & & & \alpha_1 & \beta_1 & \vdots \\ & & & -\beta_1 & \alpha_1 & \\ \vdots & & & & & 0 \\ & & & & \ddots & \\ 0 & & & 0 & & \alpha_s \quad \beta_s \\ & & & & & -\beta_s \quad \alpha_s \end{pmatrix}, $$

wobei $\lambda_1,\dots,\lambda_r$ *die reellen und* $\alpha_1+i\beta_1,\dots,\alpha_s+i\beta_s$ *die komplexen Eigenwerte von* f *sind.*

Beweis: Zunächst komplexifizieren wir (V,φ) und f wie in Kapitel 6.10.1 bis 6.10.3.
Wir wollen nun zeigen, daß $\widetilde{V}$ die orthogonale Summe der Eigenräume ist:

$$\widetilde{V}=\widetilde{V}(\lambda_1)\perp\widetilde{V}(\lambda_2)\perp\ldots\perp\widetilde{V}(\lambda_s)\perp\widetilde{V}(\lambda_{s+1})\perp\ldots\perp\widetilde{V}(\lambda_n).$$

Dazu müssen wir zwischen reellen und komplexen Eigenwerten unterscheiden:

1. Zum einen seien $\lambda_1,\dots,\lambda_r$ die *reellen* Eigenwerte und λ ein beliebiger davon mit Eigenvektor $u+iv$. Dann ist

   $$f(u)+if(v)=\widetilde{f}(u+iv)=\lambda(u+iv)=\lambda u+i\lambda v.$$

   Durch den Vergleich von Real- und Imaginärteil sehen wir, daß $f(u)=\lambda u$ und $f(v)=\lambda v$, d.h.

   $$\widetilde{V}(\lambda)=V(\lambda)+iV(\lambda)$$

   mit reellem Eigenraum $V(\lambda)$. Eine reelle Orthonormalbasis von $V(\lambda)$ ist damit auch eine Orthonormalbasis von $\widetilde{V}(\lambda)$.

2. Zum anderen seien $\alpha_1+i\beta_1,\dots,\alpha_s+i\beta_s$ die komplexen Eigenwerte ($\beta_i\neq 0$). Davon sei $\lambda=\alpha+i\beta$ ein beliebiger Eigenwert mit Eigenvektor $u+iv$. Damit gilt dann

   $$\begin{aligned}f(u)+if(v)=\widetilde{f}(u+iv)&=\lambda(u+iv)=(\alpha+i\beta)(u+iv)\\ &=\alpha u-\beta v+i(\beta u+\alpha v),\end{aligned}$$

   was genau dann gilt, wenn $f(u)=\alpha u-\beta v$ und $f(v)=\beta u+\alpha v$ gilt, also

   $$\widetilde{f}(u-iv)=\bar{\lambda}(u-iv),$$

   d.h. $\bar{\lambda}\neq\lambda$ ist ebenfalls Eigenwert. Somit erhalten wir die beiden zueinander orthogonalen Eigenräume $\widetilde{V}(\lambda)$ und $\widetilde{V}(\bar{\lambda})$.

V ist orthogonale Summe der reellen V_λ (s.o.) plus orthogonale Summe von Eigenräumen $\widetilde{V}_\lambda\perp\widetilde{V}_{\bar{\lambda}}$.
Orthonormalbasis $(w_1,\dots,w_s)$ in $\widetilde{V}_\lambda$, $w_j=a_j+ib_j$, $(\bar{w}_1,\dots,\bar{w}_s)$ in $\widetilde{V}_{\bar{\lambda}}$, $\bar{w}_j=a_j-ib_j$.

$$u_j:=\frac{1}{\sqrt{2}}(w_j+\bar{w}_j),\qquad v_j:=\frac{1}{\sqrt{2}i}(w_j-\bar{w}_j).$$

□

Beispiel für orthogonale Diagonalisierung (mit Komplexifizierung): Wir wollen nun die reelle und normale Matrix

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

orthogonal diagonalisieren. Dazu bestimmen wir zuerst das charakteristische Polynom $p(x)$ von A. Es gilt

$$p(x)=\big(x+1\big)\big(x-5-2i\sqrt{3}\big)\big(x-5+2i\sqrt{3}\big),$$

d.h. die Eigenwerte von A sind -1, $5+2i\sqrt{3}$ und $5-2i\sqrt{3}$. Da A eine reelle Matrix ist, jedoch nicht alle Eigenwerte reellwertig sind, wenden wir die letztere Ausführung des Spektralsatzes an. Wir bestimmen also zu dem reellen und zu den imaginären Eigenwerten die dazugehörigen Eigenvektoren: Nach einfacher Rechnung erhalten wir dafür

$$\frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\end{pmatrix},\quad \frac{1}{2\sqrt{3}}\begin{pmatrix}-1-i\sqrt{3}\\1-i\sqrt{3}\\2\end{pmatrix},\quad \frac{1}{2\sqrt{3}}\begin{pmatrix}-1+i\sqrt{3}\\1+i\sqrt{3}\\2\end{pmatrix}.$$

Dies führt uns schließlich zu folgender Basis

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

Mit dieser Basis gilt dann

$$B^TAB=\begin{pmatrix}-1&0&0\\0&5&-2\sqrt{3}\\0&2\sqrt{3}&5\end{pmatrix}.$$

$\Diamond$

Folgerungen:

Im folgenden wollen wir die Anwendung des Spektralsatzes auf einige Endomorphismen über euklidischen Vektorräumen betrachten. Dabei bezeichnen wir mit V einen euklidischen Vektorraum mit Skalarprodukt φ.
Aus Kapitel 6.5 wissen wir, daß die Eigenwerte eines *selbstadjungierten* Endomorphismus stets reell sind, d.h. wir können direkt den Spektralsatz in der Version des Kapitels 6.9 anwenden.
Hingegen sind die Eigenwerte eines *schiefadjungierten* Endomorphismus stets Null oder rein imaginär (Kapitel 6.6). Also gibt es eine Orthonormalbasis B von V, so daß die Darstellungsmatrix eines beliebigen *schiefadjungierten* Endomorphismus f folgende Gestalt hat:

$${}_B[f]_B=\begin{pmatrix}\boxed{\begin{matrix}0&&\\&\ddots&\\&&0\end{matrix}}&\boxed{0}&\cdots&\boxed{0}\\\boxed{0}&\boxed{\begin{matrix}0&\beta_1\\-\beta_1&0\end{matrix}}&&\vdots\\\vdots&&\ddots&\boxed{0}\\\boxed{0}&\cdots&\boxed{0}&\boxed{\begin{matrix}0&\beta_s\\-\beta_s&0\end{matrix}}\end{pmatrix}.$$

Damit gilt

$$\det f=\begin{cases}0 & \text{falls } 0 \text{ ein Eigenwert von } f \text{ ist,}\\ \beta_1^2\dots\beta_s^2 & \text{sonst.}\end{cases}$$

Betrachten wir einen *orthogonalen* Endomorphismus f, so wissen wir aus Kapitel 6.4.1, daß die Eigenwerte stets den Betrag Eins haben. Somit sind die reellen Eigenwerte ±1. Die komplexen Eigenwerte liegen auf dem Einheitskreis und für den Eigenwert λ gilt:

$$\lambda=\cos t+i\sin t=e^{it}.$$

Also gibt es zu jedem *orthogonalen* Endomorphismus eine Orthonormalbasis von V, so daß gilt:

$${}_B[f]_B=\begin{pmatrix}\boxed{\begin{matrix}1&&\\&\ddots&\\&&1\end{matrix}}&\boxed{0}&\cdots&\cdots&\boxed{0}\\\boxed{0}&\boxed{\begin{matrix}-1&&\\&\ddots&\\&&-1\end{matrix}}&&&\vdots\\\vdots&&\boxed{\begin{matrix}\cos t_1&\sin t_1\\-\sin t_1&\cos t_1\end{matrix}}&&\vdots\\\vdots&&&\ddots&\boxed{0}\\\boxed{0}&\cdots&\cdots&\boxed{0}&\boxed{\begin{matrix}\cos t_s&\sin t_s\\-\sin t_s&\cos t_s\end{matrix}}\end{pmatrix}.$$

6.11 Symmetrische Matrizen und quadratische Formen

6.11.1 Basiswechsel und Koordinatentransformation

Die Elemente eines Vektorraumes können wir einerseits als *Vektoren*, d.h. in der bisherigen Art und Weise, andererseits auch als *Punkte* auffassen. Damit können wir V auch *Punktmenge* nennen.

Definition. Sei P ein fester Punkt, der sogenannte *Ursprung*, des Vektorraums V:

- Zu dem Punkt $x\in V$ nennen wir dann $x-P$ *Ortsvektor* von x bezüglich P,
- $K=(p;B)$ *Koordinatensystem* mit *Ursprung* P und Basis B und
- falls $x-p=\sum\beta_i v_i$ heißen die β_i *Punktkoordinaten* von x im Koordinatensystem K. Dafür schreiben wir

$$x_K=(\beta_1,\dots,\beta_n).$$

Denn es gilt ja aufgrund der Definition der Koordinatenabbildung:

$$ {}_B[x-p] = {}_B[x] - {}_B[p] = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. $$

Nun wollen wir den Basiswechsel und die Koordinatentransformation für den Sonderfall $V = \mathbf{R}^n$ betrachten. Dabei fassen wir V einerseits als Vektorraum und andererseits als Punktraum auf. Die Betrachtung beliebiger Körper verläuft ganz analog.
Es seien die Basen $E = (e^1, \ldots, e^n)$ und $B = (b^1, \ldots, b^n)$ des $\mathbf{R}^n$, sowie die Koordinatensysteme $I = (0, E)$ und $K_P = K(P, B)$ mit $P \in \mathbf{R}^n$ gegeben.

Vektorkoordinaten. Ist nun $x \in \mathbf{R}^n$ ein beliebiger Vektor, so ist

$$ {}_E[x] = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} $$

die Koordinatendarstellung des Vektors x bezüglich der Basis E. Ferner ist x' durch $x = \sum_{i=1}^n x_i' b^i = Bx'$ eindeutig bestimmt und damit ist

$$ {}_B[x] = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = x' $$

die Koordinatendarstellung von x bezüglich B.

Also gilt für den *Basiswechsel* von E nach B die *Transformationsformel*

$$ x' = {}_B[x] = B^{-1}x $$

und die *Substitutionsformel*

$$ x = Bx' = B \cdot {}_B[x]. $$

Punktkoordinaten. Sei jetzt x ein Punkt aus dem $\mathbf{R}^n$ und $y := {}_B[x-p] = (y_1, \ldots, y_n)^T$. Dann gilt $x - p = Bx'$ und für die Darstellung des Punktes x im Koordinatensystem K_P:

$$ x_{K_P} = (y_1, \ldots, y_n). $$

Für den Übergang vom sogenannten *Inertialsystem* I zum Koordinatensystem K_P gilt also die *Transformationsformel*

$$ y = B^{-1}(x - P) $$

und die *Substitutionsformel*

$$ x = Bx' + P. $$

6.11.2 Quadratische Formen

Sei im folgenden $K = \mathbf{R}$ ein Innenproduktraum mit $\langle x, y \rangle = x^T y$ für $x, y \in \mathbf{R}$.
Ist eine Funktion $f\colon [-1,1] \to \mathbf{R}$ n-mal differenzierbar, dann ist f nach der Taylor-Entwicklung darstellbar durch

$$ f(x) = f(0) + f'(0) \cdot x + \frac{1}{2} f''(0) \cdot x^2 + \cdots + \frac{1}{n!} f^{(n)}(0) x^n + R(x), $$

mit einem Rest $R(x)$.
Ist $x \in \mathbf{R}^n$ so gilt für $f\colon D \to \mathbf{R}$ mit $0 \in D \subseteq \mathbf{R}^n$

$$ \begin{aligned} f(x) &= f(0) + a^T x + x^T A x + \text{Terme höherer Ordnung} + R(x) \\ &= f(0) + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n + \sum \alpha_{ij} x_i x_j + \cdots . \end{aligned} $$

Dabei ist $\alpha_i = \partial f / \partial x_i(0)$ und damit $A = \left(\partial^2 f / (\partial x_i \partial x_j)(0)\right)/2$.
Wir betrachten also $p(x) = \alpha + a^T x + \langle Ax, x \rangle$ mit $A = A^T$. Das ist ein quadratisches Polynom in den x_i.

$$ \sum_{i,j} \alpha_{ij} x_i x_j = a_{11} x_1^2 + \cdots + \alpha_{ij} x_i x_j + \alpha_{ji} x_i x_j + \cdots + \alpha_{nn} x_n^2 = a_{11} x_1^2 + \cdots + \sum (\alpha_{ij} + \alpha_{ji}) x_i x_j. $$

Zunächst betrachten wir den quadratischen Anteil:

$$ p(x) = \langle Ax, x \rangle = \sum_{i,j} \alpha_{ij} x_i x_j $$

Aufgrund der Theorie existiert eine Orthonormalbasis $B = (b^1, \ldots, b^n)$, so daß

$$ B^T A B = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n). $$

Durch Substitution mit $x = By$ erhalten wir

$$ p(x) = \langle ABy, By \rangle = \langle B^T A B y, y \rangle = \sum_i \lambda_i y_i^2 $$

(dabei sind die y_i rein quadratisch!).

6.11.3 Hauptachsentransformation

Hauptachsentransformation, orthogonale Diagonalisierung. *Sei A eine reelle und symmetrische Matrix, d.h. $A = A^T \in \mathbf{R}^{n \times n}$. Damit liegen alle Eigenwerte in $\mathbf{R}$ und es existiert eine Orthonormalbasis $(b^1, \ldots, b^n)$ aus Eigenvektoren mit $Ab^i = \lambda_i b^i$. Also ist*

$$ B^T A B = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}. $$

Vorgehen: Gegeben sei $A = A^T \in \mathbf{R}^{n\times n}$.

1. Aus $\det(xE - A) = x^n - (\operatorname{Sp} A)x^{n-1} + \cdots + (-1)^n \det A = 0$ berechnet man die Eigenwerte $\lambda_1, \ldots, \lambda_r$ ($\lambda_i \neq \lambda_j$ für $i \neq j$):

$$p_A(x) = \prod (x - \lambda_i)^{k_i}.$$

2. Zu jedem Eigenwert λ_i berechnet man den Eigenraum

$$V(\lambda_i) = \{\, x \in \mathbf{R}^n \mid (A - \lambda_i E)x = 0 \,\}.$$

Laut Theorie gilt $\operatorname{Dim} V(\lambda_i) = k_i$. Nun bestimmt man eine Basis C_i von $V(\lambda_i)$.

3. Orthonormiere die Basis zu B_i. Dabei sind die Eigenräume zu verschiedenen Eigenwerten orthogonal, da A symmetrisch ist.

4. Setze die B_i zu einer Orthonormalbasis B von $\mathbf{R}^n$ zusammen:

$$B = (b^1, \ldots, b^n).$$

Die Vektoren $b_1, \ldots, b_n$ heißen *Hauptachsen*.

6.11.4 Quadriken

Definition. Seien $A \in \mathbf{R}^n$ symmetrisch, $a \in \mathbf{R}^n$ und $\beta \in \mathbf{R}$ gegeben. Dann heißt die Menge

$$\{\, x \in \mathbf{R}^n \mid x^T Ax + a^T x + \beta = 0 \,\}$$

Quadrik oder auch *Hyperfläche zweiter Ordnung*.

Definition. Eine durch ein Polynom $p(x) = x^T Ax + a^T x + \beta$ gegebene Quadrik hat *Normalform* bzw. *Standardform*, wenn gilt:

(*1*) Der Teil $x^T Ax$ liegt rein quadratisch vor.

(*2*) Der Teil $a^T x + \beta$ kann durch affine Substitutionen nicht weiter verkürzt werden.

Die Transformation einer Quadrik auf Standardform:

1. Zunächst bestimmen wir für den quadratischen Teil $x^T Ax$ die Hauptachsen B wie in Abschnitt 6.11.3, so daß $\det B = 1$ (Rechtssystem).

2. Nun substituieren wir mit $x = By$ und erhalten so aus $p(x) = x^T Ax + a^T x + \beta = 0$ den Ausdruck

$$\tilde{p}(x) = p(By) = \sum \lambda_i y_i^2 + \sum \mu_i y_i + \gamma \qquad \text{mit } \mu_i, \gamma \in \mathbf{R}.$$

3. Jetzt ergänzen wir quadratisch mit

$$z_i := \begin{cases} y_i & \text{falls } \lambda_i = 0, \\ y_i + \mu_i/2\lambda_i & \text{sonst} \end{cases}$$

und substituieren mit $y = z - u$. Daraus erhalten wir die Standardform

$$\lambda_1 z_1^2 + \cdots + \lambda_k z_k^2 + \alpha_{k+1} z_{k+1} + \cdots + \alpha_n z_n + \eta = 0.$$

Somit haben wir insgesamt mit $x = By = B(z - u) = Bz - Bu$ substituiert, d.h. die Standardform erhalten wir im Koordinatensystem $(-Bu, B)$.

Bemerkung: Zur Gestalt der Quadriken im $\mathbf{R}^2$ und $\mathbf{R}^3$ verweisen wir auf *Höhere Mathematik 1, von Meyberg und Vachenauer*.

Beispiel für die Transformation einer Quadrik auf Standardform: Wir wollen nun ein geeignetes Koordinatensystem im $\mathbf{R}^3$ angeben, in dem folgende Quadrik Standardform hat:

$$q\colon x_1^2 + 2x_2^2 + x_3^2 + 8x_1x_2 + 4x_2x_3 + x_1 + 4x_2 + x_3 + 6 = 0.$$

Diese Darstellung ist offenbar äquivalent zu

$$q\colon x^T Ax + a^T x + \beta = 0$$

mit

$$A := \begin{pmatrix} 1 & 4 & 0 \\ 4 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \quad a = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix} \text{ und } \beta = 6.$$

Zuerst bestimmen wir die Hauptachsen von A: Für das charakteristische Polynom von A erhalten wir:

$$p(x) = x^3 - 4x^2 - 15x + 18 = (x + 3)(x - 1)(x - 6),$$

d.h. die Eigenwerte von A sind -3, 1 und 6. Damit ergeben sich folgende Eigenvektoren:

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 2 \end{pmatrix}.$$

Jetzt erhalten wir durch Orthonormierung die Hauptachsen

$$B := \frac{\sqrt{5}}{15} \begin{pmatrix} -3 & 2\sqrt{5} & 4 \\ 0 & -2\sqrt{5} & 5 \\ 6 & \sqrt{5} & 2 \end{pmatrix}.$$

Nun substituieren wir mit $x = By$:

$$q\colon y_1^2 - 3y_2^2 + 6y_3^2 + \frac{\sqrt{5}}{5} y_1 - \frac{5}{3} y_2 + \frac{26\sqrt{5}}{15} y_3 + 6 = 0$$

und erhalten nach quadratischer Ergänzung:

$$q\colon \left(y_1 + \frac{\sqrt{5}}{10}\right)^2 - 3\left(y_2 + \frac{5}{18}\right)^2 + 6\left(y_3 + \frac{13\sqrt{5}}{90}\right)^2 + \frac{50}{9} = 0.$$

Mit der Substitution

$$y = z - u = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} - \begin{pmatrix} \sqrt{5}/10 \\ 5/18 \\ 13\sqrt{5}/90 \end{pmatrix}$$

und anschließender Multiplikation mit $9/50$ erhalten wir

$$q\colon \frac{9}{50}z_1^2 - \frac{27}{50}z_2^2 + \frac{27}{25}z_3^2 + 1 = 0.$$

Das ist die *Standardform* der Quadrik q im Koordinatensystem

$$(-Bu, B) = \left(\left(-\frac{5}{18}, -\frac{1}{18}, -\frac{7}{18} \right), \frac{\sqrt{5}}{15} \begin{pmatrix} -3 & 2\sqrt{5} & 4 \\ 0 & -2\sqrt{5} & 5 \\ 6 & \sqrt{5} & 2 \end{pmatrix} \right).$$

Sie stellt ein zweischaliges Hyperboloid dar. ◊

6.11.5 Nichtorthogonale Diagonalisierung symmetrischer Matrizen

In den vorherigen Kapiteln haben wir Matrizen stets mit Hilfe einer Basis aus Eigenvektoren diagonalisiert. In der Praxis ist dabei insbesondere die Bestimmung der Eigenwerte bei größeren Matrizen recht aufwendig. Nun wollen wir ein Diagonalisierungsverfahren betrachten, bei dem wir keine Eigenwerte bestimmen brauchen.
Erhalten wir durch endlich viele elementare *Zeilenumformungen* aus einer Matrix A die Matrix $\widetilde{A}$ und aus der Einheitsmatrix E die Matrix $\widetilde{E}$, so gilt:

$$\widetilde{A} = \widetilde{E}A.$$

Setzen wir $S := \widetilde{E}$, so gilt mit analogen *Spaltenumformungen*:

$$A \rightsquigarrow AS^T.$$

Nehmen wir an einer symmetrischen Matrix gleichzeitig endlich viele elementare Zeilen- und analoge elementare Spaltenumformungen vor, so gilt

$$A^T = A \rightsquigarrow SAS^T.$$

Satz. *Sei $A = A^T \in \mathbf{R}^{n\times n}$. Die Matrix A läßt sich durch endlich viele elementare Zeilen- und analoge elementare Spaltenumformungen auf Diagonalform bringen, d.h. es existiert ein $W \in \mathrm{GL}(n, \mathbf{R})$ mit*

$$W^TAW = \mathrm{Diag}(\alpha_1, \ldots, \alpha_n).$$

Beweis: Sei $A = (\alpha_{ij})_{n\times n}$.

Fall 1: $\alpha_{11} \neq 0$.

$$A \rightsquigarrow \begin{pmatrix} \alpha_{11} & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

Fall 2: $\alpha_{11} = 0$, aber $\alpha_{ii} \neq 0$ für ein $i \in \{2, \ldots, n\}$. Vertauschungen (Zeilen bzw. Spalten-Umformungen)

$$z_1 \leftrightarrow z_i, \qquad s_1 \leftrightarrow s_i$$

führen zum Fall 1.

Fall 3: $\alpha_{ii} = 0$ für alle i und $\alpha_{ij} \neq 0$ für ein Paar $i \neq j$. Wir erhalten

$$A = \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & \alpha_{ij} & & \\ & & & \ddots & & & \\ & & \alpha_{ij} & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}.$$

Dabei führt

$$z_i \to z_i + z_j, \qquad s_i \to s_i + s_j$$

zum Fall 2 mit $\alpha'_{ii} = 2\alpha_{ij}$.

□

Beachte: Im allgemeinen ist W nicht orthogonal und damit sind die α_i's keine Eigenwerte von A!

Eine Anwendung des Satzes ist der *Positivitätstest* (vgl. Kapitel 6.11.6): $A > 0 \iff WAW^T > 0 \iff$ alle $\alpha_{ii} > 0$ (mit dem W der nichtorthogonalen Diagonalisierung).

Eine weitere Anwendung ist die Signatur einer Matrix:

Definition. Sei $A = A^T \in \mathbf{R}^{n\times n}$, p die Anzahl der positiven Eigenwerte, q die Anzahl der negativen Eigenwerte und s die Anzahl der Eigenwerte gleich 0 (gleich der Dimension des Kerns von A!). Dann heißt das Tripel

$$(p, q, s)$$

die *Signatur* (oder auch Trägheit) von A.

Trägheitssatz von Sylvester. *Sei $A \in \mathbf{R}^{n\times n}$ eine symmetrische Matrix. Dann haben die Matrizen A und WAW^T mit $W \in \mathrm{GL}(n, \mathbf{R})$ die gleiche Signatur.*

Beweis: Wenn $B^T = B^{-1}$ orthogonal, haben A und B^TAB dieselben Eigenwerte. Sei deshalb o.B.d.A.

$$A = \mathrm{Diag}(\underbrace{\lambda_1, \ldots, \lambda_p}_{EW>0}, \underbrace{-|\lambda_{p+1}|, \ldots, -|\lambda_{p+q}|}_{EW<0}, 0, \ldots, 0) \qquad \text{und}$$
$$WAW^T = \mathrm{Diag}(\underbrace{\mu_1, \ldots, \mu_t}_{EW>0}, \underbrace{-|\mu_{t+1}|, \ldots, -|\mu_{t+r}|}_{EW<0}, 0, \ldots, 0).$$

Betrachte nun für $x = (x_1, \ldots, x_n)^T$ und $x = W^T y$ das Skalarprodukt

$$\begin{aligned}\langle Ax, x\rangle &= \lambda_1 x_1^2 + \cdots + \lambda_p x_p^2 - |\lambda_{p+1}| x_{p+1}^2 - \cdots - |\lambda_{p+q}| x_{p+q}^2 \qquad (6.1)\\ &= \langle AW^T y, W^T y\rangle = \langle WAW^T y, y\rangle\\ &= \mu_1 y_1^2 + \cdots + \mu_p y_t^2 - |\mu_t| y_t^2 - \cdots - |\mu_{t+r}| y_{t+r}^2 \qquad (6.2)\end{aligned}$$

und dazu das homogene lineare Gleichungssystem

$$\begin{aligned} x_{p+1} &= 0\\ &\vdots\\ x_{p+q} &= 0\\ y_1 &= y_1(x_1, \ldots, x_n) = 0\\ &\vdots\\ y_t &= y_t(x_1, \ldots, x_n) = 0\end{aligned}$$

mit n Unbekannten $x_1, \ldots, x_n$ und $q + t$ Gleichungen.
Annahme: Sei $t < p$. Dann folgt $q + t < q + p = \operatorname{Rang} A \le n$ und das Gleichungssystem hat eine nichttriviale Lösung. Mit dieser Lösung ergibt sich oben ein Wert größer oder gleich Null, in der dritten Zeile der Gleichung ein Wert kleiner oder gleich Null. Insgesamt kommt also nur der Wert Null in Frage. Also gilt für diesen Lösungsvektor neben $x_{p+1} = \cdots = x_{p+q} = 0$ auch $\lambda_1 x_1^2 + \cdots + \lambda_p x_p^2 = 0$ (also $x_1 = \cdots = x_p = 0$) und damit eine Widerspruch zu $x \neq 0$. Aus Symmetriegründen folgt insgesamt $t = p$. Da die Anzahl der Nullen gleich dem Rang ist, gilt:

$$\operatorname{Rang} A = \operatorname{Rang} WAW^T.$$

□

Beispiel: Sei

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

Dann erhalten wir durch elementare Zeilen- und analoge Spaltenumformungen die Matrix

$$\widetilde{A} = \operatorname{Diag}(1, -4, 3, -16),$$

also hat A die Signatur $(2, 2, 0)$. ◊

6.11.6 Positive Definitheit von Matrizen

In diesem Abschnitt betrachten wir nur Matrizen mit Koeffizienten in $\mathbf{R}$. Die Vorgehensweise für komplexwertige Matrizen ist analog.

Definition. Eine Matrix $A \in \mathbf{R}^{n\times n}$ heißt genau dann *positiv definit*, wenn

$$A = A^T \quad \text{und} \quad \langle Ax, x\rangle = x^T A^T x > 0 \quad \text{für alle } x \in \mathbf{R}^n \setminus \{0\}$$

gilt. Wir schreiben dafür dann auch $A > 0$.
Ferner heißt eine Matrix $A \in \mathbf{R}^{n\times n}$ heißt genau dann *positiv semidefinit*, wenn

$$A = A^T \quad \text{und} \quad \langle Ax, x\rangle = x^T A^T x \ge 0 \quad \text{für alle } x \in \mathbf{R}^n$$

gilt. Wir schreiben dafür dann auch $A \ge 0$.

Sei $A = (\alpha_{ij})_{n\times n}$. Dann gilt

$$\begin{aligned}(Ax)^T x &= \left(\sum \alpha_{1j} x_j, \ldots, \sum \alpha_{nj} x_j\right) x\\ &= \sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} x_i x_j\right)\\ &= \sum_{i,j=1}^n (\alpha_{ij} x_i x_j).\end{aligned}$$

Charakterisierung positiv definiter Matrizen. *Für Matrizen $A = A^T \in \mathbf{R}^{n\times n}$ sind folgende Aussagen äquivalent:*

(1) *A ist positiv definit.*

(2) *Für alle $W \in \mathrm{GL}(n, \mathbf{R})$ gilt $W^T AW > 0$.*

(3) *Es gibt $B \in \mathrm{GL}(n, \mathbf{R})$ mit $A = B^T B$.*

(4) *Alle Eigenwerte von A sind positiv.*

(5) *Es gibt ein $B = B^T \in \mathbf{R}^{n\times n}$ mit $A = B^2$.*

Beweis:

(1)$\Longrightarrow$*(2)*: Es gilt $(W^T AW)^T = W^T AW$ und damit ist für $y = Wx \neq 0$ und alle $x \neq 0$

$$\langle W^T AWx, x\rangle = \langle AWx, Wx\rangle = \langle Ay, y\rangle > 0.$$

(2)$\Longrightarrow$*(1)*: Setzen wir $W = E$, so erhalten wir sofort die erste Aussage.

(1)$\Longleftrightarrow$*(4)*: Aus der Hauptachsentransformation von A erhalten wir ein $B \in \mathrm{GL}(n, \mathbf{R})$ mit $B^T = B^{-1}$, so daß $B^T AB$ Diagonalgestalt hat und die Eigenwerte $\lambda_1, \ldots, \lambda_n$ auf der Hauptdiagonalen stehen. Aufgrund der Äquivalenz der ersten Aussage ist A genau dann positiv definit, wenn $D := B^T AB$ positiv definit. Andererseits ist $D > 0$ genau dann, wenn $\langle Dx, x\rangle = \sum \lambda_i x_i^2 > 0$ für alle $x \neq 0$.

(1)$\Longleftrightarrow$*(5)*: Sei $B^T AB$ in Diagonalform mit den positiven Eigenwerten (von A) $\lambda_1, \ldots, \lambda_n$ auf der Hauptdiagonale. Es gibt dann μ_i mit $\mu_1^2 = \lambda_1, \ldots, \mu_n^2 = \lambda_n$. Mit $M = \operatorname{Diag}(\mu_1, \ldots, \mu_n)$ gilt dann $B^T AB = M^2$, d.h. $A = (BM^2 B^T)$.

□

Jacobi-Test (Hurwitz-Test). *Sei $A = A^T \in \mathbf{R}^{n\times n}$. Dann ist A genau dann positiv definit, wenn alle Hauptdeterminanten positiv sind, also*

$$\det \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1i} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ii} \end{pmatrix} > 0 \quad \textit{für alle } 1 \le i \le n.$$

Satz. *Wenn A positiv definit ist, so sind alle Werte auf der Hauptdiagonalen positiv.*

Beweis: Folgt aus der Definition mit $x = e^i$, denn wenn $\langle Ax, x\rangle > 0$ für alle $x \neq 0$, so gilt das insbesondere für $x = e^i$ und damit $\langle Ae^i, e^i\rangle = \alpha_{ii} > 0$. □

Kapitel 7

Die Jordansche Normalform

7.1 Formale Polynome

Bisher haben wir Polynomfunktionen

$$f\colon K \to K, \quad f(x) = \alpha_n x^n + \cdots + \alpha_1 x + \alpha_0$$

mit $|K| = \infty$ betrachtet. Dabei war Koeffizientenvergleich möglich. Dies ist bei endlichen Körpern hingegen im allgemeinen nicht der Fall. Betrachte z.B. $f(x) = x^2 + x$ über $\mathbf{Z}_2$, wo dieses Polynom identisch zum Nullpolynom ist. Damit sind endliche Körper für Polynomfunktionen ungeeignet.
Wir definieren uns deshalb

$$K[\tau] := \left\{ (\alpha_i)_{i\ge 0} \;\middle|\; \alpha_i \in K \quad \text{und} \quad \alpha_i = 0 \text{ für fast alle } i \right\}.$$

Dabei heißt „für fast alle i“, daß die Bedingung $\alpha_i = 0$ ab einem hinreichend großem Wert i immer erfüllt ist. Durch

$$(\alpha_i)_{i\ge 0} + (\beta_i)_{i\ge 0} := (\alpha_i + \beta_i)_{i\ge 0}$$

ist die Addition und durch

$$(\alpha_i)_{i\ge 0} \cdot (\beta_i)_{i\ge 0} := (\gamma_i)_{i\ge 0} \quad \text{mit } \gamma_i = \sum_{k=0}^{i} \alpha_k \beta_{i-k}$$

die Multiplikation erklärt.
Für $\tau = (0, 1, 0, 0, \dots)$ gilt

$$\begin{aligned} \tau^0 &= (1, 0, 0, \dots) \\ \tau^1 &= \tau \\ \tau^2 &= (0, 0, 1, 0, 0, \dots) \\ &\vdots \\ \tau^n &= (\underbrace{0, \dots, 0}_{n \text{ Nullen}}, 1, 0, \dots). \end{aligned}$$

Damit hat jedes Polynom $(a_i)_{i\ge 0} \in K[\tau]$ die Darstellung

$$p(\tau) = \sum_{i\ge 0} a_i \tau^i, \qquad \text{fast alle } a_i = 0.$$

Definition. Den Ring bzw. Vektorraum der *formalen Polynome* über K mit τ^0 als Eins definieren wir durch

$$K[\tau] := \mathrm{Lin}\big((\tau^i)_{i\in\mathbf{N}_0}\big).$$

Für $p \in K[\tau]$ schreiben wir

$$p := \sum_{i=0}^{\infty} \alpha_i \tau^i = (\alpha_0, \alpha_1, \dots) \in K^{(\mathbf{N}_0)},$$

wobei fast alle $\alpha_i = 0$ sind.
Ein konstantes Polynom hat die Form $k \cdot \tau^0$ mit $k \in K$.

Bemerkung: Der Ring der formalen Polynome ist wohldefiniert, da aus der linearen Unabhängigkeit der τ^i Fälle wie $x^2 + x \equiv 0$ nicht auftreten können.

Definition. Für $p \neq 0$ definieren wir den *Grad* von p durch

$$\mathrm{Grad}\, p := \max\big\{\, i \in \mathbf{N}_0 \;\big|\; \alpha_i \neq 0 \,\big\},$$

für $p = 0$ sei der Grad

$$\mathrm{Grad}\, 0 := -\infty.$$

Lemma 7.1: *Für $p, q \in K[\tau]$ gilt:*

(1) *Ist $p \cdot q = 0$, so ist p oder q gleich Null,*

(2) $\mathrm{Grad}(pq) = \mathrm{Grad}\, p + \mathrm{Grad}\, q$,

(3) $\mathrm{Grad}(p+q) \leq \max\{\mathrm{Grad}\, p, \mathrm{Grad}\, q\}$.

Beweis: Seien $p = \sum_{i=0}^{m} \alpha_i \tau^i$ und $q = \sum_{i=0}^{n} \beta_i \tau^i$ mit $\alpha_m, \beta_n \neq 0$. Dann folgen *(1)* und *(2)* aus

$$pq = \underbrace{\alpha_m \beta_m}_{\neq 0} \tau^m \cdot \tau^n + \text{ Terme niedrigerer Ordnung.}$$

□

Lemma 7.2: *Seien wieder $p, q \in K[\tau]$ und $\lambda \in K$. Dann gilt*

(1) $(p+q)(\lambda) = p(\lambda) + q(\lambda)$

(2) $(pq)(\lambda) = p(\lambda) \cdot q(\lambda)$

(3) $(\alpha p)(\lambda) = \alpha \cdot \big(p(\lambda)\big)$

Bemerkung: Für Polynomfunktionen wie $p(x) = x + 1$ und $q(x) = x$ gilt im allgemeinen nicht die Nullteilerfreiheit, denn z.B. ist $pq = 0$ über $\mathbf{Z}_2$.

Definition. Gibt es zu $p, q \in K[\tau]$ ein $r \in K[\tau]$ mit $pr = q$, so sagen wir „p teilt q“. In Zeichen:

$$p \backslash q.$$

Teilen mit Rest. *Sind $p, q \in K[\tau]$ und $q \neq 0$, so gibt es $a, r \in K[\tau]$ mit*

$$p = a \cdot q + r \qquad \mathrm{Grad}\, r < \mathrm{Grad}\, q.$$

Beweis: Ist $\mathrm{Grad}\, q > \mathrm{Grad}\, p$, so erfüllen $r := p$ und $a = 0$ die Forderung.
Für $n = \mathrm{Grad}\, q \leq \mathrm{Grad}\, p = m$ erhalten wir mit

$$f_1 := p - \alpha_m \beta_n^{-1} \tau^{m-n} q$$

einerseits $\mathrm{Grad}\, f_1 < \mathrm{Grad}\, p$ und andererseits $f_1 = a_1 \cdot q + r_1$ mit $\mathrm{Grad}\, r_1 < \mathrm{Grad}\, a_1$ und so über Induktion die Behauptung. □

Definition. Ist $p = \sum_{i=0}^{\infty} \alpha_i \tau^i \in K[\tau]$ und $\lambda \in K$, so verstehen wir unter

$$p(\lambda) = \sum_{i=0}^{\infty} \alpha_i \lambda^i$$

die *Einsetzung* von λ in p.

Lemma 7.3: *Ist $p = \sum \alpha_i \tau^i$ und $q = \sum \beta_i \tau^i$, so gilt $p = q$ genau dann, wenn $\alpha_i = \beta_i$ für alle i.*

Beweis: Es gilt $p - q = \sum(\alpha_i - \beta_i)\tau^i = 0$ genau dann, wenn $\alpha_i - \beta_i = 0$ für alle i. □

Definition. Als *Nullstelle* oder *Wurzel* von $p \in K[\tau]$ bezeichnen wir $\alpha \in K$, wenn $p(\alpha) = 0$.

Satz. *Ist $p \in K[\tau]$ und $\alpha \in K$ eine Wurzel von p, so gibt es $q \in K[\tau]$ derart, daß*

$$p = (\tau - \alpha)q.$$

Beweis: Aus $p = q(\tau-\alpha)+r$ und $\mathrm{Grad}\, r < \mathrm{Grad}(\tau-\alpha) = 1$, sowie $0 = p(\alpha) = q(\alpha)(\alpha-\alpha)+r(\alpha)$ folgt $r(\alpha) = 0$, d.h. $r = 0$. □

Korollar. *Ist $p \in K[\tau]$ und sind $\alpha_1, \dots, \alpha_r$ Wurzeln der Vielfachheit $n_1, \dots, n_r$, so gibt es ein $q \in K[\tau]$ mit $q(\alpha_i) \neq 0$ für $1 \leq i \leq r$, so daß*

$$p = (\tau - \alpha_1)^{n_1}(\tau - \alpha_2)^{n_2} \dots (\tau - \alpha_n)^{n_r} \cdot q.$$

Korollar. *Jedes formale Polynom über K hat in K höchstens soviele Nullstellen, wie sein Grad ist.*

Der größte gemeinsame Teiler

Definition. Seien $f, g \in K[\tau]$. Dann heißt das Polynom $d \in K[\tau]$ *größter gemeinsamer Teiler von f und g*, falls gilt:

(1) $d \backslash f$ und $d \backslash g$,

(2) aus $q \in K[\tau]$, $q \backslash f$ und $q \backslash g$ folgt $q \backslash d$.

Wir stellen fest, daß der größte gemeinsame Teiler bis auf einen konstanten Faktor eindeutig bestimmt ist: Seien $d, d' \in K[\tau]$ zwei größte gemeinsame Teiler. Dann folgt aus der zweiten Bedingung $d \backslash d'$ und $d' \backslash d$ und damit gibt es ein $\alpha \in K$ mit $d = \alpha d'$.

Definition. Ein Polynom $f \in K[\tau]$ heißt *normiert*, wenn der höchste Koeffizient gleich Eins ist.

Im folgenden verstehen wir unter dem „größten gemeinsamen Teiler" immer den „normierten größten gemeinsamen Teiler".

Satz. *Zu $f, g \in K[\tau]$, nicht beide gleich Null, existieren Polynome $h_1, h_2 \in K[\tau]$ mit*

$$ggT(f, g) = h_1 f + h_2 g.$$

Beweis: Betrachte das durch f und g erzeugte Ideal $I = (f, g)$. Da jedes Ideal ein Hauptideal ist (siehe Satz oben), gibt es ein $d \in I$ mit $fK[\tau] + gK[\tau] = (d) = dK[\tau]$. (Offensichtlich erfüllt $d = ggT(f, g)$ dies.) Da $f \in I$ und $g \in I$ gibt es Polynome $s_1, s_2 \in K[\tau]$ mit $f = s_1 d$ und $g = s_2 d$. Also gibt es zu d die Polynome h_1, h_2 mit $d = h_1 f + h_2 g$. □

Analog zur Definition oben kann auch für eine Familie von Polynomen ein größter gemeinsamer Teiler erklärt werden:

Definition. Seien $p_1, \ldots, p_n \in K[\tau]$. Dann heißt das Polynom $d \in K[\tau]$ *größter gemeinsamer Teiler der* p_i, wenn gilt:

(*1*) $d \backslash p_i$ für alle i,

(*2*) aus $q \backslash p_i$ für alle i folgt $q \backslash d$.

Analog zum Satz oben gibt es Polynome $h_i \in K[\tau]$ mit

$$ggT(p_1, \ldots, p_n) = h_1 p_1 + \cdots + h_n p_n.$$

Definition. Die Polynome $p_1, \ldots, p_n$ heißen genau dann *teilerfremd*, wenn ihr größter gemeinsamer Teiler gleich Eins ist.

Satz. *Seien $p_1, \ldots, p_n \in K[\tau]$ gegeben. Dann sind diese Polynome genau dann teilerfremd, wenn es $h_i \in K[\tau]$ gibt, mit*

$$h_1 p_1 + \cdots + h_n p_n = 1.$$

Beweis: Sind die Polynome teilerfremd, so ist ihr größter gemeinsamer Teiler gleich Eins. Also gibt es diese h_i's.
Gibt es diese h_i's, so teilt jeder gemeinsame Teiler der h_i's auch 1. □

Definition. Das nichtkonstante Polynom $p \in K[\tau] \setminus K$ heißt *irreduzibel*, wenn aus $p = gf$ folgt, daß entweder f oder g aus K ist.

Satz. *Seien $f, g, p \in K[\tau]$ gegeben. Dann ist das Polynom p irreduzibel, falls aus $p \backslash fg$ folgt, daß $p \backslash f$ oder $p \backslash g$ gilt.*

Beweis: Angenommen es gilt $p \backslash f$. Dann gilt $ggT(p, f) = 1$ und nach dem Satz gibt es $h_1, h_2 \in K[\tau]$ mit $1 = h_1 p + h_2 f$. Mit g multipliziert ergibt dies $g = h_1 pg + h_2 fg$. Da aber $p \backslash h_1 pg$ und nach Voraussetzung $p \backslash h_2 fg$ gilt, folgt $p \backslash g$. □

Zerlegung in irreduzible Polynome

Satz. *Jedes Polynom $f \in K[\tau] \setminus K$ hat bis auf Reihenfolge und konstante Faktoren eine eindeutige Darstellung in folgender Form:*

$$f = p_1^{k_1} \cdot p_2^{k_2} \ldots p_r^{k_r},$$

wobei die p_i irreduzible Polynome sind.

Beweis: *Existenz.* Induktion nach $\operatorname{Grad} f = n$. Für $n = 1$ folgt, daß $\alpha\tau + \beta$ mit $\alpha \neq 0$ irreduzibel ist.
Induktionsschritt von $k < n$ auf n: Sei $\operatorname{Grad} f = n$. Dann ist f entweder irreduzibel mit $f = p_1$ oder es gibt Polynome $g, h \in K[\tau]$ mit $f = gh$ und $\operatorname{Grad} g < \operatorname{Grad} f$, $\operatorname{Grad} h < \operatorname{Grad} f$. Aus der Induktionsvoraussetzung folgt nun die Existenz der Zerlegung für f.

Eindeutigkeit. Sei $f = p_1 \ldots p_s = q_1 \ldots q_t$ eine Zerlegung in irreduzible Polynome p_i, q_i mit $s \leq t$. Dann gilt $p_1 \backslash \prod q_i$ und daraus folgt, daß es ein $i \in \{1, \ldots, t\}$ gibt, mit $p_1 \backslash q_i$. Sei nun o.B.d.A. $i = 1$, also $p_1 \backslash q_1$ oder auch $q_1 = h_1 \cdot p_1$ für ein $h_1 \in K[\tau]$. Also gilt:

$$p_1 \cdot p_2 \ldots p_s = h_1 \cdot p_1 \cdot q_2 \ldots q_t,$$

oder auch

$$p_2 \ldots p_s = h_1 \cdot q_2 \ldots q_t.$$

Führen wir diesen Schritt mehrmals aus, so erhalten wir schließlich

$$1 = h_1 \ldots h_s \cdot q_{s+1} \ldots q_t$$

und daraus folgt nach Gradvergleich $s = t$ und $h_i \in K$. Also ist die Zerlegung bis auf die Reihenfolge und bis auf konstante Faktoren eindeutig. □

Beispiele: Allgemein gilt: $\tau^9 - 1 = (\tau - 1)(\tau^8 + \tau^7 + \cdots + \tau + 1)$.
Speziell gilt über dem Körper $\mathbf{Q}$: $\tau^9 - 1 = (\tau - 1)(\tau^2 + \tau + 1)(\tau^6 + \tau^3 + 1)$.
Über $\mathbf{C}$ gibt es nur irreduzible Polynome vom Grad 1. Sie haben also alle die Form

$$f(\tau) = c \cdot \prod (\tau - \alpha_i)^{k_i}.$$

◊

Definition. Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes echte Polynom $f \in K[\tau] \setminus K$ in Linearfaktoren zerfällt.

7.2 Primärzerlegung

Im folgenden bezeichnen wir mit f einen Endomorphismus über einem endlicherzeugten K-Vektorraum V.
Sei $p \in K[\tau]$ ein Polynom über K mit der Eigenschaft $p(f) = 0$ (also z.B. das charakteristische Polynom von f). Es besitze die faktorisierte Form

$$p(\tau) = \prod_{i=1}^{r} p_i(\tau)^{k_i}.$$

Wir definieren dann weitere Polynome $q_i(\tau)$ durch

$$q_i := \frac{p}{p_i^{k_i}} = p_1^{k_1} \dots p_{i-1}^{k_{i-1}} \cdot p_{i+1}^{k_{i+1}} \dots p_r^{k_r}.$$

Da die Polynome q_i nach Definition teilerfremd sind, gibt es Polynome $h_1, \dots, h_r$, mit

$$\sum_{i=1}^{r} h_i(\tau) q_i(\tau) = 1.$$

Definition. Sei $p(\tau)$ ein allgemeines Polynom vom Grad n,

$$p(\tau) = \alpha_n \tau^n + \alpha_{n-1} \tau^{n-1} + \dots + \alpha_1 \tau + \alpha_0.$$

Dann heißt die Abbildung $K[\tau] \to \operatorname{End} V$ mit

$$p(\tau) \mapsto p(f) = \alpha_n f^n + \alpha_{n-1} f^{n-1} + \dots + \alpha_1 f + \alpha_0 \operatorname{Id}$$

Einsetzabbildung oder *Einsetzhomomorphismus.*

Unter der Einsetzabbildung erhalten wir nun für unser Polynom p mit $p(f) = 0$

$$\operatorname{Id} = \sum_{i=1}^{r} h_i(f) q_i(f) = \sum_{i=1}^{r} c_i \quad \text{mit} \quad c_i := h_i(f) q_i(f).$$

Für $i \neq j$ gilt dabei $c_i c_j = h_i(f) q_i(f) h_j(f) q_j(f) = g(f) p(f) = 0$ und daraus folgt mit $\operatorname{Id} = \sum c_i$ sofort $c_i^2 = c_i$, da $c_i = \sum c_i c_j = c_i^2$. Somit gilt insgesamt

$$c_i c_j = \delta_{ij} c_i$$

und die c_i bilden ein vollständiges System zueinander orthogonaler Projektionen.

Satz von der Primärzerlegung. *Sei f ein Endomorphismus über dem endlich erzeugten Vektorraum V und $p(\tau) = \prod p_i(\tau)^{k_i}$ mit $p(f) = 0$ wie oben. Dann ist V die direkte Summe f-invarianter Untervektorräume:*

$$V = \bigoplus_{i=1}^{r} \operatorname{Kern} p_i(f)^{k_i}.$$

Beweis: Für alle $v \in V$ gilt

$$v = \operatorname{Id} v = \Big(\sum c_i\Big) v = \sum c_i(v) = \sum v_i, \qquad \text{mit } c_i(v) =: v_i \in V_i.$$

Also ist jeder Vektor v als Summe von Vektoren aus den V_i darstellbar. Sei nun $\sum v_i = 0$ mit $v_i = c_i(w_i)$ $(1 \le i \le r)$. Durch Anwendung von c_j erhalten wir

$$0 = \sum c_j c_i(w_i) = \sum \delta_{ij} c_i(w_i) = c_j(w_j) = v_j.$$

Also ist die Summe der Vektorräume V_i direkt.
Zuerst zeigen wir jetzt, daß $c_i(V) \subseteq \operatorname{Kern} p_i(f)^{k_i}$ ist: Sei $u \in c_i(V)$. Damit gilt $u = c_i(w) = h_i(f) q_i(f) w$ für ein $w \in V$. Anwenden von $p_i(f)^{k_i}$ ergibt

$$p_i(f)^{k_i}(u) = h_i(f) \cdot \underbrace{p_i(f)^{k_i} \cdot q_i(f)}_{= p(f) = 0}(w) = 0.$$

Damit ist $u \in \operatorname{Kern} p_i(f)^{k_i}$.
Nun müssen wir noch zeigen, daß $\operatorname{Kern} p_i(f)^{k_i} \subseteq c_i(V)$ gilt. Sei dazu $p_i(f)^{k_i} v = 0$. Da für $i \neq j$ das Polynom $p_i(\tau)^{k_i}$ als Faktor in $q_j(\tau)$ auftritt, gilt auch $q_j(f) v = 0$ für $i \neq j$, dann erst recht $c_j(v) = 0$ für $i \neq j$ und damit $v = \operatorname{Id} v = (\sum c_j) v = \sum c_j(v) = c_i(v)$. □

Beispiel: Betrachte

$$\alpha_n y^{(n)} + \alpha_{n-1} y^{(n-1)} + \dots + \alpha_1 y' + \alpha_0 y = 0,$$

mit $\alpha_i \in \mathbf{R}$ und $y := y(x)$.

$$V = \left\{ y \in C^\infty(\mathbf{R}) \;\middle|\; \alpha_n y^{(n)} + \alpha_{n-1} y^{(n-1)} + \dots + \alpha_1 y' + \alpha_0 y = 0 \right\}$$

ist $\mathbf{R}$-Vektorraum mit $\dim V = n$. Wir setzen $D\colon V \to V$, $Dy := y'$, $p(\tau) = \alpha_n \tau^n + \dots + \alpha_0$ und damit $p(D) = 0$. Faktorisieren ergibt

$$p(\tau) = \prod_{i=1}^{r} p_i(\tau)^{k_i}.$$

Aus dem Satz folgt nun

$$V = \bigoplus_{i=1}^{r} \operatorname{Kern} p_i(D)^{k_i}.$$

Für $p_i(\tau) = \tau - \alpha_i$ gilt weiter

$$V_i = \operatorname{Kern}(D - \alpha_i \operatorname{Id})^{k_i} = \left\{ y \in C^\infty(\mathbf{R}) \;\middle|\; (D - \alpha_i \operatorname{Id})^{k_i} y = 0 \right\}.$$

Ferner gilt (der Beweis ist eine einfache Übung)

$$(D - \alpha_i \operatorname{Id})^{k_i} y = 0 \iff y(x) = e^{\alpha x}(\beta_0 + \beta_1 x + \dots + \beta_{k-1} x^{k-1}).$$

◊

7.3 Nilpotente Endomorphismen

Seien V ein K-Vektorraum und f ein Endomorphismus über V. Bekanntlich heißt f dann *nilpotent*, wenn es ein m mit $f^m = 0$ gibt. (Dabei wollen wir den trivialen Fall $f = 0$ ausschließen, also $m > 1$.) Die dadurch eindeutig bestimmte Zahl m nennen wir *Nilpotenzindex*. Wie wir bereits wissen, sind alle Eigenwerte nilpotenter Abbildungen Null, das charakteristische Polynom ist somit einfach $\chi_f(\tau) = \tau^n$. Die Spur der Darstellungsmatrix und ihre Determinante sind ebenfalls immer Null.

Satz. *Sei f ein nilpotenter Endomorphismus mit $f^m = 0$ und $f^{m-1} \neq 0$. Dann sind für alle $a \in V$ mit $f^{m-1}(a) \neq 0$ die Vektoren*

$$\big(a, f(a), f^2(a), \dots, f^{m-1}(a)\big)$$

linear unabhängig.

Beweis: Wir betrachten die Linearkombination $\alpha_0 a + \alpha_1 f(a) + \dots + \alpha_{m-1} f^{m-1}(a) = 0$. Anwenden von f^{m-1} ergibt $\alpha_0 f^{m-1}(a) + 0 + \dots + 0 = 0$. Also folgt sofort aus $f^{m-1}(a) \neq 0$, daß $\alpha_0 = 0$ gilt. Durch anwenden von f^{m-2}, f^{m-3}, etc. folgt analog $\alpha_1 = 0$, $\alpha_2 = \dots = \alpha_{m-1} = 0$. □

Nach diesem Satz stellen also die Vektoren $\big(a, f(a), f^2(a), \ldots, f^{m-1}(a)\big)$ eine Basis eines m-dimensionalen, von a zyklisch erzeugten, f-invarianten Vektorraum dar, den wir mit

$$Z(a,f) := \mathrm{Lin}\big(a, f(a), f^2(a), \ldots, f^{m-1}(a)\big)$$

bezeichnen. Die Matrixdarstellung von $f|_Z$ bezüglich dieser Basis hat folgendes Aussehen:

$$\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \ldots & 0 & 1 & 0 \end{pmatrix}.$$

Um die Einser auf die obere Nebendiagonale zu verschieben, führen wir folgende Substitution durch:

$$b_i := f^{m-i}(a), \qquad 1 \le i \le m.$$

Damit ist $B = (b_1, \ldots, b_m)$ Basis von $Z(a,f)$ mit

$${}_B[f|_Z]_B = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \ldots & 0 & 0 & 0 \end{pmatrix}.$$

Eine solche Matrix heißt *nilzyklisch*.

Hauptsatz für nilpotente Abbildungen. *Seien V ein endlicherzeugter Vektorraum und f ein nilpotenter Endomorphismus über V. Dann gilt*

$$V = \bigoplus_{i=1}^{r} Z(a_i, f),$$

d.h. V ist direkte Summe von f-invarianten, zyklischen Untervektorräumen. Somit gibt es eine Basis $B = (B_1, \ldots, B_r)$ von V mit

$${}_B[f]_B = \begin{pmatrix} \boxed{N_1} & 0 & \ldots & 0 \\ 0 & \boxed{N_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \boxed{N_r} \end{pmatrix}$$

mit nilzyklischen N_i.

Beweis mit vollständiger Induktion: Wir zeigen $Z(a,f)$ – dabei sei a wie oben durch $f^m = 0$ und $f^{m-1}(a) \neq 0$ bestimmt – besitzt einen f-invarianten, direkten Summanden, d.h.

$$V = Z(a,f) \oplus W,$$

wobei W ein f-invarianter Unterraum von V ist.
Sei $U := Z(a,f)$ mit Basis $B = (b_1, \ldots, b_m)$, $b_i = f^{m-i}(a)$ wie bisher; $b := b_1 \neq 0$. Dann existiert $b^* \in V^*$: $b^*(b) \neq 0$, d.h. $b^*\big(f^{m-1}(a)\big) \neq 0$, $(f^{m-1})^*(b^*)(a) \neq 0 \Longrightarrow (f^*)^{m-1}(a) \neq 0$ und damit $(f^*)^{m-1} \neq 0$. Nun ist $U^* = \mathrm{Lin}\big(b^*, f^*(b^*), \ldots, (f^*)^{m-1}(b^*)\big) = Z(b^*, f^*)$ ein f^*-invarianter Unterraum. Betrachten wir nun

$$(U^*)^\perp = \big\{\, v \in V \;\big|\; u^*(v) = 0 \quad \text{für alle } u^* \in U^*. \,\big\}$$

Da U^* f^*-invariant ist, ist auch $(U^*)^\perp$ f-invariant und zusammen mit $v \in (U^*)^\perp$ folgt

$$u^*(fv) = (f^*u^*)(v) = 0 \quad \text{für } u^* \in U^*,$$

d.h. $f(v) \in (U^*)^\perp$.
Nun müssen wir nur noch zeigen, daß $V = Z(a,f) \oplus (U^*)^\perp$ gilt. Zunächst geht aus der Definition hervor, daß $Z(a,f) \cap (U^*)^\perp = \{0\}$. Daraus folgt:

$$Z(a,f) + (U^*)^\perp = Z(a,f) \oplus (U^*)^\perp.$$

Dimensionsbetrachtungen zeigen $\dim V = \dim Z(a,f) + \dim(U^*)^\perp$:

$$\dim U^* + \dim(U^*)^\perp = \dim V = \dim Z(a,f)$$

nach Konstruktion. Daraus folgt die Behauptung. □

7.4 Jordan-Normalform

Im folgenden beschränken wir uns auf Vektorräume über dem Körper $K = \mathbf{C}$. Sei f nun wieder ein gewöhnlicher Endomorphismus (nicht notwendigerweise nilpotent) mit dem charakteristischen Polynom $p(\tau)$. Über dem Körper $\mathbf{C}$ zerfällt dieses Polynom in Linearfaktoren:

$$p(\tau) = \prod_{i=1}^{r} (\tau - \alpha_i)^{n_i}.$$

Für das Minimalpolynom gilt entsprechend

$$\chi(\tau) = \prod_{i=1}^{r} (\tau - \alpha_i)^{k_i} \qquad \text{mit } 1 \le k_i \le n_i.$$

Die Primärzerlegung lautet dazu

$$V = \bigoplus_{i=1}^{r} \mathrm{Kern}(f - \alpha_i \,\mathrm{Id})^{k_i},$$

mit $\dim \mathrm{Kern}(\tau - \alpha_i)^{k_i} = k_i$. Die *Primärkomponente* zum Eigenwert α_i bezeichnen wir mit V_i. Die Darstellungsmatrix von f nennen wir F. Nun folgt aus $(F - \alpha_i E)^{k_i} = 0$, daß $(F - \alpha_i E)$ nilpotent ist. Durch mehrmaliges Anwenden des Satzes aus dem vorhergehenden Kapitel erhalten wir nun

eine Basis B_i, so daß für die einzelnen Eigenräume die $N_i := A - \alpha_i E$ nilzyklisch sind. Dann gilt für die einzelnen Blöcke $B_i^{-1} F B_i = \alpha_i E + B_i^{-1} N_i B_i$. Jeder einzelne Block hat somit die Gestalt

$$J_i := \begin{pmatrix} \alpha_i & 1 & 0 & \dots & 0 \\ 0 & \alpha_i & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \alpha_i & 1 \\ 0 & \dots & 0 & 0 & \alpha_i \end{pmatrix}.$$

Insgesamt ergibt sich damit folgende Darstellung von f:

$$\begin{pmatrix} \boxed{B_1} & 0 & \dots & 0 \\ 0 & \boxed{B_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boxed{B_r} \end{pmatrix}^{-1} \begin{pmatrix} \boxed{A_1} & 0 & \dots & 0 \\ 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boxed{A_r} \end{pmatrix} \begin{pmatrix} \boxed{B_1} & 0 & \dots & 0 \\ 0 & \boxed{B_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boxed{B_r} \end{pmatrix}$$

$$= \begin{pmatrix} \boxed{B_1^{-1} A_1 B_1} & 0 & \dots & 0 \\ 0 & \boxed{B_2^{-1} A_2 B_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boxed{B_r^{-1} A_r B_r} \end{pmatrix} = \begin{pmatrix} \boxed{J_1} & 0 & \dots & 0 \\ 0 & \boxed{J_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boxed{J_r} \end{pmatrix}.$$

Diese Darstellung heißt *Jordannormalform von* f.

Beispiel: Sei

$$A := \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix}, \qquad \chi_A = (\tau - 2)^4,$$

mit Eigenwert $\alpha = 2$ und Vielfachheit 4. $\operatorname{Rang}(A - 2E) = 1$, $\operatorname{Dim}\operatorname{Kern}(A - 2E) = 3$, also gibt es drei nilzyklische Kästen:

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

ist Jordan-Normalform von A zu geeigneter Basis. ◊

7.5 Komplexifizierung

In den vorherigen Abschnitten haben wir uns auf den Körper $\mathbf{C}$ beschränkt, da dort das charakteristische Polynom in Linearfaktoren zerfällt. In diesem Abschnitt wollen wir diese Einschränkung aufheben und Endomorphismen über dem Körper $\mathbf{R}$ behandeln.
Sei dazu $A \in \mathbf{R}^{n\times n}$ bzw. f ein Endomorphismus über dem $\mathbf{R}$-Vektorraum V. Dann komplexifizieren wir, d.h. wir setzen

$$V_C = V \oplus iV$$

mit $(\alpha + i\beta)(u + iv) = (\alpha u - \beta v) + i(\alpha v + \beta u)$ und $\tilde{f}(u + iv) = f(u) + if(v)$.
Zu reellen Eigenwerten $\alpha \in \mathbf{R}$ erhalten wir reelle Basisvektoren und entsprechend

$$J_i = \begin{pmatrix} \alpha & 1 & 0 & \dots & 0 \\ 0 & \alpha & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \alpha & 1 \\ 0 & \dots & 0 & 0 & \alpha \end{pmatrix}.$$

Zu komplexen Eigenwerten $\zeta = \alpha + i\beta$ mit $\beta \neq 0$ erhalten wir Basisvektoren u_l $(1 \le l \le k)$. Es gilt $(\tilde{f} - \zeta\,\mathrm{Id}) = N$.

$$\begin{aligned} \tilde{f}u_l &= \zeta u_l + u_{l-1}, \\ \tilde{f}\bar{u}_l &= \zeta \bar{u}_l + \bar{u}_{l-1}, \\ u_l &= a_l + ib_l \quad \text{mit } a_l, b_l \in V \text{ reell}. \end{aligned}$$

Sind u_l und $\bar{u}_l$ linear unabhängig, so sind auch a_l und b_l linear unabhängig. Es gilt

$$f(a_l) + if(b_l) = \alpha a_l - \beta b_l + i(\beta a_l + \alpha b_l) + a_{l-1} + ib_{l-1}.$$

Vergleich von Real- und Imaginärteil ergibt:

$$\begin{aligned} f_{a_l} &= \alpha a_l - \beta b_l + a_{l-1}, \\ f_{b_l} &= \beta a_l + \alpha b_l + b_{l-1}. \end{aligned}$$

Setzen wir nun aus den a_i und b_i die reelle Basis $B = (a_1, b_1, \dots, a_l, b_l)$ zusammen, so erhalten wir in der *reellen Jordannormalform* den zu $\zeta = \alpha + i\beta$ gehörenden „Jordan-Kasten“ (um die Darstellung übersichtlicher zu gestalten sind einige Nullen weggelassen):

$$J := \begin{pmatrix} \alpha & \beta & 1 & & & & \\ -\beta & \alpha & 0 & 1 & & & \\ 0 & 0 & \alpha & \beta & 1 & & \\ & 0 & -\beta & \alpha & 0 & \ddots & \\ & & 0 & \ddots & \ddots & \ddots & 1 \\ & & & \ddots & 0 & \alpha & \beta \\ & & & & 0 & -\beta & \alpha \end{pmatrix}.$$

7.6 Anwendungen

Satz. *Jede komplexe, quadratische Matrix ist zu ihrer Transponierten ähnlich.*

Beweis: Zu einer beliebigen Matrix $A \in \mathbf{C}^{n\times n}$ existiert $W \in \mathrm{GL}(n, \mathbf{C})$ derart, daß

$$W^{-1} A W = A^T,$$

gilt, denn A und A^T haben die gleiche Jordan-Normalform. □

Satz über die Tridiagonalisierung. *Zu $A \in \mathbf{R}^{n\times n}$ existiert ein $W \in \mathrm{GL}(n,\mathbf{R})$, mit*

$$W^{-1}AW = \begin{pmatrix} * & * & & 0 \\ * & \ddots & \ddots & \\ & \ddots & \ddots & * \\ 0 & & * & * \end{pmatrix}.$$

Beweis von A. Gerstner-Bunse (1980): Vergleiche

$$\begin{pmatrix} \alpha & \beta & 1 & 0 & \dots & & 0 \\ -\beta & \alpha & 0 & 1 & \ddots & & \\ 0 & 0 & \alpha & \beta & 1 & \ddots & \vdots \\ \vdots & \ddots & -\beta & \alpha & 0 & \ddots & 0 \\ & & & \ddots & \ddots & \ddots & 1 \\ \vdots & & & \ddots & 0 & \alpha & \beta \\ 0 & \dots & & \dots & 0 & -\beta & \alpha \end{pmatrix} \quad \text{mit} \quad \begin{pmatrix} \alpha & \beta & 0 & \dots & & \dots & 0 \\ -\beta & \alpha & 0 & \ddots & & & \vdots \\ 0 & r_1 & \alpha & \beta & & & \\ \vdots & \ddots & -\beta & \alpha & 0 & \ddots & \vdots \\ & & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & r_\nu & \alpha & \beta \\ 0 & \dots & & \dots & 0 & -\beta & \alpha \end{pmatrix}.$$

Sind $\beta \neq 0$ und $r_i \neq 0$, so ist die linke Matrix die reelle Jordan-Normalform der rechten Matrix. □

Im folgenden wollen wir uns mit der Matrix-Exponentialfunktion beschäftigen.
Sei dazu $A \in \mathbf{C}^{n\times n}$. Bilde nun analog zu einer konvergenten Potenzreihe die *Matrix-Reihe*

$$\sum_{k=0}^{\infty} \alpha_k A^k = \alpha_0 E + \alpha_1 A + \cdots + \alpha_n A^n + \cdots.$$

Die Konvergenz ist hier komponentenweise erklärt.
Betrachten wir nun den Spezialfall

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = E + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots.$$

Behauptung: Die Reihe konvergiert.
Beweis: Wir führen dazu die *Maximumsnorm* ein: Für $A = (\alpha_{ij}) \in \mathbf{C}^{n\times n}$ ist

$$\|A\| := \max_{i,j} |\alpha_{ij}|.$$

Hierfür haben wir folgende Rechenregeln:

$$\begin{gathered} \|A+B\| \le \|A\| + \|B\|, \\ \|tA\| = |t| \cdot \|A\| \quad \text{für } t \in \mathbf{C}, \\ \|AB\| \le n\|A\|\,\|B\|, \\ \|A^k\| \le \big(n\|A\|\big)^k. \end{gathered}$$

Daraus folgt sofort

$$\Big\| \sum_{k=0}^{N} \frac{1}{k!} A^k \Big\| \le e^{n\|A\|}.$$

Da $e^{tA} = \sum t^k/k! \cdot A^k$ absolut und gleichmäßig auf jedem endlichen Teilintervall $I \subset \mathbf{R}$ konvergiert, folgt

$$\frac{d}{dt} e^{tA} = \sum \frac{1}{k!} \frac{d}{dt} t^k A^k = A \cdot e^{tA} = e^{tA} \cdot A.$$

□

Es gelten folgende Rechenregeln:

$$\begin{gathered} e^0 = E, \\ \frac{d}{dt} e^{tA} = A \cdot e^{tA} = e^{tA} \cdot A, \\ AB = BA \implies e^{A+B} = e^A e^B = e^B e^A, \\ (e^A)^{-1} = e^{-A} \quad \text{mit } e^A \in \mathrm{GL}(n,\mathbf{C}), \\ e^A = W e^{W^{-1}AW} W^{-1} \quad \text{mit } W \in \mathrm{GL}(n,\mathbf{C}). \end{gathered}$$

Beweis des letzten Punktes:

$$\begin{aligned} W^{-1}\Big(\sum \alpha_i A^i\Big)W &= W^{-1}(\alpha_0 E + \alpha_1 A + \cdots)W \\ &= \alpha_0 E + \alpha_1 W^{-1}AW + \alpha_2 (W^{-1}AW)^2 + \cdots \\ &= \sum \alpha_i (W^{-1}AW)^i. \end{aligned}$$

Sei nun $x(t) \in \mathbf{C}^n$ mit $t \in \mathbf{R}$.
Satz. *Das lineare Differentialgleichungssystem*

$$\dot{x} = Ax \quad \textit{mit } A \in \mathbf{C}^{n\times n} \quad \textit{und} \quad x(0) = c \in \mathbf{C}^{n\times n}$$

hat die eindeutige Lösung

$$x(t) = e^{tA}c.$$

Beweis: Aufgrund der Theorie der Differentialgleichungen existiert genau eine Lösung. Offenbar erfüllt $e^{tA}c$ die Differentialgleichung.
Zur Berechnung von e^{tA}:

1. über Schur-Normalform: Es habe $B := W^{-1}AW$ Schur-Normalform, d.h. obere Dreiecksgestalt mit den Eigenwerten auf der Hauptdiagonalen. Dann ist

$$e^A = W e^B W^{-1} = W\Big(E + B + \frac{1}{2!}B^2 + \cdots\Big)W^{-1} = W \begin{pmatrix} e^{\lambda_1} & * & \dots & * \\ 0 & e^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & e^{\lambda_n} \end{pmatrix} W^{-1}.$$

Außerdem ist

$$\det e^A = \det e^B = e^{\lambda_1} \dots e^{\lambda_n} = e^{\mathrm{Sp}\,A}.$$

2. über Jordan-Normalform: Sei nun

$$A = \begin{pmatrix} \boxed{J_1} & 0 & \dots & 0 \\ 0 & \boxed{J_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \boxed{J_n} \end{pmatrix}, \quad \text{also } e^A = \begin{pmatrix} \boxed{e^{J_1}} & 0 & \dots & 0 \\ 0 & \boxed{e^{J_2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \boxed{e^{J_n}} \end{pmatrix}.$$

Damit ist nur noch

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{pmatrix} = \lambda E + N \quad \text{mit } N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix}$$

interessant. Also ist mit $N^{k-1} \neq 0$, aber $N^k = 0$,

$$e^J = e^{\lambda E} e^N = e^\lambda \left(E + N + \frac{1}{2!}N^2 + \dots + \frac{1}{(k-1)!}N^{k-1} \right)$$
$$= e^\lambda \begin{pmatrix} 1 & 1 & 1/2 & \dots & 1/(k-1)! \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 1/2 \\ & & & \ddots & 1 \\ 0 & & & & 1 \end{pmatrix}$$

und entsprechend

$$e^{tJ} = e^\lambda \begin{pmatrix} 1 & t & t^2/2 & \dots & t^{k-1}/(k-1)! \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & t^2/2 \\ & & & \ddots & t \\ 0 & & & & 1 \end{pmatrix}.$$

□

Kapitel 8

Unendliche Vektorräume

8.1 Definitionen

Sei V ein unendlichdimensionaler K-Vektorraum und S eine Teilmenge von V. Ferner sei x eine Linearkombination aus S:

$$x = \sum_{s \in S}{}' \alpha_s s,$$

wobei $\sum'$ bedeutet, daß fast alle α_s Null sind. Also gibt es dann endlich viele $s_1, \dots, s_k$ und $\alpha_1, \dots, \alpha_k \in K$ mit

$$x = \alpha_1 s_1 + \dots + \alpha_k s_k.$$

Definition. Die *lineare Hülle* von S definieren wir wie folgt:

$$\operatorname{Lin} S := \left\{ \sum_{s \in S}{}' \alpha_s s \;\middle|\; \alpha_s \in K \right\}.$$

Dann heißt S *Erzeugendensystem* für $\operatorname{Lin} S$.
Zur Vermeidung von Fallunterscheidungen definieren wir

$$\operatorname{Lin} \emptyset := \{0\}.$$

Definition. Sei S eine Teilmenge von V. Dann heißt S *linear unabhängig*, wenn jedes endliche System $(s_1, \dots, s_k) \in S$ linear unabhängig ist, was genau dann gilt, wenn

$$\sum_{s \in S}{}' \alpha_s s = 0 \quad \Longleftrightarrow \quad \alpha_s = 0 \text{ für alle } s.$$

Definition. Die Teilmenge $B \subseteq V$ heißt *Basis*, wenn B linear unabhängig ist und $\operatorname{Lin} B = V$ gilt.

Satz über die Charakterisierung einer Basis. *Folgende Aussagen sind äquivalent:*

(1) B *ist Basis von* V.

(2) B *ist eine maximal linear unabhängige Familie.*

(3) B *ist minimales Erzeugendensystem.*

Beweis:

(1)$\Longrightarrow$(2): Zu zeigen ist, daß aus $B \subseteq S$ und S linear unabhängig folgt, daß $B = S$ gilt. Sei dazu $s \in S \setminus B$, $s = \sum'_{b\in B} \alpha_b b$ (denn B ist Basis!), dann folgt, daß $B \cup \{s\}$ linear abhängig ist, aber $B \subseteq S$ ist ein Widerspruch!

(2)$\Longrightarrow$(1): Sei $v \in V \setminus B$. Dann folgt $\alpha_v v + \sum'_{b\in B} \alpha_b b = 0$ und daraus $\alpha_v \neq 0$, da die b_i linear unabhängig und $B \cup \{v\}$ linear abhängig sind. Daraus folgt nun $v = -(1/\alpha_v) \sum \alpha_b b$. Somit ist B eine Basis.

□

Satz. *Sei V ein K-Vektorraum und $S \subseteq V$ linear unabhängig. Dann gibt es eine Basis B von V (Existenz einer Basis) und für jede Basis C von V gilt $S \subseteq C$ (Basisergänzung).*

Beweis: Seien $V \neq \{0\}$ und

$$\mathcal{M} := \{\, X \subseteq V \mid X \text{ linear unabhängig} \,\} \neq \emptyset.$$

Sei K eine Kette in M:

$$Y := \bigcup_{X\in K} X$$

ist linear unabhängig, also in $\mathcal{M}$. Da $X \subseteq Y$ für alle $X \in K$ ist Y eine obere Schranke. Also besitzt $\mathcal{M}$ nach dem Lemma von Zorn ein maximales Element B, das somit maximal linear unabhängig und somit eine Basis ist. □

Beispiel: Die Menge $\mathbf{R}$ ist ein $\mathbf{Q}$-Vektorraum und es existiert eine $\mathbf{Q}$-Basis. ◊

Anwendung des Lemmas von Zorn: Jeder Vektorraum besitzt eine Basis.

Basisergänzung. *Jede linear unabhängige Menge $S \subseteq V$ kann zu einer Basis B mit $S \subseteq B$ ergänzt werden.*

Beweis: Sei $X = \{\, A \subseteq V \mid A \text{ linear unabhängig}, S \subseteq A \,\} \neq \emptyset$. Ferner sei K eine Kette in X. Dann gilt:

$$Z := \bigcup_{U\in K} U \in X.$$

□

Satz. *Je zwei Basen sind gleichmächtig.*

Frage: Wie sieht Basis aus? (offen)
Beispiel: $\mathbf{R}$ als $\mathbf{Q}$-Vektorraum. Keine Basis ist explizit bekannt. Sei die Menge

$$S := \{\, \ln p \mid p \text{ prim} \,\}$$

linear unabhängig. Denn aus $\sum \alpha_k \ln p_k = 0$ folgt

$$\ln \prod p_k^{\alpha_k} = 0 \quad\Longrightarrow\quad \prod p_k^{\alpha_k} = 1$$

und daraus, daß alle α_k Null sind.

8.2 Normierte Vektorräume

Sei im folgendem V ein Vektorraum über dem Körper $\mathbf{R}$ oder $\mathbf{C}$.

Definition. Die Abbildung $\|\ \|\colon V \to \mathbf{R}$ heißt *Norm auf* V, wenn für alle $u, v \in V$ und $\alpha \in K$

$$\|v\| \geq 0 \quad\text{und}\quad \|v\| = 0 \Longleftrightarrow v = 0, \tag{N1}$$
$$\|\alpha v\| = |\alpha| \cdot \|v\| \quad\text{und} \tag{N2}$$
$$\|u + v\| \leq \|u\| + \|v\| \qquad \text{(Dreiecksungleichung)} \tag{N3}$$

gilt. Wir nennen dann $(V, \|\ \|)$ einen *normierten Vektorraum*.

Beispiele:

1. Seien $v \in K^n$ und $p \in \mathbf{N}$. Dann ist

$$\|v\|_p := \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}$$

   eine Norm. Für $p = 2$ heißt diese auch *euklidische Norm*.

2. Mit $v \in K^n$ und der *Maximumsnorm*

$$\|v\|_{\max} := \max\{\, |v_i| \mid 1 \leq i \leq n \,\}$$

   ist $(K^n, \|\ \|)$ ein normierter Vektorraum.

3. Die Menge

$$\ell^p := \left\{ (x_k)_{k\geq 0} \;\middle|\; x_k \in K,\ \sum_{k=0}^{\infty} |x_k|^p < \infty \right\}$$

   ist ein K-Vektorraum mit möglicher Norm

$$\|x\| := \left(\sum_{k=0}^{\infty} |x_k|^p\right)^{1/p}.$$

4. Die Menge

$$\ell^\infty := \{\, (x_k)_{k\geq 0} \mid x_k \in K,\ (x_k)_{k\geq 0} \text{ beschränkt} \,\}$$

   ist ebenfalls ein K-Vektorraum mit möglicher Norm

$$\|x\| := \sup\{\, |x_k| \mid k \geq 0 \,\}.$$

5. Gleiches gilt für

$$L^p := \left\{ f\colon I \to K \;\middle|\; \int_I |f(t)|^p dt < \infty \right\}.$$

◊

Definition. Seien $(V, \|\ \|)$ ein normierter Vektorraum und $u, v \in V$. Dann heißt die Funktion

$$d(u, v) := \|u - v\|$$

der *Abstand von u und v* bzw. auch *Metrik*. Ferner heißt die Folge $(x_k)_{k\geq 0}$ *konvergent gegen den Grenzwert* x, wenn gilt:

$$\lim_{k\to\infty} x_k = x \quad \Longleftrightarrow \quad \lim_{k\to\infty} \|x_k - x\| = 0.$$

Definition. Die Folge $(x_k)_{k\geq 0}$ heißt *Cauchy-Folge*, wenn es zu jedem $\epsilon > 0$ ein $n_0 \in \mathbf{N}$ mit

$$\|x_k - x_l\| < \epsilon \qquad \text{für alle } k, l \geq n_0$$

gibt.

Definition. Ein normierter Vektorraum $(V, \|\ \|)$ heißt *vollständig* oder *Banachraum*, wenn jede Cauchy-Folge in V konvergiert.

Definition. Sei $A \subseteq V$. Wir definieren die Menge

$$\bar{A} := \left\{ v \in V \;\middle|\; \text{es gibt } (a_k)_{k\geq 0},\ a_k \in A \text{ und } \lim_{k\to\infty} a_k = v \right\}.$$

Ferner heißt die Menge A *abgeschlossen*, wenn $A = \bar{A}$ gilt.

Satz. *Seien $(V, \|\ \|)$ ein normierter, vollständiger Vektorraum und $A \subseteq V$. Die Menge A ist genau dann vollständig, wenn $A = \bar{A}$ gilt, also wenn A abgeschlossen ist.*

Beweis: Zunächst nehmen wir an, A sei vollständig. Sei dann $a \in \bar{A}$ und (a_k) eine Cauchy-Folge mit Grenzwert a. Nach Voraussetzung liegt der Grenzwert in A. Also folgt $\bar{A} \subseteq A$ und daraus $A = \bar{A}$.
Nehmen wir nun $A = \bar{A}$ an. Sei (a_k) eine Cauchy-Folge in A und damit auch eine Cauchy-Folge in V, so hat sie nach Voraussetzung einen Grenzwert a. Damit gilt:

$$a_k \to a \in \bar{A} = A.$$

□

8.3 Hilberträume

Definition. Ein Innenproduktraum $(V, \langle\ ,\ \rangle)$ heißt *Hilbertraum*, wenn V bezüglich der Norm

$$\|v\| := \sqrt{\langle v, v\rangle}$$

vollständig ist.

Satz. *Jeder endlichdimensionale Innenproduktraum ist ein Hilbertraum.*

Beweis: Wir bezeichnen den Innenproduktraum mit $(V, \langle\ ,\ \rangle)$. Sei $(v_1, \ldots, v_n)$ eine Orthonormalbasis von V. Jeder Vektor $v \in V$ besitzt die Darstellung

$$v = \sum_{k=1}^{n} \langle v, v_k\rangle v_k.$$

Sei $(x_k)_{k\geq 0}$ eine Cauchy-Folge in V, dann kann man die x_k darstellen als

$$x_k = \sum_{j=1}^{n} \langle x_k, v_j\rangle v_j.$$

Für die Folge der j-ten Koeffizienten gilt:

$$\left|\langle x_k, v_j\rangle - \langle x_l, v_j\rangle\right| = \left|\langle x_k - x_l, v_j\rangle\right| \leq \underbrace{\|v_j\|}_{=1} \cdot \|x_k - x_l\| \to 0,$$

da die x_k eine Cauchy-Folge bilden. Da K vollständig ist, hat die Koeffizientenfolge als Cauchy-Folge einen Grenzwert und somit hat auch (x_k) einen Grenzwert in V. □

Im folgenden bezeichnen wir mit K den Körper der reellen bzw. komplexen Zahlen.
Definition. Jeder Innenproduktraum $(V, \langle\ ,\ \rangle)$ heißt *Prähilbertraum*.

Satz. *Jeder Prähilbertraum läßt sich zu einem Hilbertraum vervollständigen.*

Beispiele:

1. Die Menge

$$\left\{ (x_n)_{n\geq 0} \;\middle|\; x_n \in K, \sum_{k=0}^{\infty} |x_n|^2 < \infty \right\}$$

   mit Skalarprodukt $\langle a, b\rangle = \sum_{k=0}^{\infty} a_k \bar{b}_k$, wobei $a = (a_n)$, $b = (b_n)$ und Norm $\|a\| = \sqrt{\langle a, a\rangle} = \left(\sum |a_k|^2\right)^{1/2}$, ist ein Hilbertraum.

2. Die Menge

$$C^0(I) = \left\{ f \;\middle|\; I \to K,\ f \text{ stetig} \right\}$$

   mit Skalarprodukt $\langle f, g\rangle = \int_{-1}^{1} f(t)\bar{g}(t)\,dt$ ist ein zu vervollständigender Prähilbertraum.

◊

Lemma 8.1: *Ist $(V, \langle\ ,\ \rangle)$ ein Prähilbertraum mit $u, v \in V$, so ist*

$$u \perp v \quad \Longleftrightarrow \quad \langle u, v\rangle = 0 \qquad \textit{und}$$
$$u \perp v \quad \Longrightarrow \quad \|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Satz. *Für alle $v \in V$ gilt:*

$$\|v\|^2 = \sum_{k=1}^{n} \left|\langle v, v_n\rangle\right|^2 + \left\| v - \sum_{k=1}^{n} \langle v, v_k\rangle v_k \right\|^2.$$

Beweis: Sei $u = \sum_{k=1}^{n} \langle v, v_k\rangle v_k$. Damit ist $(v - u) \perp u$, denn

$$\left\langle v - \sum_{k=1}^{n} \langle v, v_k\rangle v_k, v_j \right\rangle = \langle v, v_j\rangle - \sum_{k=1}^{n} \langle v, v_k\rangle \langle v_k, v_j\rangle = 0.$$

Aus $\|v\|^2 = \|u\|^2 + \|v - u\|^2 = \sum |\langle v, v_k\rangle|^2 + \|v - u\|^2$ folgt die Behauptung. □

Korollar. *Sei $(v_1, \dots, v_n)$ eine Orthonormalbasis von V. Dann gilt*

$$\|v^2\| \geq \sum_{k=1}^{n} \big|\langle v, v_k\rangle\big|^2 \qquad \textit{(Bessel-Ungleichung).}$$

Die Koeffizienten $\langle v, v_k\rangle$ heißen Fourierkoeffizienten.

Satz. *Mit einer Orthonormalbasis $(v_1, \dots, v_n)$ gilt stets*

$$\left\| v - \sum_{k=1}^{n} \langle v, v_k\rangle v_k \right\| \leq \left\| v - \sum_{k=1}^{n} \alpha_k v_k \right\|,$$

d.h. die beste Approximation wird mit den Fourierkoeffizienten erreicht.

Beweis: Seien $u := \sum \langle v, v_k\rangle v_k$, $u_1 := \sum \alpha_k v_k$ und damit $w := v - u$, $w_1 := v - u_1$. Dann ist $w + u = v = w_1 + u_1$ und so $w_1 = (u - u_1) + w$. Da $(u - u_1) \perp w$, folgt nach dem Satz des Pythagoras $\|w_1\|^2 = \|u - u_1\|^2 + \|w^2\|$, also $\|w\| \leq \|w_1\|^2$. □

Parallelogrammungleichung. *Sei $\big(V, \langle\,,\,\rangle\big)$ ein Prähilbertraum. Dann gilt*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Beweis:

$$\begin{aligned}\langle u + v, u + v\rangle + \langle u - v, u - v\rangle &= \langle u, u\rangle + 2\langle u, v\rangle + \langle v, v\rangle + \langle u, u\rangle - 2\langle u, v\rangle\langle v, v\rangle \\ &= 2\big(\langle u, u\rangle + \langle v, v\rangle\big).\end{aligned}$$

□

Satz. *Ist $\big(V, \langle\,,\,\rangle\big)$ ein Hilbertraum, so gilt*

(*1*) *Das Skalarprodukt $\langle\,,\,\rangle\colon V \times V \to K$ ist stetig*

(*2*) *Die Norm $\|\ \|\colon V \to K$ ist stetig.*

Beweis:

(*1*) Für geeignete x', y' gilt:

$$\begin{aligned}|\langle x, y\rangle - \langle x', y'\rangle| &= |\langle x - x' + x', y\rangle - \langle x', y'\rangle| \\ &\leq |\langle x - x', y\rangle| + |\langle x', y - y'\rangle| \\ &\leq |y|\|x - x'\| + \|x'\|\,\|y - y'\| \\ &< \epsilon.\end{aligned}$$

(*2*) Aus

$$\big|\|x\| - \|x'\|\big| \leq \|x' - x\| < \epsilon$$

folgt sofort die Stetigkeit mit $\|x' - x\| < \delta = \epsilon$.

□

Definition. Eine Menge $A \subseteq V$ heißt *konvex*, wenn für zwei Punkte $u, v \in V$ stets der Punkt $\alpha u + (1 - \alpha)v$ in A liegt $(0 \leq \alpha \leq 1)$.

Approximationssatz. *Sei $\big(V, \langle\,,\,\rangle\big)$ ein Hilbertraum und $A \subseteq V$ konvex und abgeschlossen. Ferner sei*

$$d = d(x, A) := \inf\big\{\, \|x - a\| \;\big|\; a \in A \,\big\}.$$

Dann gibt es genau ein $a_0 \in A$ mit $d = \|x - a_0\|$.

Beweis: Zu $n \in \mathbf{N}$ gibt es ein $a_n \in A$ mit $\|x - a_n\| \leq d + 1/n$, d.h. es existiert eine Folge $(a_n)_{n\geq 1}$ in A mit $\lim \|x - a_n\| = d$ für $x \to \infty$. Nun müssen wir zeigen, daß (a_n) ist eine Cauchy-Folge darstellt. Sei dazu $u := a_n - x$ und $v := a_m - x$. Dann ist

$$\|a_n - a_m\|^2 = 2\|a_n - x\|^2 + 2\|a_m - x\|^2 - \|a_n + a_m - 2x\|^2$$

Also ist

$$\|2(a_n + a_m)/2 - 2x\| = 2\|(a_n + a_m)/2 - x\| \geq 2d.$$

Aus

$$\|a_n - a_m\|^2 \leq 2\|a_n - x\|^2 + 2\|a_m - x\|^2 - (2d)^2$$

folgt dann, daß die Folge eine Cauchy-Folge darstellt, denn aufgrund der Vollständigkeit von V existiert $\lim a_n = a_0 \in \bar{A} = A$ für $n \to \infty$ und damit ist wegen der Stetigkeit der Norm

$$d = \lim \|a_n - x\| = \|\lim a_n - x\| = \|a_0 - x\|.$$

Nun bleibt uns nur noch die Eindeutigkeit zu zeigen. Seien $a_0, b \in A$ und $d = \|x - a_0\| = \|x - b\|$. Dann gilt

$$\begin{aligned}\|a_0 - b\|^2 &= \|(a_0 - x) + (x - b)\|^2 \\ &= 2\|a_0 - x\|^2 + 2\|b - x\|^2 - 4\|(a_0 + b)/2 - x\|^2 \\ &\leq 2d^2 + 2d^2 - 4d^2 = 0.\end{aligned}$$

Also ist $a_0 = b$. □

Korollar. *Sei $A \subseteq V$ konvex und abgeschlossen. Dann existiert genau ein $a_0 \in A$ mit*

$$a_0 = \inf\big\{\, \|a\| \;\big|\; a \in A \,\big\}$$

Definition. Ist $A \subseteq V$ beliebig, so nennen wir

$$A^\perp = \big\{\, v \in V \;\big|\; \langle v, a\rangle = 0 \quad \text{für alle } a \in A \,\big\}$$

orthogonales Komplement.

Satz. *Es gilt:*

(*1*) $\{0\}^\perp = V$ *und* $V^\perp = \{0\}$

(*2*) *Für jedes $A \subseteq V$ ist $A^\perp$ ein abgeschlossener Unterraum von V.*

Beweis:

(*1*) Zum einen ist $\langle v, 0\rangle = 0$ für alle $v \in V$, d.h. $\{0\}^\perp = V$ und zum anderen gilt $\langle x, v\rangle = 0$ für alle $v \in V$ genau dann, wenn $x = 0$, also ist $V^\perp = \{0\}$.

(*2*) Zunächst zeigen wir, daß $A^\perp$ ein Unterraum ist: Offenbar gilt $0 \in A$. Sind $u, v \in A^\perp$ und $\alpha \in K$, so folgt mit der Definition des Skalarprodukts, daß auch $u + \alpha \cdot v \in A^\perp$ ist, d.h. $A^\perp$ ist ein Unterraum.

□

Satz. *Sei $(H, \langle\,,\,\rangle)$ ein Hilbertraum und $A \subseteq H$ abgeschlossen und konvex. Dann existiert zu jedem $x \in H$ genau ein $a_0 \in A$ mit*

$$\|x - a_0\| = d(x, A) = \inf\{\, \|x - a\| \mid a \in A \,\}.$$

Satz. *Sei wieder $(H, \langle\,,\,\rangle)$ ein Hilbertraum, $U \subseteq H$ ein Unterraum, sowie $v_0 \in U$ und $x \in H$. Dann gilt*

$$\|x - v_0\| = d(x, U) \quad \Longleftrightarrow \quad (x - v_0) \in U^\perp.$$

Beweis: Zunächst nehmen wir an, daß $x - v_0 \in U^\perp$ ist und $v \in U$. Dann ist $\|x - v\|^2 = \|x - v_0\|^2 + \|v_0 - v\|^2 = \|x - v_0\|$, denn es ist $x - v_0 \in U^\perp$ und $v_0 - v \in U$.
Sei nun $\|x - v_0\| = d(x, U)$, d.h. $\|x - v_0\|^2 \le \|x - v\|^2$ für alle $v \in U$. Somit ist

$$\begin{aligned} \|x - v_0\|^2 &\le \|x - (v_0 + \lambda v)\|^2 \quad \text{für alle } \lambda \in K \\ &= \|x - v_0\|^2 - 2\,\mathrm{Re}(\lambda\langle v, x - v_0\rangle) + |\lambda|^2\|v\|^2. \end{aligned}$$

Mit $\lambda = \alpha\langle v, x - v_0\rangle$ für $\alpha \in \mathbf{R}$ gilt

$$2\alpha\left|\langle v, x - v_0\rangle\right|^2 \le \alpha^2\left|v, x - v_0\right|^2\|v\|^2.$$

Falls $\alpha \neq 0$ dürfen wir kürzen und erhalten so mit $\alpha \to 0$, daß $\langle v, x - v_0\rangle = 0$. Daraus folgt die Behauptung. □

Projektionssatz. *Ist $(H, \langle\,,\,\rangle)$ ein Hilbertraum und U ein abgeschlossener Unterraum von H, so gilt*

(1) $H = U \oplus U^\perp$ *und*

(2) $(U^\perp)^\perp = U$.

Beweis:

(1) Sei $x \in H$. Nach dem Approximationssatz existiert $v_0 \in U$, so daß $\|x - v_0\| = d(x, U)$. Sei weiter $x - v_0 =: v_1 \in U^\perp$, also $x = v_0 + v_1$. Daraus folgt $H = U + U^\perp$.
Ist nun $v \in U \cap U^\perp$, so ist $\langle v, v\rangle = 0$ und damit auch $v = 0$. Insgesamt gilt also $H = U \oplus U^\perp$.

(2) Zunächst folgt aus der Definition von $U^\perp$, daß $U \subseteq (U^\perp)^\perp$.
Mit $u := v_0 + v_1 \in (U^\perp)^\perp$ und $v_0 \in U, v_1 \in U^\perp$ ist $v_1 = u - v_0 \in (U^\perp)^\perp \cap U^\perp = \{0\}$ und damit erkennen wir die Behauptung als wahr.

□

Satz von Riesz. *Ist $(H, \langle\,,\,\rangle)$ ein Hilbertraum so gelten die beiden folgenden Aussagen:*

(1) Zu jedem $a \in H$ ist $\varphi_a \colon H \to K$, $\varphi_a(v) = \langle v, a\rangle$ eine stetige Linearform.

(2) Jede stetige Linearform auf H hat folgende Form:

$$\lambda(v) = \langle v, a\rangle \quad \textit{für alle } v \in H$$

für ein gewisses $a \in H$ und $\lambda \in H^*$ stetig.

Beweis:

(1) Trivial.

(2) Sei $U = \operatorname{Kern}\lambda$ und o.B.d.A. $\lambda \neq 0$. Dann ist U abgeschlossen, da λ stetig. Somit ist nach dem Projektionssatz $H = U \oplus U^\perp$. Nun wählen wir ein $b \in U^\perp$ mit $\|b\| = 1$ und setzen $\alpha := \lambda(b) \neq 0$. Für $v \in H$ ist damit $\lambda(v)b - \alpha v \in \operatorname{Kern}\lambda = U$. Daraus folgt $\langle\lambda(v)b - \alpha v, b\rangle = 0$, denn $b \in U^\perp$. Daraus folgt $\lambda(v) = \langle v, \alpha b\rangle$ und somit die Behauptung.

□

8.4 Vollständige Orthogonalsysteme (Hilbertraumbasen)

Definition. Sei $(H, \langle\,,\,\rangle)$ ein Prähilbertraum. Dann heißt genau dann $S \subseteq H$ *Orthonormalsystem*, wenn für $s, t \in S$

$$\langle s, t\rangle = \delta_{st}$$

gilt.

Beispiele:

1. Betrachten wir den Prähilbertraum ℓ^2 mit Skalarprodukt

$$\langle a, b\rangle = \sum_{i=1}^{\infty} \alpha_i \beta_i,$$

so ist $S = (e_i)_{i \in \mathbf{N}}$ ein Orthonormalsystem.

2. Zu L^2 mit innerem Produkt

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\overline{g(t)}\,dt$$

ist $S = \{1, \cos t, \sin t, \cos 2t, \sin 2t, \ldots\}$ ein Orthonormalsystem.

◊

Definition. Ein Orthonormalsystem $S \subseteq H$ heißt *vollständig* oder *Hilbertraumbasis*, wenn

$$\overline{\operatorname{Lin} S} = H.$$

Satz. *Jeder Hilbertraum besitzt eine Hilbertraumbasis.*

Beweis: Sei H ein nichtleerer Hilbertraum und $\mathcal{M} = \{\mathcal{S} \subseteq \mathcal{H} \mid \mathcal{S} \text{ Orthonormalsystem}\}$. Offenbar ist $\mathcal{M}$ nichtleer, denn ist $y \in H$, so ist $\{y/\|y\|\} \in \mathcal{M}$. Ist nun K eine Kette in $\mathcal{M}$ und $T := \bigcup_{S \in K}$ die obere Schranke von K und M, so existiert nach dem Lemma von Zorn ein maximales Element $S_0 \in \mathcal{M}$.
Wir wollen nun zeigen, daß S_0 eine Hilbertraumbasis von H ist: Wir zeigen dies durch einen indirekten Beweis. Nehmen wir an, daß S_0 nicht vollständig ist, so ist $\overline{\operatorname{Lin} S_0} \neq H$ und so $H = \overline{\operatorname{Lin} S_0} \oplus (\overline{\operatorname{Lin} S_0})^\perp = \overline{\operatorname{Lin} S_0} \oplus S_0^\perp$. Also ist $H = \overline{\operatorname{Lin} S_0} \oplus S_0^\perp$ mit $S_0^\perp \neq \{0\}$. Damit existiert ein $s \neq 0$ in $S^\perp$ und so ist $S_0 \cup \{s/\|s\|\} \in \mathcal{M}$ im Widerspruch zur Maximalität von S_0. □

Bemerkung: Zu jedem Hilbertraum gibt es einerseits eine sogenannte *algebraische Basis (Hamelbasis)* und zum anderen eine *Hilbertraumbasis*. Im allgemeinen sind beide voneinander unabhängig.
Bemerkung: Sei $(H, \langle\ ,\ \rangle)$ ein Hilbertraum, I eine Indexmenge und $S = (v_i)_{i \in I}$ eine Hilbertraumbasis. Dann gilt: $\sum_{i \in I} v_i = v$ genau dann, wenn es zu jedem $\epsilon > 0$ eine endliche Indexmenge $I_0 \subseteq I$, so daß für alle endlichen Mengen J die Beziehung $I_0 \subseteq J \subseteq I$ gilt, d.h.

$$\left\| \sum_{i \in J} v_i - v \right\| < \epsilon.$$

Für $I = \mathbf{N}$ gilt somit:

$$\sum_{i=1}^{\infty} v_i = v \Longleftrightarrow \left\| \sum_{i=1}^{n} v_i - v \right\| < \epsilon \text{ für alle } n \geq n_0 \in \mathbf{N}.$$

Satz. *Sei wieder $(H, \langle\ ,\ \rangle)$ ein Hilbertraum mit Hilbertraumbasis S. Dann gilt*

$$v = \sum_{s \in S} \langle v, s \rangle s,$$

$$\langle u, v \rangle = \sum_{s \in S} \langle u, s \rangle \langle s, v \rangle,$$

$$\|v\|^2 = \sum_{s \in S} |\langle v, s \rangle|^2 \quad \textit{Parsevalsche-Gleichung,}$$

$$\|v\|^2 \geq \sum_{s \in T} |\langle v, s \rangle|^2 \quad \textit{mit } T \subseteq S \textit{ und } T \textit{ endlich.}$$

Kapitel 9

Multilineare Algebra

9.1 Bilineare Abbildungen und Tensorprodukte

9.1.1 Bilineare Abbildungen

Definition. Es seien U, V und W Vektorräume über dem Körper K. Die Abbildung

$$\varphi\colon U \times V \to W$$

heißt genau dann *bilinear*, wenn

$$\varphi(u + u', v) = \varphi(u, v) + \varphi(u', v), \tag{B1}$$

$$\varphi(u, v + v') = \varphi(u, v) + \varphi(u, v') \text{ und} \tag{B2}$$

$$\alpha\varphi(u, v) = \varphi(\alpha u, v) = \varphi(u, \alpha v) \tag{B3}$$

für alle $u, u' \in U$, $v, v' \in V$ und $\alpha \in K$ gilt.

Bemerkung: Ist v fest und φ eine bilineare Abbildung, so gilt offenbar

$$u \mapsto \varphi(u, v) \in \operatorname{Hom}(U, W).$$

Analog gilt bei festem u

$$v \mapsto \varphi(u, v) \in \operatorname{Hom}(V, W).$$

Definition. Wir definieren

$$\varphi(U, V) := \{\varphi(u, v) \mid u \in U, v \in V\}$$

und das *Bild von* φ durch

$$\operatorname{Bild} \varphi := \operatorname{Lin} \varphi(U, V).$$

Bemerkung: Im Gegensatz zum Bild linearer Abbildungen (Kapitel 3, Lemma 3.1.2) ist das Bild bilinearer Abbildungen im allgemeinen *kein* Unterraum von W!
Beispiel: Ein Beispiel für eine bilineare Abbildung ist:

$$\varphi\colon K^2 \times K^2 \to K^4, \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \mapsto \begin{pmatrix} x_1 y_1 \\ x_1 y_2 \\ x_2 y_1 \\ x_2 y_2 \end{pmatrix}.$$

◊

Lemma 9.1: *Eine bilineare Abbildung ist eindeutig durch die Bilder der Basen bestimmt.*

Beweis: Die Behauptung folgt aus der Definition von bilinearen Abbildungen: Ist $\varphi\colon U \times V \to W$ bilinear und ist $(u_1, \dots, u_n)$ eine Basis von U, sowie $(v_1, \dots, v_m)$ eine Basis von V, so gilt mit $u := \sum \alpha_i u_i$ und $v := \sum \beta_j v_j$

$$\varphi(u, v) = \sum_{i=1}^{n} \alpha_i \varphi(u_i, v) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \varphi(u_i, v_j).$$

Also ist φ durch die Bilder der Basen eindeutig bestimmt. □

Definition. Seien $(u_1, \dots, u_n)$ und $(v_1, \dots, v_m)$ Basen der Vektorräume U und V respektive und (w_{ij}) eine Familie aus dem Vektorraum W, sowie $u := \sum \alpha_i u_i$ und $v := \sum \beta_j v_j$. Dann nennen wir die nach obigem Lemma eindeutig bestimmte bilineare Abbildung

$$\varphi\colon U \times V \to W, \quad (u, v) \mapsto \sum_{ij} \alpha_i \beta_j w_{ij}$$

die *bilineare Fortsetzung* von $(u_i, v_j) \mapsto w_{ij}$.

Definition. Wir definieren

$$\operatorname{Hom}(U, V; W) := \left\{ \varphi\colon U \times V \to W \;\middle|\; \varphi \text{ bilinear} \right\}.$$

Lemma 9.2: $\operatorname{Hom}(U, V; W)$ *ist ein* K*-Vektorraum.*

Beweis: Für $\varphi, \psi \in \operatorname{Hom}(U, V; W)$ und $\alpha \in K$ gilt

$$(\varphi + \psi)(u, v) = \varphi(u, v) + \psi(u, v),$$
$$(\alpha\varphi)(u, v) = \alpha \cdot \varphi(u, v).$$

□

Lemma 9.3: *Es gilt*

$$\operatorname{Hom}(U, V; W) \cong \operatorname{Hom}\big(U, \operatorname{Hom}(V, W)\big).$$

Beweis: Sei $\varphi \in \operatorname{Hom}(U, V; W)$. Dann ist $\varphi_u\colon V \to W, v \mapsto \varphi(u, v)$ mit festem $u \in U$ linear, d.h. $\varphi_u \in \operatorname{Hom}(V, W)$. Betrachten wir nun die lineare Abbildung

$$\Phi\colon U \to \operatorname{Hom}(V, W), \quad u \mapsto \varphi_u.$$

Damit definieren wir nun $F\colon \operatorname{Hom}(U, V; W) \to \operatorname{Hom}\big(U, \operatorname{Hom}(V, W)\big)$ durch $F(\varphi) := \Phi$. Wir müssen nun zeigen, daß F ein Isomorphismus ist.

1. F ist aufgrund der Konstruktion linear.
2. F ist injektiv, denn aus $F(\varphi) = 0$ folgt $\Phi = 0$ und damit $\Phi(u) = \varphi_u = 0$ für alle u. Daraus folgt schließlich $\varphi(u, v) = 0$ (für alle u, v) und $\varphi = 0$.
3. F ist surjektiv, denn betrachten wir $g\colon U \to \operatorname{Hom}(V, W)$ linear mit

   $$\big[g(u)\big](v) =: \varphi(u, v),$$

   wobei φ bilinear ist, so gilt $F(\varphi) = g$.

□

Korollar. *Sind U, V und W endlichdimensionale Vektorräume, so gilt*

$$\operatorname{Dim} \operatorname{Hom}(U, V; W) = \operatorname{Dim} U \cdot \operatorname{Dim} V \cdot \operatorname{Dim} W.$$

Bilinearformen

Definition. Sei V ein Vektorraum über dem Körper K. Eine bilineare Abbildung der Form

$$\varphi\colon V \times V \to K$$

heißt *Bilinearform.*

Beispiele:

1. Die Abbildung $\varphi(x, y) = 3x_1y_2 - 4x_3y_1 + x_2y_3$ ist über dem K^3 eine Bilinearform.
2. Seien $\lambda, \mu \in V^*$. Dann ist

   $$\varphi(u, v) := \lambda(u) \cdot \mu(v)$$

   eine Bilinearform.
3. Über dem $K^{n\times n}$ ist mit festem $A \in K^{n\times n}$

   $$\varphi(X, Y) := \operatorname{Sp}(XAY)$$

   eine Bilinearform.
4. Ein weiteres Beispiel für eine Bilinearform ist das kanonische oder natürliche Skalarprodukt des K^n:

   $$\langle x, y \rangle := \sum_{i=1}^{n} x_i y_i = x^T y.$$

   Hierfür gilt $\langle x, y \rangle = \langle y, x \rangle$ und gilt $\langle x, y \rangle = 0$ für alle y, so folgt daraus $x = 0$.
5. Sei ein $A \in K^{n\times n}$ gegeben. Dann ist durch

   $$\varphi_A(x, y) := \langle x, Ay \rangle = x^T A y$$

   eine Bilinearform definiert. Im allgemeinen ist $\varphi_A(x, y) \neq \varphi_A(y, x)$ und es gibt ein u mit $\varphi_A(u, x) = 0$ für alle x.

◊

Im folgenden sei φ stets eine Bilinearform über dem endlich erzeugten K-Vektorraum V.

Definition. Ist $B = (v_1, \ldots, v_n)$ eine Basis von V, so bezeichnen wir die *Darstellungsmatrix* von φ bezüglich B mit

$$ {}_B[\varphi] := \big(\varphi(v_i, v_j)\big)_{i,j}. $$

Basisdarstellung: Sei V ein Vektorraum mit Basis $B = (v_1, \ldots, v_n)$. Für jede Bilinearform φ über V und $u := \sum \alpha_i v_i$, $v := \sum \beta_j v_j$ gilt dann

$$ \varphi(u, v) = \sum_{i,j} \alpha_i \beta_j \varphi(v_i, v_j) = \big\langle {}_B[u],\ {}_B[\varphi] \cdot {}_B[v] \big\rangle. $$

Denn es ist ja

$$ {}_B[\varphi] \cdot {}_B[v] = \begin{pmatrix} \sum \varphi(v_1, v_j)\beta_j \\ \vdots \\ \sum \varphi(v_n, v_j)\beta_j \end{pmatrix}. $$

Basiswechsel: Ist $C = (w_1, \ldots, w_n)$ eine weitere Basis von V, so gilt nach dem Darstellungssatz ${}_C[v] = {}_C[\mathrm{Id}]_B \cdot {}_B[v]$. Es ist also

$$ \varphi(u, v) = \big\langle {}_C[u],\ {}_C[\varphi] \cdot {}_C[v] \big\rangle = \big\langle {}_B[u],\ {}_C[\mathrm{Id}]_B{}^{T} \cdot {}_C[\varphi] \cdot {}_C[\mathrm{Id}]_B \cdot {}_B[v] \big\rangle. $$

Daraus können wir folgern, daß

$$ {}_B[\varphi] = {}_C[\mathrm{Id}]_B{}^{T} \cdot {}_C[\varphi] \cdot {}_C[\mathrm{Id}]_B $$

ist.

Beispiel: Betrachten wir nun über dem K^n die Bilinearform $\varphi_A : (x, y) \mapsto x^T A y$. Dabei sei $A \in K^{n\times n}$. Mit der Basis $B = (e^1, \ldots, e^n)$ gilt dann offenbar

$$ A = {}_B[\varphi]. $$

Für die Darstellung von φ bezüglich einer Basis $C = (w_1, \ldots, w_n)$ des K^n erhalten wir dann:

$$ {}_C[\varphi] = C^T A C. $$

$\Diamond$

Definition. Für eine Bilinearform φ über dem Vektorraum V und den Unterraum $U \subseteq V$ definieren wir

- den *Linkskern* $U^\perp := \big\{ v \in V \,\big|\, \varphi(u, v) = 0 \quad \text{für alle } u \in U \big\}$ und
- den *Rechtskern* ${}^\perp U := \big\{ u \in V \,\big|\, \varphi(u, v) = 0 \quad \text{für alle } v \in U \big\}$.

Im allgemeinen sind der Linkskern und der Rechtskern verschieden.

Definition. Eine Bilinearform φ über dem Vektorraum V heißt *nicht ausgeartet*, wenn

$$ V^\perp = \{0\} \text{ und } {}^\perp V = \{0\} $$

gilt.

Satz. *Ist B eine Basis von V, so sind folgende Aussagen äquivalent:*

(1) φ *ist nicht ausgeartet.*

(2) $V^\perp = \{0\}$

(3) ${}^\perp V = \{0\}$

(4) ${}_B[\varphi]$ *ist invertierbar.*

Beweis:

(3) $\Longrightarrow$ *(4)*: Aus $\varphi(u, v) = \big\langle {}_B[u],\ {}_B[\varphi] \cdot {}_B[v] \big\rangle = 0$ für alle $u \in V$ folgt zum einen ${}_B[\varphi] \cdot {}_B[v] = 0$ und zum anderen nach Voraussetzung $v = 0$, d.h. ${}_B[\varphi]$ ist invertierbar.

(2) $\Longleftrightarrow$ *(4)*: $F \in \mathrm{GL}(n, K) \Longleftrightarrow F^T \in \mathrm{GL}(n, K)$.

(3) $\Longleftrightarrow$ *(4)*: $\varphi(u, v) = \langle [u], F[v] \rangle = 0$ für alle $u \in V \quad \Longleftrightarrow \quad F[v] = 0$.

□

Darstellungssatz von Riesz. *Sei φ eine nicht ausgeartete Bilinearform über dem endlichdimensionalen Vektorraum V. Dann gibt es zu jeder Linearform $\lambda \in V^*$ genau ein $l \in V$ mit*

$$ \lambda(v) = \varphi(l, v) \quad \textit{für alle } v. $$

Beweis: Sei $u \in V$ fest, dazu $\varphi_u(v) := \varphi(u, v)$ ($\varphi_u : V \to K$), $\varphi_u \in V^*$. Die Abbildung $u \mapsto \varphi_u$ ($V \to V^*$) ist linear und injektiv, denn

$$ \varphi_u = 0 \Longrightarrow \varphi(u, v) = 0 \quad \text{für alle } v \qquad \Longrightarrow \qquad u = 0. $$

Da $\dim V = \dim V^* < \infty$ ist diese Abbildung sogar bijektiv. Insbesondere kommt jedes $\lambda \in V^*$ als Bild vor, d.h. zu $\lambda \in V^*$ existiert ein $l \in V$ mit $\lambda = \varphi_l$ bzw. gilt

$$ \lambda(v) = \varphi_l(v) = \varphi(l, v). $$

Da $\dim V < \infty$ und φ u.a. eine Bilinearform ist, gibt es zu jedem $\lambda \in V^*$ genau ein $l \in V$ mit

$$ \lambda(v) = \varphi(l, v) \quad \text{für alle } v. $$

□

Korollar. *Zu jeder Bilinearform φ über dem endlich erzeugtem Vektorraum V mit Basis $B = (v_1, \ldots, v_n)$ existiert genau eine Basis $C = (w_1, \ldots, w_n)$, so daß $\varphi(w_i, v_j) = \delta_{ij}$ gilt.*

Beweis: Zu B wähle duale Basis $B^* = (\lambda_1, \ldots, \lambda_n)$ von V^*:

$$ \lambda_i(v_j) = \delta_{ij}. $$

Nach vorherigem Satz existiert ein $w_i \in V$ mit $\lambda_i(v) = \varphi(w_i, v)$ für alle v. Insbesondere gilt

$$ \delta_{ij} = \lambda_i(v_j) = \varphi(w_i, v_j). $$

□

Definition. Die nach obigem Korollar eindeutig bestimme Basis C des Vektorraums V heißt die zu B *duale Basis.*

Definition. Die Bilinearform φ über V heißt

- *symmetrisch*, wenn gilt
$$\varphi(u,v) = \varphi(v,u) \quad \text{für alle } u,v \in V \quad \Longleftrightarrow \quad {}_B[\varphi] = {}_B[\varphi]^T.$$
- *schiefsymmetrisch*, wenn gilt
$$\varphi(u,v) = -\varphi(v,u) \quad \text{für alle } u,v \in V \quad \Longleftrightarrow \quad {}_B[\varphi] = -{}_B[\varphi]^T.$$

Lemma 9.4: *Ist φ eine Bilinearform über dem Vektorraum V mit $\varphi(u,v) = \varphi(v,u) = 0$, so ist φ entweder symmetrisch oder es gilt $\varphi(u,u) = 0$ für alle $u \in V$.*

Lemma 9.5: *Aus $\varphi(v,v) = 0$ für alle $v \in V$ folgt $\varphi(u,v) = -\varphi(v,u)$ für alle $u,v \in V$.*

Beweis:
$$0 = \varphi(u+v,u+v) = \varphi(u,u) + \varphi(v,u) + \varphi(u,v) + \varphi(v,v) = \varphi(v,u) + \varphi(u,v).$$
Umgekehrt gilt (falls $2 \neq 0$ in K!):
$$\varphi(u,v) = -\varphi(v,u) \Longrightarrow \varphi(v,v) = -\varphi(v,v) \Longrightarrow 2\varphi(v,v) = 0 \Longrightarrow \varphi(v,v) = 0.$$
□

Aufgrund dieses Satzes wollen wir im folgenden nur die Fälle betrachten, wo φ entweder symmetrisch oder schiefsymmetrisch ist.

Satz. *Sei $U \subseteq V$ ein Unterraum. Dann gilt*
$$V = U \oplus U^\perp.$$

Beweis: Seien $v \in V$ beliebig, $\varphi_v\colon U \to K$, $\varphi_v(u) = \varphi(v,u)$ ist in U^*, da φ auf U nicht ausgeartet ist, gibt es ein $u' \in U$ mit $\varphi(v,u) = \varphi(u',u)$ für alle $u \in U$. Daraus folgt
$$\varphi(v-u',u) = 0 \quad \text{für alle } u \in U,$$
bzw. $(v-u') =: u_0 \in U^\perp$, somit $v = u_0 + u'$ und damit $V = U + U^\perp$.
Diese Summe ist direkt, da aus $w \in U \cap U^\perp \Longrightarrow \varphi(w,u) = 0$ für alle $u \in U$ direkt $w = 0$ (f ist auf U nicht ausgeartet). □

Satz. *Zu jeder symmetrischen Bilinearform φ über V existiert eine Basis $B = (v_1,\dots,v_n)$ mit*
$$\varphi(v_i,v_j) = 0 \qquad \text{für } i \neq j.$$

Beweis über vollständige Induktion: Für $n = 1$ ist nichts zu beweisen.
Induktionsschritt von $n-1$ nach n: Sei o.E. $\varphi \neq 0$. Dann existiert ein $u_1 \in V$ mit $\varphi(u_1,u_1) \neq 0$. Also ist φ auf $U = Kn_1$ nicht ausgeartet und es gilt
$$V = Kn_1 \oplus (Kn_1)^\perp.$$
Nun wenden wir die Induktionsvoraussetzung auf
$$\varphi\colon (Kn_1)^\perp \times (Kn_1)^\perp$$
an. □

Satz. *Zu jeder schiefsymmetrischen Bilinearform φ existiert eine Basis B von V mit*
$${}_B[\varphi] = \begin{pmatrix} \begin{pmatrix} 0 & \alpha_1 \\ -\alpha_1 & 0 \end{pmatrix} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \begin{pmatrix} 0 & \alpha_s \\ -\alpha_s & 0 \end{pmatrix} & 0 \\ 0 & \cdots & 0 & \begin{pmatrix} 0 & & \\ & \ddots & \\ & & 0 \end{pmatrix} \end{pmatrix}.$$

Beweis über vollständige Induktion: Für $n = 1$ ist nichts zu beweisen.
Induktionsschritt von $n-1$ nach n: Sei o.E. $\varphi \neq 0$. Dann existieren $a_1, b_1 \in V$ mit $\alpha_1 := \varphi(a_1,b_1) \neq 0$ (beachte, daß a_1 und b_1 linear unabhängig sind; dies folgt aus schiefsymmetrisch!). Damit ist φ auf $Ka_1 \oplus Kb_1$ nicht ausgeartet und es gilt:
$$V = (Ka_1 \oplus Kb_1) \oplus (Ka_1 \oplus Kb_1)^\perp,$$
wende nun die Induktionsvoraussetzung auf den zweiten Summanden an, $v_1 := a_1$, $v_2 := b_1$. □

Bemerkung: Zu jeder Basis $B = (v_1,\dots,v_n)$ existiert eine Dualbasis $C = (w_1,\dots,w_n)$, so daß $\varphi(v_i,w_j) = \delta_{ij}$. Auch über $\bar{\mathbf{R}}$ werden indefinite Bilinearformen zugelassen.
Beispiel: Wir betrachten nun den $\mathbf{R}^4$: Sei
$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ t_1 \end{pmatrix} \text{ und } y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ t_2 \end{pmatrix},$$
sowie
$$\varphi(x,y) = x_1y_1 + x_2y_2 + x_3y_3 - (ct_1ct_2).$$
Dann ist
$$\varphi(x,x) = x_1^2 + x_2^2 + x_3^2 - (ct)^2 \qquad \text{Lorenz-Metrik.}$$
◊

Satz. *Zu einem Endomorphismus f gibt es auch in der allgemeinen Situation eine adjungierte Abbildung f^*.*

Beweis: Seien V ein Vektorraum, $B = (v_1,\dots,v_n)$ eine Basis und $C = (w_1,\dots,w_n)$ eine Dualbasis, sowie $u,v \in V$. Aus der eindeutigen Darstellung $u = \sum \alpha_i v_i$, folgt $fu = \sum \varphi(u,v_i) f w_i$ und daraus $\varphi(fu,v) = \sum \varphi(u,v_i)\varphi(fw_i,v)$ mit $f^*v = \sum \varphi(fw_i,u)v_i$. □

Es gilt:

$$f^{**} = f \tag{9.1}$$
$$(gf)^* = f^* g^* \tag{9.2}$$
$$(\alpha f + g)^* = \alpha f^* + g^* \tag{9.3}$$

9.1.2 Bilineare Tensorprodukte

Sei $\varphi\colon U \times V \to W$ bilinear und $f\colon W \to W'$ linear. Dann ist $\tilde{\varphi}\colon U \times V, \tilde{\varphi}(u,v) := (f \circ \varphi)(u,v)$ ebenfalls bilinear. Ferner habe der Vektorraum U die Basis $B = (u_1, \ldots, u_n)$ und der Vektorraum V die Basis $C = (v_1, \ldots, v_m)$. Nun wählen wir einen Vektorraum T der Dimension $m \cdot n$ mit Basis $(t_{ij}, 1 \le i \le n, 1 \le j \le m)$ und setzen $(u_i, v_j) \mapsto t_{ij}$ bilinear fort. Dann heißt die Abbildung $\tau\colon U \times V \to T, \tau(u,v) = \sum \alpha_i \beta_j t_{ij}$ *Tensorprodukt* mit $u = \sum \alpha_i u_i$ und $v = \sum \beta_j v_j$. Nun sei $\varphi\colon U \times V \to W$ bilinear und $w_{ij} = \varphi(u_i, v_j) \in W$. Sei $t_{ij}\colon T \to W$ die lineare Fortsetzung von t_{ij} nach w_{ij} und $f\Big(\sum \alpha_{ij} t_{ij}\Big) = \sum \alpha_{ij} w_{ij}$. Es gilt dann $(f \circ \tau)(u,v) = \varphi(u,v)$ für alle $u \in U$ und $v \in V$.

Definition. Seien U, V und T Vektorräume über dem Körper K und

$$\tau\colon U \times V \to T$$

bilinear. Dann heißt (T, τ) *Tensorprodukt*, wenn

$$T = \mathrm{Lin}\big\{\, \tau(u,v) \;\big|\; u \in U \text{ und } v \in V \,\big\},$$

und es zu jedem bilinearen $\varphi\colon U \times V \to W$ genau eine lineare Abbildung $f\colon T \to W$ mit $(f \circ \tau)(u,v) = \varphi(u,v)$ gibt.

Satz. *Zu je zwei Vektorräumen U mit Basis $(u_1, \ldots, u_n)$ und V mit Basis $(v_1, \ldots, v_m)$ existiert ein Tensorprodukt (T, τ). Dabei ist $(t_{ij}, 1 \le i \le n, 1 \le j \le m)$ die Basis von T und $\tau\colon U \times V \to T$ die bilineare Fortsetzung von $(u_i, v_j) \mapsto t_{ij}$.*

Satz. *Sind (T, τ) und (T', τ') Tensorprodukte der Vektorräume U und V, so gibt es einen eindeutig bestimmten Isomorphismus $g\colon T \to T'$ mit $(g \circ \tau)(u,v) = \tau'(u,v)$ für alle $u \in U$ und $v \in V$.*

Beweis: Es gibt genau einen Homomorphismus g mit $g\colon T \to T'$ und $g \circ \tau = \tau'$. Durch Vertauschung erkennen wir die Bijektivität von g, denn

$$g' \circ g\colon T \to T \text{ und } g' \circ g \circ \tau = g' \circ \tau' = \tau, \text{ sowie}$$
$$g \circ g'\colon T' \to T' \text{ und } g \circ g' \circ \tau' = \tau'$$

Damit ist $\mathrm{Id} = gg' = g'g$, d.h. g und g' sind zwei zueinander inverse Isomorphismen. Damit ist die Behauptung verifiziert. □

Da nach obigem Satz alle Tensorprodukte bis auf Isomorphie gleich sind, führen wir folgende Bezeichnung ein:

Definition. Sind U und V Vektorräume, so schreiben wir

$$T = U \otimes V$$
$$\tau(u,v) = u \otimes v \quad \text{für } u \in U, v \in V$$

und sagen U Tensor V.

Wir geben nun ein paar einfache Rechenregeln an. Dazu seien $u, u' \in U$, $v, v' \in V$ und $\alpha \in K$, wobei U und V Vektorräume sind.

$$(u + u') \otimes v = u \otimes v + u' \otimes v,$$
$$u \otimes (v + v') = u \otimes v + u \otimes v',$$
$$\alpha(u \otimes v) = \alpha u \otimes v = u \otimes \alpha v.$$

Satz. *Ist $(u_1, \ldots, u_n)$ eine Basis des Vektorraumes U und $(v_1, \ldots, v_m)$ eine Basis des Vektorraumes V, so bildet $(u_i \otimes v_j)_{1 \le i \le n,\, 1 \le j \le m}$ eine Basis von $U \otimes V$.*

Beweis: Sei (T, τ) ein beliebiges Tensorprodukt mit Basis (t_{ij}) und $(u_i, v_j) \mapsto t_{ij}$ gegeben. Dann existiert ein Isomorphismus g mit $g(u_i \otimes v_j) = t_{ij}$. Daraus folgt die Behauptung. □

Korollar. *Sind $(u_1, \ldots, u_n)$ und $(v_1, \ldots, v_m)$ linear unabhängig, so ist auch $(u_i \otimes v_j)$ linear unabhängig.*

Korollar. *Genau dann ist $(u \otimes v) = 0$, wenn $u = 0$ oder $v = 0$.*

Beispiele zu Tensorprodukten:

1. Wir betrachten die Vektorräume $U = K^m$ und $V = K^n$ mit ihren natürlichen Basen. Dazu wählen wir $T = K^{m \times n}$ mit der Basis $t_{ij} = E_{ij} = e^i \otimes e^j = e^i (e^j)^T$ und die bilineare Fortsetzung

$$x \otimes y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (y_1, \ldots, y_n) = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \ldots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \ldots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \ldots & x_m y_n \end{pmatrix}$$

2. Sei K ein Körper. Dann betrachten wir die beiden K-Vektorräume K und V. Zum einen gilt $K \otimes V = V \otimes K = V$, zum anderen gilt mit

$$\tau\colon K \times V \to V, (\lambda, x) \mapsto \lambda x \text{ bilinear,}$$
$$\varphi\colon K \times V \to W \text{ bilinear und}$$
$$h\colon V \to W, h(x) := \varphi(1, x) \text{ linear, daß}$$
$$\varphi(\lambda, x) = \lambda\varphi(1, x) = \lambda h(x) = h(\lambda x) = h\big(\tau(\lambda, x)\big) = (h \circ \tau)(\lambda, x) = h(\lambda \otimes x) \text{ ist.}$$

3. Nun betrachten wir einen beliebigen $\mathbf{R}$-Vektorraum V: Es gilt $\mathbf{C} \otimes_{\mathbf{R}} V = \widetilde{V}_{\mathbf{C}}$. Weiter ist $\tau(a + ib, v) := (av, bv)$, $\varphi\colon \mathbf{C} \times V \to W$ ist $\mathbf{R}$-bilinear und $h\colon \widetilde{V}_{\mathbf{C}} \to W, h(av + bv) := a + ib, v)$, sowie $h \circ \tau = \varphi$.

◊

Satz. *Sind U, V und W Vektorräume, so gilt*

$$\mathrm{Hom}\big(U, \mathrm{Hom}(V,W)\big) \cong \mathrm{Hom}(U \otimes V, W).$$

Beweis: Sei $\varphi\colon U \times V \to$ bilinear. Dann gibt es genau einen Homomorphismus $h_\varphi\colon U \otimes V \to W$ mit $\varphi(u,v) = h_\varphi(u \otimes v)$. Ferner sei $h\colon \mathrm{Hom}\big(U, \mathrm{Hom}(V,W)\big) \to \mathrm{Hom}(U \otimes V, W)$ mit $\varphi \mapsto h_\varphi$ gegeben. Wir zeigen nun, daß h ein Isomorphismus ist: Wegen

$$\begin{aligned} h_{\alpha\varphi+\psi}(u,v) = (\alpha\varphi + \psi)(u,v) = \alpha\varphi(u,v) + \psi(u,v) &= \alpha h\varphi(u \otimes v) + h\psi(u \otimes v) \\ &= (\alpha h\varphi + h\psi)(u \otimes v) \end{aligned}$$

ist h linear.
Ferner ist h injektiv. Denn aus $h(\varphi) = 0$ folgt $h_\varphi(u \otimes v) = 0$ für alle u und v, d.h. $\mathrm{Kern}\, h = \{0\}$.
Schließlich erkennen wir, daß h auch surjektiv ist. Es folgt nämlich mit $A\colon U \otimes V \to W$ linear und $\varphi(u,v) := A(u \otimes v)$ bilinear, daß $h_\varphi(u \otimes v) = A$ ist, d.h. $h_\varphi = A$. □

Korollar. *Sind U und V Vektorräume, so gilt*

$$(U \otimes V)^* \cong \mathrm{Hom}(U, V^*).$$

9.2 Multilineare Abbildungen und Tensorprodukte

9.2.1 Multilineare Abbildungen

Definition. Sind U_1 $U_2, \ldots, U_r$ und V Vektorräume und ist $f\colon U_1 \times U_2 \times \cdots \times U_r \to V$, so heißt f *multilinear* oder r-*linear*.
Die Menge aller r-linearen Abbildungen bezeichnen wir wie folgt:

$$\mathrm{Hom}(U_1, U_2, \ldots, U_r, V) := \big\{\, f\colon U_1 \times U_2 \times \cdots \times U_r \to V \;\big|\; f \text{ ist } r\text{-linear} \,\big\}.$$

Die 1-, 2- und 3-linearen Abbildungen bezeichnen wir respektive als *lineare*, *bilineare* und *trilineare* Abbildungen.

Beispiele:

1. Die Determinante ist ein bereits bekannte multilineare Abbildung (siehe Kapitel 4).
2. Die Abbildung

$$K^{m\times n} \times K^{n\times p} \times K^{p\times q} \to K^{m\times q}, \quad (A,B,C) \mapsto ABC$$

ist ebenfalls multilinear.

◊

Lemma 9.6: *Ist $\varphi \in \mathrm{Hom}(U_1, \ldots, U_r, V_0)$ und $\psi \in \mathrm{Hom}(V_0, \ldots, V_l, W)$, so ist*

$$(u_1, \ldots, u_r, v_1, \ldots, v_l) \mapsto \psi\big(\varphi(u_1, \ldots, u_r), v_1, \ldots v_l\big)$$

$(r+l)$-linear.

Beweis: Die Behauptung folgt sofort aus der Definition von multilinearen Abbildungen. □

9.2.2 Multilineare Tensorprodukte

Definition. Sind U_1, U_2, …, U_r und V Vektorräume und ist $\tau\colon U_1 \times U_2 \times \cdots \times U_r \to V$ eine r-lineare Abbildung, so heißt τ *Tensorprodukt*, wenn

$$\mathrm{Lin}\big(\tau(U_1 \times U_2 \times \cdots \times U_r)\big) = V$$

gilt und es genau einen Homomorphismus $h\colon V \to W$ mit $h \circ \tau = \varphi$ gibt.

Satz. *Es existiert bis auf Isomorphie genau ein Tensorprodukt von endlich vielen Vektorräumen.*

Definition. Wir führen folgende Bezeichnungen ein:

$$V = U_1 \otimes U_2 \otimes \cdots \otimes U_r := \bigotimes_{i=1}^{r} U_i,$$

$$\tau(u_1, u_2, \ldots, u_r) = u_1 \otimes u_2 \otimes \cdots \otimes u_2 := \bigotimes_{i=1}^{r} U_i.$$

Diese Definition ist nach dem vorangegangenen Satz wohldefiniert.

Satz. *Sind X, Y und Z Vektorräume über dem Körper K, so gilt*

$$X \otimes Y \otimes Z \cong X \otimes (Y \otimes Z) \cong (X \otimes Y) \otimes Z.$$

Korollar. *Für die K-Vektorräume $U_1, U_2, \ldots, U_r, U_{r+1}, \ldots, U_l$ gilt*

$$U_1 \otimes U_2 \otimes \cdots \otimes U_l \cong (U_1 \otimes U_2 \otimes \cdots \otimes U_r) \otimes (U_{r+1} \otimes U_{r+2} \otimes \cdots \otimes U_l).$$

9.3 Tensorielle Produkte linearer Abbildungen

Satz. *Es existiert genau ein Monomorphismus*

$$\begin{aligned} \tau\colon\ & \mathrm{Hom}(V_1, W_1) \otimes \mathrm{Hom}(V_2, W_2) \otimes \cdots \otimes \mathrm{Hom}(V_n, W_n) \\ & \to \mathrm{Hom}(\, V_1 \otimes V_2 \otimes \cdots \otimes V_n,\ W_1 \otimes W_2 \otimes \cdots \otimes W_n \,) \end{aligned}$$

mit $\tau(\varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_n) = \varphi_1 \mathbin{\underline{\otimes}} \varphi_2 \mathbin{\underline{\otimes}} \cdots \mathbin{\underline{\otimes}} \varphi_n$.
Sind die Vektorräume $V_1, V_2, \ldots V_n$ und $W_1, W_2, \ldots, W_n$ endlichdimensional, so ist τ sogar ein Isomorphismus.

9.4 Tensormultiplikation

Definition. Ist V ein beliebiger Vektorraum und V^* sein Dualraum, so bezeichnet

$$V^{[p,q]} := \underbrace{V \otimes \cdots \otimes V}_{p \text{ mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q \text{ mal}}$$

das Tensorprodukt von p mal den Vektorraum V und q mal den Vektorraum V^*.
Die Elemente des Tensorraums $V^{[p,q]}$ nennen wir p-*fach kontravariante* und q-*fach kovariante Tensoren* oder auch *Tensoren der Stufe* (p,q).

9.5 Algebren

Im folgenden betrachten wir nun den Sonderfall $U = V = W$:

Definition. Eine bilineare Abbildung $\varphi\colon V \times V \to V$ heißt *Produkt auf* V. Anstatt von $\varphi(u, v)$ schreiben wir $u \cdot v$.

Bemerkung: Die Linearität von φ bedeutet

$$(\alpha u + u') \cdot v = \alpha u \cdot v + u' \cdot v,$$
$$u \cdot (\alpha v + v') = \alpha u \cdot v + u \cdot v'.$$

Definition. Ein Vektorraum über einem Körper K mit einem Produkt heißt eine K-*Algebra*.

Beispiele:

1. Sei V der *triviale Vektorraum* K.
2. Der Vektorraum $K^{n\times n}$ mit $(A, B) \mapsto AB$ ist eine *assoziative Algebra*.
3. Der Vektorraum $K^{n\times n}$ mit der *Lie-Klammer* (auch *Kommutator* genannt):

   $$(A, B) \mapsto [A, B] := AB - BA.$$

   Hierfür gilt:

   $$[A, B] = -[B, A],$$
   $$\big[A, [B, C]\big] + \big[B, [C, A]\big] + \big[C, [A, B]\big] = 0 \qquad \text{(Jacobi-Identität)}.$$

   Die so definierte Algebra heißt auch *Lie-Algebra*.

◊

Beispiele: Die ersten drei Beispiele stellen assoziative Algebren mit Eins dar.

1. $K^{n\times n}$ mit $(A, B) \mapsto AB$.
2. $\operatorname{End} V$ mit $(f, g) \mapsto f \circ g$.
3. $T(V)$.
4. $A = K^{n\times n}$ mit einem neuem Produkt $u \cdot v := uv + vu$, $(A, \cdot) =: A^+$.
5. $A^- := (A, [\,])$ mit $[u, v] := uv - vu$. A^- heißt *Lie-Algebra*.

◊

Definition. Seien A, B zwei K-Algebren. Dann heißt die Abbildung $f\colon A \to B$ *Algebra-Homomorphismus* oder einfach *Homomorphismus*, wenn gilt:

(*1*) f ist linear,

(*2*) $f(uv) = f(u)f(v)$.

Mit $\operatorname{Kern} f$ bezeichnen wir folgende Menge:

$$\operatorname{Kern} f := \big\{\, u \in A \;\big|\; f(u) = 0 \,\big\}.$$

Dabei ist $\operatorname{Kern} f$ ein Unterraum von A.

Definition. Eine Teilmenge $I \subseteq A$ heißt *Ideal*, wenn gilt:

(*1*) $I \subseteq A$ ist Unterraum.

(*2*) Aus $u \in I$ und $v \in A$ folgt $uv, vu \in I$.

Damit ist z.B. der Kern von f ein Ideal!

Definition. Zum Ideal bilden wir den Faktorraum

$$A/I := \big\{\, v + I \;\big|\; v \in A \,\big\}.$$

Ferner ist dies ein Vektorraum, denn es gilt

$$(v + I) + (v' + I) := (v + v') + I,$$
$$\alpha(v + I) := \alpha v + I.$$

Weiter ist

$$(v + I)(v' + I) := vv' + I$$

ein wohldefiniertes Produkt. Damit stellt der Faktorraum die *Faktoralgebra* dar.

Tensorprodukte

Die Tensorprodukte $U \otimes V$, $V_1 \otimes V_2 \otimes \cdots \otimes V_k$ sind universell bzgl. K-linearen Abbildungen.

$$\varphi(v_1, \ldots, v_k) = F(v_1 \otimes \cdots \otimes v_k).$$

Es gilt:

$$T(V) = \bigoplus_{k\geq 0} V^{\otimes k},$$

mit

$$\Big(\sum_{k\geq 0} t_k\Big)\Big(\sum_{s\geq 0} l_s\Big) = \sum_{k,s} t_k \otimes l_s.$$

$T(V)$ ist eine assoziative K-Algebra mit Eins.

Äußere Algebra oder Graßmann-Algebra

Definition. Eine Abbildung $\varphi\colon V_1 \times \cdots \times V_k \to W$ heißt *alternierend* K-*linear*, falls gilt:

1. φ ist K-linear.
2. $\varphi(v_1, \ldots, v_k) = 0$ falls es i, j gibt, mit $i \neq j$ und $v_i = v_j$.

Beispiel: $\det(a^1, \ldots, a^n)$.
Wir spezialisieren zu $V_K = V$.

Definition. Sei V ein K-Vektorraum. Ein Vektorraum L zusammen mit einer alternierenden K-linearen Abbildung $\lambda\colon V \times \cdots \times V \to L$ heißt k-*faches äußeres Produkt von* V, wenn es zu einem alternierenden K-linearen φ ein f gibt, mit $\varphi = f\lambda$.

Wie für Tensorprodukte ist (L, λ) bis auf natürliche Isomorphie *eindeutig bestimmt.*
Existenz: Sei V gegeben. Betrachte in $V^{|k|}$ den Unterraum U_k, erzeugt von allen $v_1 \otimes \cdots \otimes v_k$ „mit zwei gleichen v_i". Sei

$$L := V^{|k|}/U_k$$

und $\lambda\colon V \times \cdots \times V \to L$, $\lambda(v_1, \ldots, v_k) := v_1 \otimes \cdots \otimes v_k + U_k$ K-linear, alternierend. (L, V) hat die gewisse universelle Eigenschaft. $\pi(t) = t + U_k$. $F(v_1 \otimes \cdots \otimes v_k) = \varphi(v_1, \ldots, v_k)$, da φ alternierend: $U_k \subseteq \operatorname{Kern} f$. $f(t + U_k) := F(t)$, $\varphi = f \circ \lambda$. Deshalb ist $f(t + U_k) := F(t)$ wohldefiniert.
Graßmann's Idee war, ein äußeres Produkt zu gegebenen Vektoren $a, b, c \in \mathbf{R}^3$ zu finden, daß das *Parallelogramm* bzw. den *Spat* repräsentiert, das von diesen Vektoren aufgespannt wird.
Sei $L_k = V^{|k|}/U_k$. $\lambda(v_1, \ldots, v_k) = v_1 \otimes \cdots \otimes v_k + U_k =: v_1 \wedge v_2 \wedge \ldots \wedge v_k$ (ließ Dach oder Zack!).
Einführung von „$\wedge$" von hinten: Seien K, $V = L_1$, $L_2 = V^{|2|}/U_2$ und $v_1, v_2 \in V$ gegeben. Wir setzen

$$v_1 \wedge v_2 := v_1 \oplus v_2 + U_2.$$

Für $k = 3$ gilt: $L_3 = V^{|3|}/U_3$ und $v_1, v_2, v_3 \in V$.

$$L_3 \ni v_1 \wedge v_2 \wedge v_3 =: v_1 \otimes v_2 \otimes v_3 + U_3.$$

Übliche Bezeichnung:

$$L_k = \bigwedge^k V$$

heißt k-*faches äußeres Produkt.*
Eigenschaften:

1. $u \wedge v = -v \wedge u$ $(k = 2)$.
2. $v_1 \wedge \ldots \wedge v_i \wedge \ldots \wedge v_j \wedge \ldots \wedge v_k = -v_1 \wedge \ldots \wedge v_j \wedge \ldots \wedge v_i \wedge \ldots \wedge v_k$.
3. Sei $\sigma \in S_n$ eine Permutation mit der Signatur l (die Anzahl der Vertauschungen). Dann gilt:

$$v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(n)} = (-1)^l v_1 \wedge \ldots \wedge v_n.$$

4. Dim $V = n$, $m > n$. Dann gilt:

$$\bigwedge^m V = \{0\}.$$

Zu zeigen: $v_1 \wedge \ldots \wedge v_k = 0$. Dies folgt aus $v_1, \ldots, v_k$ linear unabhängig, daraus folgt $v_1 \wedge \ldots \wedge v_k = 0$.

Beweis: Ohne Einschränkung sei $v_1 = \alpha_2 v_2 + \cdots + \alpha_k v_k$, einsetzen ergibt:

$$v_1 \wedge \ldots \wedge v_k = \sum_{i=2}^{n} \alpha_i \underbrace{(v_i \wedge v_2 \wedge \ldots \wedge v_k)}_{=0}$$

(jeder Summand enthält zwei gleiche Vektoren). □

Sei nun Dim $V < \infty$.

Satz. *Die Vektoren $v_1, \ldots, v_k$ sind genau dann linear unabhängig, wenn*

$$v_1 \wedge \ldots \wedge v_k \neq 0$$

gilt.

Beweis: Ergänze zu einer Basis $B = (v_1, \ldots, v_n)$ und zeige, daß

$$v_1 \wedge \ldots \wedge v_n \neq 0$$

gilt. Betrachte dazu die n-lineare, alternierende Abbildung δ mit

$$\delta(w_1, \ldots, w_n) := \det({}_B[w_1], \ldots, {}_B[w_n]).$$

Es gilt $\delta(v_1, \ldots, v_n) = 1$ und speziell: $f(v_1 \wedge \ldots \wedge v_n) = 1$, insbesondere gilt $v_1 \wedge \ldots \wedge v_n \neq 0$ (folgt sofort aus der Linearität von f).
Die andere Richtung folgt aus dem obigen Beweis. □

Namen- und Sachverzeichnis

Abbildung 3
–, adjungierte 70–72
–, Bild 119
–, bilineare 119–128
–, duale 40–42
–, Einschränkung 3, 62
–, Fortsetzung 3
– in den Grundkörper 40
–, Komposition 3
–, längentreue 73
–, lineare 29–45
–, multilineare 128–129
–, Rang 33–34
–, Restriktion 3, 62
–, winkeltreue 73
Abbildungsmatrix 36, 38–39
Abel, Niels Henrik (1802–1829) 7
abelsch 7
abgeschlossene Menge 112
Abgeschlossenheit 18
–, algebraische 99
abzählbare Menge 3
Addition 17
adjungierte Abbildung 70–72
–, Beispiel 70
–, Determinante 71
–, Eindeutigkeit 71
–, Matrixdarstellung 71
–, Rechenregeln 71–72
–, Spur 71
Adjunkte 52
Ähnlichkeit
–, Matrix zur Transponierten 105
– von Matrizen 37, 57
Aleph 3
Algebra 130–133
–, äußere 132–133
–, Faktor- 131
–, Graßmann- 132–133
–Homomorphismus 131
–, Lie- 130
–, multilineare 119–133
algebraische Abgeschlossenheit 99
algebraische Basis 118
algebraische Vielfachheit 58, 62
allgemeine lineare Gruppe 13
allgemeine Lösung 12
alternierend 47
Antinomie von Russell 1–2
antisymmetrisch 4
Approximation 114
Approximationssatz 115
Äquivalenzklasse 5
Äquivalenzrelation 5–6, 37
Assoziativität 6
äußere Algebra 132–133
äußeres Produkt 132
Auswahlaxiom 6
Automorphismus 31, 73

Banach, Stefan (1892–1945) 112
Banachraum 112
Basis 21–24
–, algebraische 118
–, Charakterisierung 21, 109
–, duale 40
–ergänzung 22, 110
–, Existenz 22, 110
–, Hamel- 118
–, Invarianz der Länge 22
–, kanonische 21
– vom Kern 33
–, natürliche 21
– eines unendlichen Vektorraumes 109
– eines Vektorraums 21
Basisdarstellung 122
Basisergänzung 22, 23, 110
Basisvektoren 32
Basiswechsel 37–39, 85–86, 122
–, Satz 37
Bessel, Friedrich Wilhelm (1784–1846) 114
Bessel-Ungleichung 114
Betrag eines Vektors 67–68
Bewegung 74
bijektiv 3, 31, 34
Bild 29, 119
Bildbereich 3
bilinear 119
bilineare Abbildung 119–128
bilineare Tensorprodukte 126–128
Bilinearform
–, Charakterisierung 65
–, schiefsymmetrische 124
–, symmetrische 65, 124
Bilinearformen 121–126
Blockform 15
Brüche 1

Cantor, Gabriel (1845–1918) 1
Cauchy, Augustin Louis, Baron de (1789–1857) 67
Cauchy-Folge 112
Cauchy-Schwarz-Ungleichung 67
Cayley, Arthur (1821–1895) 62, 63
Charakterisierung
– einer Basis 21, 109
– einer Bilinearform 65
– positiv definiter Matrizen 93
charakteristisches Polynom 56–58, 62
–, Beispiel 57, 59
Cramer, Gabriel (1704–1752) 52
Cramer-Regel 52

Darstellung eines Vektors 20
Darstellungsmatrix 36, 37
–, Ähnlichkeit 38
–, Beispiel 38–39
Darstellungssatz 36–37, 123
Definitionsbereich 3
Determinante 25, 46–53
– einer adjungierten Abbildung 71
–, Anwendungen 52–53
–, Berechnung 46
– in Dreiecksform 47
– von Elementarmatrizen 48
– eines Endomorphismus 53
–, Entwicklung 46, 50–52
–, Linearität 47
–, Multiplikationssatz 48
–, Rechenregeln 47–50
–, rekursive Definition 46
–, Symmetrie 49
–, Vandermonde- 25
Determinantentest zur Invertierbarkeit 48
Diagonalisierung 61–62
–, nichtorthogonale 90–92
–, orthogonale 79, 82
Diagonalmatrix 9, 61
Differentialgleichung 107
Differenz von Mengen 2
Dimension 21–24
–, Beispiel 23
–, unendlicher Vektorraum 23
– eines Vektorraumes 23–24
Dimensionsformel 23
direkte Summe 24
disjunkt 2
Distributivität 7
Division von Polynomen 26, 97
Divisionsalgorithmus für Polynome 27
Drehachse 75
Drehgruppe 73
Drehmatrix 32
Drehung 32, 75
–, Drehachse 75
–, Drehwinkel 75
Drehwinkel 75
Dreiecksform 11
Dreiecksgestalt einer Matrix 10
duale Abbildung 40–42
duale Basis 40
Dualraum 40–42
Durchschnitt 2

echte Dreiecksgestalt 10
echte Teilmenge 1
Eigenraum 54, 58
eigentlich orthogonal 72, 73
Eigenvektor 54–64
–, Beispiel 55, 59
–, Berechnung 58
–, Eigenschaften 55–56
Eigenwert 54–64, 91
–, algebraische Vielfachheit 58
–, Beispiel 55, 59
–, Berechnung 56–59
–, Eigenschaften 55–56
–, geometrische Vielfachheit 58
– eines orthogonalen Endomorphismus 73

– eines schiefadjungierten Endomorphismus 77
– eines selbstadjungierten Endomorphismus 76
eineindeutig 3
Einheitsmatrix 9, 16
Eins 6
Eins-zu-Eins-Abbildung 3
Einschränkung einer Abbildung 3, 62
Einselement 6
Einsetzabbildung 100
Einsetzhomomorphismus 100
Einsetzung 97
Element 1
–, inverses 7
–, neutrales 6
elementare Spaltenumformung 14, 15, 35, 48, 90–91
elementare Zeilenumformung 14, 15, 35, 48, 90–91
elementare Zeilenumformungen 16
Elementarmatrix 14–16
–, Menge der Produkte 16
elementfremd 2
endlich erzeugt 19, 23
endlichdimensional 23
endliche Menge 3
endlicher Körper 24
Endomorphismus 31
–, Determinante 53
–, Fortsetzung 81–82
–, idempotenter 31
–, komplexifizierter 82
–, nilpotenter 31, 56, 101
–, normaler 77–78, 81
–, orthogonaler 72–75, 77, 81, 85
–, schiefadjungierter 76–77, 81, 84
–, selbstadjungierter 75–77, 81, 84
–, unitärer 72–75, 77, 81
–, Vektorraum 31
Entwicklung einer Determinanten 46, 50–52
– nach einer Spalte 50
– nach einer Zeile 50
Epimorphismus 31, 44
–, kanonischer 44
–, natürlicher 44
erweiterte Matrix 11
Erzeugendensystem 19, 24, 33, 109
–, minimales 21, 109
Euklid (um 300 v.Chr.) 66
euklidischer Vektorraum 65–94
Existenz einer Basis 22, 110
Exponentialfunktion
–, Matrix- 106

Faktoralgebra 131
Faktorraum 43–44, 131
Familie
–, kartesisches Produkt 2
Familie von Mengen 2
Familie von Untervektorräumen 18
fast alle 95
Fläche eines Parallelogramms 52
Folge 112
–, Cauchy- 112
Form 40
–, bilineare 121–126
–, hermitsche 65, 66
formale Polynome
–, Ring 96
–, Vektorraum 96
Formales Polynom 95–99
Fortsetzung
– einer Abbildung 3
– eines Endomorphismus 81–82
– des Skalarprodukts 80–81
– des Vektorraumes 80
Fourierkoeffizienten 114

Gauß, Carl Friedrich (1777–1855) 12
Gauß-Algorithmus 12
general linear group 13
Geometrie 74–75
geometrische Vielfachheit 58, 62
geordnete Menge 4
Gerstner-Bunse, Angelika 106
gleichmächtig 3
Grad eines Polynoms 26, 96
Gradsatz 26
Graßmann, Hermann Günter (1808–1877) 132
Graßmann-Algebra 132–133
größter gemeinsamer Teiler 97
Gruppe 6–7
–, Beispiel 7
–, Dreh- 73
–, Halbgruppe 6
–, orthogonale 73
–, spezielle orthogonale 73
–, symmetrische 51
–, unitäre 73

Halbgruppe 6, 7
–, multiplikative 7
Hamel, Georg Karl Wilhelm (1877–1954) 118
Hamelbasis 118
Hamilton, Sir William Rowan (1805–1865) 62, 63
Hauptachsen 88
Hauptachsentransformation 87–88
Hauptformel der Matrizenmultiplikation 10
Hauptideal 27, 98
Hauptsatz
– endlicher Vektorräume 22
– nilpotenter Endomorphismen 102
Hermite, Charles (1822–1901) 65
hermitsche Form 65, 66
hermitsche Matrix 75, 76
Hilbert, David (1862–1943) 112
Hilbertraum 112
Hilbertraumbasen 117–118
homogen 11, 13
Homomorphiesatz 44
Homomorphismus 31, 37
–, Algebra- 131
–, Vektorraum 31
Hurwitz, Adolf (1859–1919) 94
Hyperebene 74
–, Dimension 74
Hyperfläche zweiter Ordnung 88

Ideal 27, 98, 131
–, Beispiel 28
idempotenter Endomorphismus 31
Identität 31
indefinit 66
Index 8
–, Spalten- 8
–, Zeilen- 8
Indexmenge 2
Inertialsystem 86
inhomogen 11
injektiv 3, 30, 31, 34, 73
Inklusion 4
Innenproduktraum 66
innere Komposition 6
inneres Produkt 65–67
–, Beispiel 66
invarianter Unterraum 39, 63
invariantes Komplement 39
Invarianz der Basislänge 22
inverse Matrix
–, Berechnung 16, 52
inverses Element 7
invertierbar 14, 15
invertierbare Matrix 13–14, 16, 35–36
–, Eigenschaften 16
Involution 31
irrationale Zahlen 1
irreduzibel 98
irreduzible Polynome 99
Isometrie 74
isomorphe Vektorräume 42
Isomorphie 42–43
Isomorphismus 31, 42–44
–, Beispiel 42

Jacobi, Carl Gustav (1804–1851) 94
Jordan, Camille (1838–1922) 95
Jordan-Normalform 95–108
–, Komplexifizierung 104

kanonische Basis 21
kanonischer Epimorphismus 44
kanonisches Skalarprodukt 66
kartesisches Produkt von Mengen 2
Kästchensatz 50, 62
Kegelschnitte 53
Kern 29
–, Basis 33
–, Links- 122
–, Rechts- 122
Kette 4
Klasse 6
–, Vertreter 6
Koeffizienten 12
–, Eindeutigkeit von 24
Koeffizientenmatrix 11
Koeffizientenvergleich 25, 95
kollinear 53
kommutativ 7
kommutativer Ring 7
– mit Eins 7, 26

komplexe Zahlen 1
komplexifizierter Endomorphismus 82
Komplexifizierung 80–85
–, Jordan-Normalform 104
Komponenten 17
Komposition
– von Abbildungen 3
–, innere 6
Konjugation 65
konstantes Polynom 96
Konstruktionssatz 33
Konvergenz der Matrix-Reihe 106
konvexe Menge 115
Koordinatenabbildung 32, 36, 42, 85
Koordinatensystem 85
Koordinatentransformation 85–86
Koordinatenvektor 32
Körper 7, 8, 17
–, algebraisch abgeschlossener 99
–, endlicher 24
–, unendlicher 25
Kreisgleichung 53
Kronecker, Leopold (1823–1891) 41
Kronecker Symbol 41

längentreue Abbildung 73
Laplace, Pierre Simon, Marquis de (1749–1827) 50
leere Menge 1
Leibniz, Gottfried Wilhelm (1646–1716) 51
Lemma von Zorn 6, 110
Lie, Marius Sophus (1842–1899) 130
Lie-Algebra 130
linear 29
linear abhängig 19
linear unabhängig 19
lineare Abbildung 29–45
–, Beispiele 31–32
–, Bild 29
–, Darstellung durch eine Matrix 36–39
–, Eigenschaften 30–31
–, Kern 29
–, Rang 33–34
lineare Fortsetzung 32–33
lineare Hülle 19, 109
linearer Raum 17
linearer Teilraum 18
lineares Gleichungssystem 11–13
–, Dreiecksform 11
–, homogenes 13
–, Lösung 11
–, Sonderformen 11
–, Staffelform 12
–, Stufenform 12
Linearform 40, 65
–, Beispiel 40
Linearkombination 19, 109
Linksinverses 6
Linkskern 122
Lösung 11
–, allgemeine 12
–, spezielle 12
–, triviale 13
Lösungsraum 35

Mächtigkeit einer Menge 3
Matrix 8–16, 32
–, Addition 8
–, Ähnlichkeit 37, 57
–, Ähnlichkeit zur Transponierten 105
–, Darstellung adjungierter Abb. 71
–, Darstellung linearer Abb. 36–39
–, Diagonal- 9
–, Dreiecksgestalt 10
–, Einheits- 9
–, Elementar- 14–16
–, erweiterte 11
–, Gleichheit 8
–, hermitsche 75, 76
–, invertierbare 13–14, 16, 35–36
–, Invertierbarkeitstest 48
–, Jordan-Normalform 95–108
–, Koeffizienten- 11
–, Matrizenmultiplikation 10
–, Multiplikation 10
–, nilzyklische 102
–, normale 77
–, Null- 9
–, orthogonale 72
–, positive definite 92–94
–, positive semidefinite 93
–, Rang 35
–, Rechenregeln 11
–, schiefsymmetrische 76
–, Schur-Normalform 107
–, Signatur 91
–, Skalarmultiplikation 8
–, Spur 40
–, Subtraktion 8
–, symmetrische 75, 76, 85–94
–, Trägheit 91
–, transponierte 8
–, unitäre 72
Matrix-Exponentialfunktion 106
Matrix-Reihe 106
–, Konvergenz 106
Matrixdarstellung
– adjungierter Abbildungen 71
– linearer Abbildungen 36–39
Matrizenmultiplikation 10
–, Hauptformel 10
maximale linear unabhängige Familie 21, 109
maximales Element 4
Maximumsnorm 106
Menge 1–6
–, abgeschlossene 112
–, abzählbare 3
– aller Mengen 1
–, Differenz 2
–, disjunkte 2
–, Durchschnitt 2
–, Element 1
– der Elementarmatrizenprodukte 16
–, elementfremde 2
–, endliche 3
–, Familie 2
–, Gleichheit 1
–, kartesisches Produkt 2
–, konvexe 115
–, leere 1
–, Mächtigkeit 3
– der Nebenklassen 43
–, obere Schranke 5
–operationen 2
–, Ordnung 4
–, orthogonales Komplement 69–70, 115
– der Polynome 25
–, Potenz- 1
–, Quotienten- 6
– der Spaltenvektoren 10
–, Teil- 1
–, untere Schranke 5
–, Vereinigung 2
– der Zeilenvektoren 10
Menge aller Mengen 1
Mengenoperationen 2
minimales Element 4
minimales Erzeugendensystem 21, 109
Minimalpolynom 63, 103
Monomorphismus 31
multilineare Abbildungen 128–129
multilineare Algebra 119–133
multilineare Tensorprodukte 128–129
Multiplikation zweier Polynome 26
Multiplikationssatz für Determinanten 48
multiplikative Halbgruppe 7

natürliche Basis 21
natürliche Zahlen 1
natürlicher Epimorphismus 44
natürliches Skalarprodukt 66
Nebendiagonale 102
Nebenklasse 43
–, Menge der 43
negativ definit 66
negativ semidefinit 66
neutrales Element 6
–, Eindeutigkeit 6
nicht ausgeartet 122
nichtorthogonale Diagonalisierung 90–92
–, Beispiel 92
nilpotenter Endomorphismus 31, 56, 101
–, Hauptsatz 102
Nilpotenzindex 101
nilzyklische Matrix 102
Norm
–, Maximums- 106
–, Polarisierung 68
– eines Vektors 67–68
normale Matrix 77
normaler Endomorphismus 77–78, 81
Normalform 15
– einer Quadrik 88
–, Schur- 107
normierte Vektorräume 111–112
normiertes Polynom 25, 98
Null 6
Nullabbildung 31
Nullelement 6
Nullmatrix 9
Nullpolynom 24, 26
Nullspalte 13
Nullstelle 27, 97

Nullteilerfreiheit von Polynomen 96
Nullzeile 13

obere Dreiecksgestalt 10
obere Schranke 5
one-to-one Abbildung 3
Ordnung 4
–, Beispiel 4
–, vollständige 4
orthogonal 72
Orthogonalbasis 68
orthogonale Abbildung
–, Eigenwerte 73
orthogonale Diagonalisierung 79, 82
orthogonale Gruppe 73
orthogonale Matrix 72
orthogonale Summe 70
orthogonale Trigonalisierung 78–79
orthogonale Untervektorräume 69
orthogonale Vektoren 67
orthogonaler Endomorphismus 72–75, 77, 81, 85
–, Eigenwerte 73
orthogonales Komplement 69–70, 115
Orthogonalisierung 67–70
Orthogonalisierungssatz 68
Orthogonalisierungsverfahren 68–70
Orthogonalsystem 68
–, vollständiges 117–118
Orthonormalbasis 68, 69
Orthonormalisierung 67–70
Orthonormalsystem 68, 117
–, vollständiges 118
Ortsvektor 85

Parallelogramm 52, 132
Parallelogrammungleichung 114
Parseval des Chênes, Marc-Antoine (1755–1836) 118
Permutation 51
–, Signatur 51
Pivot 12
Polarisierung 68
Polynom 25, 56
–, Addition 25, 95
–, charakteristisches 56–58, 62
–division 26, 97
–, Divisionsalgorithmus 27
–, Einsetzung 97
–, formales 95–99
–, Grad 26, 96
–, größter gemeinsamer Teiler 97
–, irreduzibles 98
–, Koeffizientenvergleich 95
–, kommutativer Ring mit Eins 26
–, konstantes 96
–, Menge der 25
–, Minimal- 63
–, Multiplikation 26, 95
–, normiertes 25, 98
–, Nullstelle 27, 97
–, Nullteilerfreiheit 96
–ring 26
–, Ring der formalen Polynome 96
–, Skalarmultiplikation 25
–, Teilen mit Rest 26, 97
–, teilerfremdes 98
–, Vektorraum der formalen Polynome 96
–, Wurzel 27, 97
–, Zerlegung in irreduzible Faktoren 99
Polynomdivision 26, 97
Polynomfunktion 24–28, 95
Polynomring 26
positiv definit 66, 92–94
–, Charakterisierung von Matrizen 93
positiv semidefinit 66, 93
Positivitätstest 91
Potenzmenge 1, 4
Potenzreihe 106
Prähilbertraum 113
Primärkomponente 103
Primärzerlegung 99
–, Satz 100
Primzahl 7
Produkt von Elementarmatrizen 16
Projektion 31, 56
Projektionssatz 116
Punkt 85
Punktkoordinate 85, 86
Punktmenge 85
Pythagoras von Samos (ca. 580–500 v.Chr.) 67, 114

quadratische Formen 87
Quadrik 88–90
–, Beispiel 89–90
–, Normalform 88
–, Standardform 88
–, Transformation 88
Transformation einer Quadrik 88
Quotientenmenge 6
Quotientenraum 43–44

Rang einer Abbildung 33–34
Rangformeln 33–34
Rangsatz 33
–, Verallgemeinerung 34
rationale Zahlen 1
Rechenregeln
– für Adjungierte 71–72
– für Determinanten 47–50
– für Matrizen 11
Rechtsinverses 6
Rechtskern 122
reflexiv 4, 5
Regel von Sarrus 46
rekursive Definition der Determinante 46
Relation 4
Repräsentant 6
Restklassengruppe 7
Restriktion einer Abbildung 3, 62
Riesz, Frigyes (1880–1956) 117, 123
Ring
– der formalen Polynome 96
–, kommutativer 7
– mit Eins 11
Ringe 7
Ringisomorphismus 42
Russell, Bertrand (1872–1970) 1

Sarrus, Pierre (1798–1861) 46
Satz
–, Approximations- 115
– von Cayley und Hamilton 63
– über die Charakterisierung einer Basis 21, 109
–, Darstellungs- 36
–, Darstellungssatz von Riesz 123
– über die Dimensionsformel 23
– von der dualen Basis 40
–, Hauptsatz für endliche Vektorräume 22
–, Konstruktionssatz 33
– über nilpotente Endomorphismen 102
– von der Primärzerlegung 100
–, Projektionssatz 116
– des Pythagoras 114
– von Pythagoras 67
– von Riesz 117, 123
– von Schur 78
– über die Tridiagonalisierung 106
– über den Koeffizientenvergleich 25
– vom Basiswechsel 37
schiefadjungiert 76–77, 81, 84
schiefsymmetrisch 76
schiefsymmetrische Bilinearform 124
Schmidt, Erhard (1876–1959) 68
Schmidtsches Orthogonalisierungsverfahren 68–70
Schur, Issai (1875–1941) 59, 61, 78
Schur-Normalform 59, 107
Schwarz, Hermann Amandus (1843–1921) 67
selbstadjungiert 75–77, 81, 84
Signatur
– einer Matrix 91
– einer Permutation 51
Skalarmultiplikation 8, 17
Skalarprodukt 65–67
–, Beispiel 66
–, Fortsetzung 80–81
–, kanonisches 66
–, natürliches 66
–, Polarisierung 68
Spaltenindex 8
Spaltenumformung 14, 15, 35, 48, 90–91
Spaltenvektoren
–, Menge 10
Spat 52, 53
Spatprodukt 52
special linear group 16
Spektralsatz 79, 82
spezielle lineare Gruppe 16
spezielle Lösung 12
spezielle orthogonale Gruppe 73
Spiegelung 74
–, Darstellungsmatrix 74
–, Determinante 74
Spur
– einer adjungierten Abbildung 71
– einer Matrix 40
Staffelform 12, 13
Standardform einer Quadrik 88
Stufenform 12
Substitutionsformel 86
Summe

–, direkte 24
–, orthogonale 70
Summenraum 18
surjektiv 3, 31, 34
Sylvester, James Joseph (1814–1897) 91
Symmetrie des Skalarproduktes 65
Symmetrie von Determinanten 49
symmetrisch 4, 5
symmetrische Bilinearform 65, 124
symmetrische Gruppe 51
symmetrische Matrix 75, 76, 85–94

Taylor, Brook (1685–1731) 87
Taylor-Entwicklung 87
Teilen mit Rest von Polynomen 26, 97
teilerfremde Polynome 98
Teilmenge 1
–, echte 1
Tensor
–multiplikation 129
tensorielle Produkte lineare Abbildungen 129
Tensorprodukt 119–128, 131
–, Beispiel 127
–, bilineares 126–128
–, multilineare 128
–, multilineares 129
Tetraeder 52
Trägheit 91
Trägheitssatz von Sylvester 91–92
Transformationsformel 86
transitiv 4, 5
Translation 74
transponierte Matrix 8
–, Ähnlichkeit 105
Triangulierung 59–61
Tridiagonalisierung 106
trigonalisierbar 59
Trigonalisierung 59–61
–, orthogonale 78–79
triviale Lösung 13

Übergangsmatrix 37
unendlicher Körper 25
unendlicher Vektorraum 109–118
–, Basis 109
–, Linearkombination 109
Ungleichung von Cauchy-Schwarz 67
unitär 72
unitäre Gruppe 73
unitäre Matrix 72
unitärer Endomorphismus 72–75, 77, 81
unitärer Vektorraum 65–94
untere Dreiecksgestalt 10
untere Schranke 5
Unterraum 18–20, 29
–, invarianter 39, 63
Unterraum-Test 18
Untervektorraum 18, 19, 30
–, Durchschnitt 18
–, Familie 18
Ursprung 85

Vandermonde-Determinante 25
Vektor 17, 85
–, Betrag 67–68
–, Darstellung 20, 21
–, Norm 67–68
–, orthogonal 67
Vektorkoordinate 86
Vektorraum 8, 17–28, 85
–axiome 17
–, Banachraum 112
–, Basis 21
–, Beispiel 17
– der Endomorphismen 31
– der Homomorphismen 31
–, Dimension 23–24
–, euklidischer 65–94
– der formalen Polynome 96
–, Fortsetzung 80
–, Hauptsatz für endlich erzeugte 22
–, isomorpher 42
–, normierter 111–112
–, orthogonales Komplement 69–70
– der Polynome 24–28
–, unendlicher 23, 109–118
–, unitärer 65–94
–, Unter- 18
–, vollständiger 112
Vektorraumaxiome 17
Vereinigung 2
Verknüpfung 6
Vertreter 6
Vertretersystem 6
Vielfachheit
–, algebraische 58, 62
–, geometrische 58, 62

vollständige Ordnung 4
vollständiger Vektorraum 112
vollständiges Orthogonalsystem 117–118
vollständiges Orthonormalsystem 118
Volumen 52
– eines Spats 52, 53
– eines Tetraeders 52
Volumenverzerrungsfaktor 53

winkeltreue Abbildung 73
Wurzel 27, 97

Zahlen 1
Zeilenindex 8
Zeilenstufenform 16
Zeilenumformung 14, 15, 35, 48, 90–91
Zeilenvektoren
–, Menge 10
Zorn, Max (1906–1993) 6
Zornsches Lemma 6, 110