Lineare Algebra

Notizen zur Vorlesung, WS 08/09 und SS 09, TU Berlin

Michael Joswig

Technische Details:

(i) Organisatorisches Konzept
  - ▷ Vorlesung
  - ▷ Übung
  - ▷ Tutorien

(ii) Hausaufgaben
  - ▷ Zweiergruppen
  - ▷ 50% der erreichbaren Punkte notwendig zur Klausurzulassung

(iii) Klausuren
  - ▷ Scheinklausuren am Ende des Semesters
  - ▷ Modulklausur im Sommer

(iv) Sprechstunden

(v) Literatur
  - ▷ kein Skript, aber Notizen
  - ▷ Greub: Linear Algebra: v. 23 (Graduate Texts in Mathematics): v. 23, 4. Auflage, Springer 1995
  - ▷ Jähnich: Lineare Algebra: Mit 110 Testfragen (Taschenbuch), 11. Auflage, Springer 2008
  - ▷ Scherfner und Volland: Lineare Algebra für das erste Semester (Broschiert), Pearson 2006

Diese Notizen sind aus verschiedenen meiner Vorlesungen an der TU Darmstadt und der TU Berlin erwachsen. Inspririert haben mich diverse Bücher und ein Vorlesungsskript von Karl-Herrmann Neeb.

Inhaltsverzeichnis

KAPITEL 0

Motivation

0.1. Zahlbereiche

$\mathbb{N} := \{0, 1, 2, \dots\}$ = Menge der *natürlichen Zahlen* mit Addition "+", Multiplikation "+" und Anordnung "<".

$\mathbb{Z} := \mathbb{N} \cup (-\mathbb{N})$ = Menge der *ganzen Zahlen*. Gleichungen der Art $a + x = b$ können in $\mathbb{Z}$ für beliebige $a, b \in \mathbb{N}$ (oder auch $a, b \in \mathbb{Z}$) gelöst werden.

$\mathbb{Q} := \{\frac{p}{q} : p, q \in \mathbb{Z},\ q \neq 0\}$ = Menge der *rationalen Zahlen* $\rightarrow$. Gleichungen der Art $ax+b = c$ können in $\mathbb{Q}$ für beliebige $a, b, c \in \mathbb{Z}$ mit $a \neq 0$ gelöst werden (bzw. $a, b, c \in \mathbb{Q}$).

$\mathbb{R}$ = Menge der *reellen Zahlen*, entsteht aus $\mathbb{Q}$ durch Vervollständigung ($\uparrow$ Analysis).

$\mathbb{C}$ = Menge der *komplexen Zahlen* ($\uparrow$ Analysis). Gleichungen der Art $ax^2 + bx + c = d$ können gelöst werden für beliebige $a, b, c, d \in \mathbb{R}$ (bzw. $\mathbb{C}$) mit $(a, b) \neq (0, 0)$.

Es gilt:

$$\mathbb{N} \subsetneq \underbrace{\mathbb{Z} \subsetneq \overbrace{\mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}}^{\text{Körper}}}_{\text{Ringe}}$$

0.2. Vektorrechnung in $\mathbb{R}^d$ (bzw. $\mathbb{C}^d$)

motiviert z.B. durch die Physik

—BILD—

Addition "+" Subtraktion von *Vektoren*, Skalarmultiplikation.

0.3. Vektorräume

- ▷ mathematische Konzeption zur Grundlegung der "Vektorrechnung"
- ▷ zwei Abstraktionsrichtungen
  - ◇ beliebig viele Dimensioen, nicht nur $1, 2, 3$; sogar ∞-dimensional
  - ◇ *beliebige Körper* als Koordinatenbereiche (wichtig z.B. für Kryptographie).

0.4. Skalarprodukte und Determinanten

- ▷ Zusätzlich zur Vektorrechnung *Winkel* zwischen Vektoren und *Vektorlängen*
- ▷ dadurch: Grundlegung der Elementargeometrie.

0.5. Lineare Gleichungssysteme

—BILD—

Gesucht: Menge aller Lösungen (x, y) für das lineare Gleichungssystem:

$$\begin{cases} 2y = x + 2 \\ 2y = 4 - x \end{cases}$$

Eliminiere y:

$$4 - x = x + 2 \Leftrightarrow 2 = 2x \Leftrightarrow 1 = x.$$

Setze x ein:

$$2y = 1 + 2 = 3 \Leftrightarrow y = \frac{3}{2}.$$

0.6. Symmetriebegriffe

—BILD—

KAPITEL 1

Grundlagen

1.1. Aussagenlogik und Quantoren

1.1.1. Aussagen. Die (Aussagen-) Logik handelt von mathematischen *Aussagen*, die nach gewissen Regeln aus einer gewissen Menge von Symbolen gebildet werden. Eine wohlgeformte Aussage hat entweder den *Wahrheitswert* "wahr" oder "falsch".

BEISPIEL 1.1.1. Aussagen:

- ▷ "$0 \in \mathbb{Z}$", bzw. "0 ist eine ganze Zahl"
- ▷ "$2 + 2 = 5$"
- ▷ "$a + a = 2 \cdot a$ gilt für alle natürlichen Zahlen a"

BEMERKUNG 1.1.2. Ob eine Aussage wahr oder falsch ist, hängt vom axiomatischen Kontext ab.

Aus bereits vorhandenen Aussagen A und B lassen sich wie folgt neue Aussagen bilden:

- ▷ *Negation*: $\neg A$ ist wahr genau dann, wenn A falsch ist; lies als: "nicht A"
- ▷ *Konjunktion*: $A \& B$ ist wahr genau dann, wenn A und B beide wahr sind; lies als: "A und B". Manchmal schreibt man auch "$A \wedge B$".
- ▷ *Disjunktion*: $A \vee B$ ist wahr genau dann, wenn mindestens A oder B wahr sind; lies als: "A oder B"
- ▷ *Subjunktion*: $A \Rightarrow B$ ist definiert als $(\neg A) \vee B$; lies als: "aus A folgt B"
- ▷ *Äquivalenz*: $A \Leftrightarrow B$ ist wahr genau dann, wenn die Wahrheitswerte von A und B gleich sind (d.h. beide wahr oder beide falsch); lies als: "A ist äquivalent zu B"

1.1.2. Wahrheitstafeln. Diese Konstruktionen neuer Aussagen aus bestehenden lassen sich durch die folgende *Wahrheitstafel* zusammenfassen:

A	B	$\neg A$	$A\&B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Durch Inspektion aller möglichen Wahrheitswertbelegungen lassen sich zum Beispiel folgende *Sätze der Aussagenlogik* (*Tautologien*) beweisen:

(i) $A \vee (\neg A)$ *Tertium non datur*

(ii) $\neg(A \vee B) \Leftrightarrow (\neg A)\&(\neg B)$ und $\neg(A\&B) \Leftrightarrow (\neg A) \vee (\neg B)$ *De Morganschen Regeln*

(iii) $\left.\begin{array}{l}(A\&B) \Leftrightarrow (B\&A) \\ (A \vee B) \Leftrightarrow (B \vee A)\end{array}\right\}$ *Kommutativität*

(iv) $\left.\begin{array}{l}(A\&B)\&C \Leftrightarrow A\&(B\&C) \\ (A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)\end{array}\right\}$ *Assoziativität*

$$\text{(v)}\quad \left.\begin{array}{l} A\&(B \vee C) \Leftrightarrow (A\&B) \vee (A\&C) \\ A \vee (B\&C) \Leftrightarrow (A \vee B)\&(A \vee C) \end{array}\right\} \textit{Distributivgesetze}$$

Zusätzlich gelten die folgenden *Regeln des logischen Schließens*, die man ebenfalls durch Inspektion der Wahrheitstafel verifizieren kann:

Satz 1.1.3 (Direkter Schluss). $[A\&(A \Rightarrow B)] \Rightarrow B$

Satz 1.1.4 (Indirekter Schluss). $[(\neg B)\&(A \Rightarrow B)] \Rightarrow \neg A$

Beweis.

A	B	$\neg B$	$A \Rightarrow B$	$(\neg B)\&(A \Rightarrow B)$	$\neg A$	$[(\neg B)\&(A \Rightarrow B)] \Rightarrow \neg A$
w	w	f	w	f	f	w
w	f	w	f	f	f	w
f	w	f	w	f	w	w
f	f	w	w	w	w	w

□

Satz 1.1.5 (Kontraposition). $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$

Der folgende Satz dient dazu, eine andere Beweismethode als durch Wahrheitstafeln zu demonstrieren.

Beispiel 1.1.6. Es gilt $[A \Rightarrow (B \vee C)] \Leftrightarrow [(A\&\neg B) \Rightarrow C]$.

Beweis.

$$\begin{aligned} &[A \Rightarrow (B \vee C)] \\ &\Leftrightarrow [\neg A \vee (B \vee C)] \Leftrightarrow [(\neg A \vee B) \vee C] \\ &\Leftrightarrow [\neg(\neg A \vee B) \Rightarrow C] \Leftrightarrow [(A\&\neg B) \Rightarrow C] \end{aligned}$$

□

Beispiel 1.1.7. Die folgende Aussage gilt z.B. über $\mathbb{Q}$:

$$\underbrace{\lambda\mu = 0}_{A} \Rightarrow \underbrace{\lambda = 0}_{B} \quad \text{oder} \quad \underbrace{\mu = 0}_{C}$$

Dies ist gleichwertig zu:

$$\underbrace{\lambda\mu = 0}_{A} \quad \text{und} \quad \underbrace{\lambda \neq 0}_{\neg B} \Rightarrow \underbrace{\mu = 0}_{C}$$

1.1.3. Quantoren. Sei I eine Menge und $(A_i)_{i\in I}$ eine Familie von Aussagen.

Definition 1.1.8 (Allquantor). "$\forall i \in I : A_i$" ist eine Aussage, die wahr ist genau dann, wenn die Aussage A_i wahr ist für jedes $i \in I$; lies als: "für alle $i \in I$ …".

Definition 1.1.9 (Existenzquantor). "$\exists i \in I : A_i$" ist eine Aussage, die wahr ist genau dann, wenn es (mindestens) ein $i \in I$ gibt, so dass A_i wahr ist; lies als: "es existiert ein $i \in I$ …".

Definition 1.1.10 (Quantor zur eindeutigen Existenz). "$\exists! i \in I : A_i$" ist eine Aussage, die wahr ist genau dann, wenn es genau ein $i \in I$ gibt, so dass A_i wahr ist (d.h. für alle anderen $j \in I\setminus\{i\}$ ist A_j falsch).

Beispiel 1.1.11. $[\exists n \in \mathbb{N} : n^2 = 5]$ ist falsch.

Beispiel 1.1.12. $[\forall q \in \mathbb{Q} : \exists r \in \mathbb{Q} : 2r = q]$ ist wahr.

1.2. Mengenlehre

Hier wird nur ein naiver Zugang zur Mengenlehre angeboten.

Definition 1.2.1.

(i) Eine *Menge* ist eine Zusammenfassung bestimmter, wohlunterscheidbarer Objekte zu einem ganzen. Die Objekte heißen *Elemente* der Menge.
(ii) *Extensionalität*: $A = B :\Leftrightarrow (\forall x : x \in A \Leftrightarrow x \in B)$
(iii) *leere Menge*: $\emptyset := \{x\colon x \neq x\}$
(iv) $A \subseteq B :\Leftrightarrow (\forall x : x \in A \Rightarrow x \in B)$
(v) $A \subsetneq B :\Leftrightarrow (A \subseteq B)\&(A \neq B)$.
(vi) *Potenzmenge* $\mathcal{P}(A) := \{x\colon x \subseteq A\}$

Bemerkung 1.2.2. Ungehemmte naive Mengenbildung führt zu Widersprüchen; z.B. (Russell): Es sei $R := \{x\colon x \notin x\}$. Die naive Frage: 'Gilt "$R \in R$" oder "$R \notin R$"?' führt auf einen Widerspruch. Die *axiomatische Mengenlehre* löst dieses Problem, indem sie solche Fragen als *syntaktisch unzulässig* ausschließt.

Ähnlich wie man Aussagen miteinander verknüpfen kann, kann man auch Mengen miteinander verknüpfen. Es seien im Folgenden A und B Mengen.

Definition 1.2.3.

(i) $A \backslash B := \{x\colon x \in A \text{ und } x \notin B\}$ *Differenzmenge*
(ii) $A \cap B := \{x\colon x \in A \text{ und } x \in B\}$ *Schnittmenge*
(iii) $A \cup B := \{x\colon x \in A \text{ oder } x \in B\}$ *Vereinigungsmenge*

Satz 1.2.4. *Es gilt:*

(i) $A \cap B = B \cap A$
(ii) $A \cup B = B \cup A$
(iii) $A \cap (B \cap C) = (A \cap B) \cap C$
(iv) $A \cup (B \cup C) = (A \cup B) \cup C$
(v) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
(vi) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Vergleichen Sie mit den in Abschnitt 1.1.2 aufgeführten Sätzen der Aussagenlogik.

1.3. Relationen und Funktionen

Definition 1.3.1 (Geordnetes Paar). $(x, y) := \{\{x, y\}, \{x\}\}$

Satz 1.3.2. $(x, y) = (x', y') \Leftrightarrow x = x'$ *und* $y = y'$.

Beweis. Sei $(x, y) = (x', y')$. Dann folgt $(\{x, y\} = \{x', y'\}$ und $\{x\} = \{x'\})$ oder $(\{x, y\} = \{x'\}$ und $\{x\} = \{x', y'\})$. Wir unterscheiden zwei Fälle.
Fall 1: $x = y$. Dann ist $\{x\} = \{x', y'\}$ und damit $x = x' = y'$
Fall 2: $x \neq y$. Dann ist $\{x, y\} \neq \{x'\}$ und damit $\{x, y\} = \{x', y'\}$. Außerdem gilt $\{x\} = \{x'\}$, also $x = x'$ und dann auch $y = y'$.
Die umgekehrte Inklusion folgt direkt. □

Definition 1.3.3. Zu zwei Mengen X, Y heißt

$$x \times y \ := \ \{(x, y) : \ x \in X \text{ und } y \in Y\}$$

das *kartesische Produkt* von X und Y.

Definition 1.3.4. Eine *(binäre) Relation* auf zwei Mengen X und Y ist eine beliebige Teilmenge von $X \times Y$. Falls $R \subseteq X \times Y$, so setzen wir

$$xRy \ :\Leftrightarrow \ (x, y) \in R\,.$$

Definition 1.3.5. Sei R eine Relation auf M

(i) R heißt *reflexiv*, falls

$$\forall x \in M : xRx$$

(ii) R heißt *symmetrisch*, falls

$$\forall x, y \in M : xRy \Leftrightarrow yRx$$

(iii) R heißt *antisymmetrisch*, falls

$$\forall x, y \in M : xRy \& yRx \Rightarrow x = y$$

(iv) R heißt *transitiv*, falls

$$\forall x, y, z \in M : xRy \& yRz \Rightarrow xRz.$$

Definition 1.3.6.

- ▷ Eine *Äquivalenzrelation* ist reflexiv, symmetrisch und transitiv.
- ▷ Eine *Ordnungsrelation* ist reflexiv, antisymmetrisch und transitiv.

Definition 1.3.7. Sei M eine Menge und $P \subseteq \mathcal{P}(M)$. Dann heißt P *Partition* von M, falls die folgenden Bedingungen erfüllt sind:

- ▷ $\forall m \in P : m \neq \emptyset$
- ▷ $\forall m, m' \in P : m \neq m' \Rightarrow m \cap m' = \emptyset$
- ▷ $\bigcup\{m \colon m \in P\} = M$

Sei $\sim$ eine Äquivalenzrelation auf eine Menge M. Zu $x \in M$ setze

$$[x]_\sim \ := \ \{y \in M : \ x \sim y\} \ \textit{Äquivalenzklasse}$$

Satz 1.3.8. *Die Äquivalenzklassen $\{[x]_\sim \colon x \in M\}$ einer beliebigen Äquivalenzrelation $\sim$ partitionieren M. Umgekehrt definiert jede Partition eine Äquivalenzrelation.*

Beweis. Übung. □

Beispiel 1.3.9. Betrachte die Menge der ganzen Zahlen $\mathbb{Z}$, und wähle $m \in \mathbb{Z}$.

$$x \equiv_m y \ :\Leftrightarrow \ \exists k \in \mathbb{Z} : km = x - y$$

Die Relation $\equiv_m$ wird als "*kongruent mod m*" gelesen. Die Kongruenzrelation ist eine Äquivalenzrelation, denn:

- ▷ $\equiv_m$ ist reflexiv:
  Sei $x \in \mathbb{Z} \Rightarrow 0 \cdot m = x - x \Rightarrow x \equiv_m x$.
- ▷ $\equiv_m$ ist symmetrisch:
  Seien $x, y \in \mathbb{Z}$ mit $x \equiv_m y \Rightarrow ex \cdot k \in \mathbb{Z}$:

$$km = x - y \Rightarrow (-k)m = y - x \Rightarrow y \equiv_m x.$$

- ▷ $\equiv_m$ ist transitiv:
  Seien $x, y, z \in \mathbb{Z}$ mit $x \equiv_m y$ und $y \equiv_m z \Rightarrow ex \cdot k, l \in \mathbb{Z}$, so dass $km = x - y$ und $lm = y - z \Rightarrow (k + l)m = x - z \Rightarrow x \equiv_m z$.

▷ Die resultierenden Partitionen sehen so aus:

$$\mathbb{Z} \;=\; \{0, 2, -2, 4, -4, \dots\} \cup \{1, -1, 3, -3, \dots\} \text{ für } m = 2$$

und

$$\mathbb{Z} \;=\; \{0, 3, -3, 6, -6, \dots\} \cup \{1, -2, 4, -5, 7, -8, \dots\} \cup \{2, -1, 5, -4, 8, -7, \dots\} \text{ für } m = 3$$

▷ Äquivalenzklassen:

$$[x]_{\equiv_m} \;=\; \{x, x \pm m, x \pm 2m, \dots, \} \;=\; x + m\mathbb{Z}.$$

DEFINITION 1.3.10. Eine Relation $R \subseteq X \times Y$ heißt *Funktion* von X nach Y, falls gilt

$$\forall x \in X : \exists y \in Y : xRy$$

und

$$\forall x \in X : \forall y \in Y : \forall y' \in Y : xRy \& xRy' \Rightarrow y = y'\,.$$

Man spricht von *Rechtseindeutigkeit*. Notation: “$R : X \to Y$”.

Die vorherige Definition verlangt, dass eine Funktion jedem Element des *Definitionsbereichs* X ein Element des *Wertebereichs* Y zuordnet. Lässt man diese Bedingung weg, redet man von einer *partiellen Funktion*.

DEFINITION 1.3.11.

(i) Eine Funktion $f : X \to Y$ heißt *injektiv*, falls

$$\forall x, x' \in X : f(x) = f(x') \Rightarrow x = x'\,.$$

(ii) Eine Funktion $f : X \to Y$ heißt *surjektiv*, falls

$$\forall y \in Y : \exists\, x \in X : f(x) = y\,.$$

(iii) Eine Funktion $f : X \to Y$ heißt *bijektiv*, falls f injektiv und surjektiv ist.

(iv) Zu $X' \subseteq X$ ist

$$f(X) \;:=\; \{f(x) :\; x \in X\}$$

das *Bild* von X unter f.

(v) Zu $Y' \subseteq Y$ ist

$$f(Y')^{-1} \;:=\; \{x \in X :\; f(x) \in Y'\}$$

das *volle Urbild* von x unter f.

DEFINITION 1.3.12. Seien A, B, C Mengen und $f : B \to C$ und $g : A \to B$ Abbildungen. Die Abbildung

$$f \circ g : A \to C : a \mapsto f(g(a))$$

heißt *Verkettung* von f und g. [hier auch als “erst g, dann f”].

Seien A, B, C, D Mengen und $f : C \to D$, $g : B \to C$, $h : A \to B$ Abbildungen.

SATZ 1.3.13.

$$f \circ (g \circ h) \;=\; (f \circ g) \circ h\,.$$

BEWEIS. Sei $a \in A$. Dann gilt

$$\begin{aligned}[f \circ (g \circ h)](a) &= f((g \circ h)(a)) = f(g(h(a))) \\ &= (f \circ g)(h(a)) = [(f \circ g) \circ h](a)\end{aligned}$$

□

1.4. Gruppen, Ringe, Körper

Definition 1.4.1. Eine (binäre) *Verknüpfung* auf einer Menge M ist eine Abbildung

$$* : M \times M \to M : (m, m') \mapsto m * m$$

Bemerkung 1.4.2. Je nach Verknüpfung sind teils sehr unterschiedliche Notationen üblich.

Beispiel 1.4.3.
(i) $M = \mathbb{Z}$, $* = + \leftarrow$ Addition
(ii) $M = \mathbb{R}$, $* = \cdot \leftarrow$ Multiplikation
(iii) M beliebig, $\cup$ und $\cap$ Verknüpfungen auf $\mathcal{P}(M)$.

Definition 1.4.4. Seien A, B Mengen. Dann bezeichnet

$$A^B \ := \ \{f : \ f : B \to A\}$$

die Menge aller Abbildungen von B nach A.

Beispiel 1.4.5. Sei M eine Menge. Dann ist die Verkettung

$$\circ : M^M \times M^M \to M^M$$

eine Verknüpfung auf M^M.

Definition 1.4.6. Eine Verknüpfung $* : M \times M \to M$ heißt

▷ *kommutativ* $:\Leftrightarrow \forall x, y \in M : x * y = y * x$
▷ *assoziativ* $:\Leftrightarrow \forall x, y, z \in M : (x * y) * z = x * (y * z)$.

Definition 1.4.7. Ein Tupel $(M, *)$ aus einer nicht-leeren Menge $M \neq \emptyset$ und einer assoziativen Verknüpfung $*$ heißt *Halbgruppe*.

Beispiel 1.4.8. Dies sind Beispiele für Halbgruppen:

(i) $(\mathbb{Z}, +)$
(ii) $(\mathbb{Z}, \cdot)$
(iii) $(\mathbb{N}, +)$
(iv) $(\mathbb{N}\backslash\{0\}, +)$
(v) $(2\mathbb{N}, +)$
(vi) $(\mathbb{R}, +)$
(vii) $(\mathbb{R}\backslash\{0\}, \cdot)$

Definition 1.4.9. Eine Halbgruppe $(G, *)$ heißt *Gruppe*, falls gilt

(i) existiert $e \in G$, so dass für alle $g \in G$ gilt:

$$e * g \ = \ g$$

und
(ii) zu jedem $g \in G$ existiert ein $h \in G$, so dass

$$h * g \ = \ e\,.$$

Bemerkung 1.4.10. Man kann zeigen, dass in einer Gruppe genau ein Element e mit der Eigenschaft i existiert. Dies heißt das *Neutralelement* von G. Ebenso existiert zu jedem $g \in G$ genau ein $h \in G$ mit der Eigenschaft ii. Dieses Element heißt das zu g *Inverse* und wird oft mit g^{-1} bezeichnet.

Beispiel 1.4.11.

(i) $(\mathbb{Z}, +)$ ist eine Gruppe. Das Neutralelement ist 0, das zu $a \in \mathbb{Z}$ Inverse ist $-a$.

(ii) $(\mathbb{R}, +)$ ist eine Gruppe. Außerdem ist $(\mathbb{R}\backslash\{0\}, \cdot)$ auch eine Gruppe. In diesem zweiten Fall ist das Neutralelement 1, und das zu $x \in \mathbb{R} \setminus \{0\}$ Inverse ist $\frac{1}{x}$.
(iii) Sei E eine Menge mit genau einem Element $e \in E$. Dann ist $(E, *)$ eine Gruppe für $* : (e, e) \to e$. Wir nennen $(E, *)$ eine *triviale Gruppe*.
(iv) Sei $\mathbb{F}_2 := \{0, 1\}$. Durch die folgende Tafel sei eine Verknüpfung $+ : \mathbb{F}_2 \times \mathbb{F}_2 \to \mathbb{F}_2$ definiert:

+	0	1
0	0	1
1	1	0

Dann ist $(\mathbb{F}_2, +)$ eine Gruppe mit dem Neutralelement 0.
(v) Sei M eine nicht-leere Menge und sei $\mathrm{Sym}(M) := \{f : (f : M \to M \text{ bijektiv})\}$. Dann ist $(\mathrm{Sym}(M), \circ)$ eine Gruppe mit einem Neutralelement

$$\mathrm{id}_m : M \to M : m \mapsto m\,.$$

Die Elemente von $\mathrm{Sym}(M)$ heißen auch *Permutationen*. Die Gruppe selbst heißt *symmetrische Gruppe*.

Beweis. Wir betrachten hier nur das letzte Beispiel.

Die Verkettung $\circ$ ist eine Verknüpfung auf Sym (M), weil die Verkettung von bijektiven Abbildungen bijektiv ist. Aus Satz 1.3.13 folgt, dass $(\mathrm{Sym}(M), \circ)$ eine Halbgruppe ist.

Sei $f : M \to M$ eine beliebige bijektive Abbildung, und $m \in M$. Es gilt $(\mathrm{id}_M \circ f)(m) = f(\mathrm{id}_M(m)) = f(m)$, also $\mathrm{id}_M \circ f = f$. Da f bijektiv ist, existiert die Umkehrabbildung $f^{-1} : M \to M$ (und die ist ebenfalls bijektiv).

Es gilt $(f^{-1} \circ f)(m) = f(f^{-1}(m)) = m = \mathrm{id}_M(m)$, also $f^{-1} \circ f = \mathrm{id}_M$. □

Definition 1.4.12. Sei R eine Menge mit zwei Verknüpfungen $+ : R \times R \to R$ und $\cdot : R \times R \to R$. Das Tripel $(R, +, \cdot)$ heißt *Ring*, falls gilt:

(i) $(R, +)$ ist eine kommutative Gruppe.
(ii) $(R, \cdot)$ ist eine Halbgruppe.
(iii) Es ist $(a + b)c = ac + bc$ sowie $a(b + c) = ab + ac$ für alle $a, b, c \in R$.

Beispiel 1.4.13.

(i) $(\mathbb{Z}, +, \cdot)$ ist ein Ring.
(ii) $(\mathbb{R}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sind Ringe.

Definition 1.4.14. Ein Ring $(K, +, \cdot)$ mit additiven Neutralelementen $0 \in K$ heißt *Körper*, falls $(K\backslash\{0\}, \cdot)$ eine kommutative Gruppe ist. Das Neutralelement 0 heißt *Nullelement*, und das mult. Neutralelement heißt *Einselement*.

Beispiel 1.4.15.

(i) $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sind Körper
(ii) $(\mathbb{Z}, +, \cdot)$ ist *kein* Körper.
(iii) Wenn man auf der Menge $\mathbb{F}_2\backslash\{0\} = \{1\}$ aus Beispiel 1.4.11(iv) die triviale Multiplikation $\cdot$ betrachtet, so ist $(\mathbb{F}_2, +, \cdot)$ ein Körper mit genau zwei Elementen. Es ist definiert $0 \cdot 0 = 0$ und $0 \cdot 1 = 1 \cdot 0 = 0$.

Bemerkung 1.4.16. Man kann für die lineare Algebra, die Körper wesentlich als Koordinatenbereiche benötigt, weitgehend auf die Kommutativität der Multiplikation verzichten; dies führt zu linearen Algebra über *Schiefkörpern*.

1.4.1. Restklassenringe und endliche Körper von Primzahlordnung. Sei $m \in \mathbb{Z}$. In Beispiel 1.3.9 wurde gezeigt, dass

$$x \equiv_m y :\Leftrightarrow \; \exists k \in \mathbb{Z} : km = x - y$$

eine Äquivalenzrelation ist. Die Äquivalenzklassen $[x] \equiv_m = x + m\mathbb{Z}$ heißen *Kongruenzklassen modulo m*.
Auf der Menge $\mathbb{Z}/m\mathbb{Z} := \mathbb{Z}/\equiv_m$ definieren wir zwei Verknüpfungen:

$$(x + m\mathbb{Z}) + (y + m\mathbb{Z}) \;:=\; (x+y) + m\mathbb{Z} \quad \text{und} \quad (x + m\mathbb{Z}) \cdot (y + m\mathbb{Z}) \;:=\; (xy) + m\mathbb{Z}\,.$$

Es ist die Wohldefiniertheit zu zeigen:
Seien $x, x', y, y' \in \mathbb{Z}$ mit $x \equiv_m x'$ und $y \equiv_m y'$. Dann gibt es $h, l \in \mathbb{Z}$ mit $km = x - x'$ und $lm = y - y'$. Es folgt weiter

$$\begin{aligned}(x' + y') + m\mathbb{Z} &= (x' + m\mathbb{Z}) + (y' + m\mathbb{Z}) \\ &= (x - km + m\mathbb{Z}) + (y - lm + m\mathbb{Z}) \\ &= (x + m\mathbb{Z}) + (y + m\mathbb{Z}) = (x+y) + m\mathbb{Z}\end{aligned}$$

Analog folgt $(x'y') + m\mathbb{Z} = (xy) + m\mathbb{Z}$.

Satz 1.4.17. $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ *ist ein kommutativer Ring.*

Beweis.

- ▷ $0 + m\mathbb{Z}$ ist das additive Neutralelement.
- ▷ Addition und Multiplikation sind beide assoziativ und kommutativ.
- ▷ Es gibt ein multiplikatives Neutralelement $1 + m\mathbb{Z}$.
- ▷ Es gelten die Distributivgesetze.

□

Satz 1.4.18. $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ *ist ein Körper genau dann, wenn m eine Primzahl ist.*

Beweis. Übungsaufgabe. □

1.4.2. Bemerkungen.

- ▷ Statt $x + m\mathbb{Z}$ schreibt man oft auch $\bar{x}$ oder sogar nur x.
- ▷ Für m prim schreibt man auch $\mathbb{F}_m := \mathbb{Z}/m\mathbb{Z}$. Verifizieren Sie, dass für $m = 2$ diese Notation auch zu Beispiel 1.4.11(iv) passt.

Die “Gleichartigkeit” algebraischer Strukturen wird durch den *Isomorphiebegriff* präzisiert. In diesem Sinn gelten (bis auf Isomorphie) die folgenden Eindeutigkeitsaussagen:

- ▷ $\mathbb{Z}/0\mathbb{Z} \cong \mathbb{Z}$
- ▷ $\mathbb{Z}/1\mathbb{Z} \cong \{0\}$ trivialer Ring
- ▷ $\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z} \Leftrightarrow m \pm n$
- ▷ Für m prim ist $\mathbb{F} = \mathbb{Z}/m\mathbb{Z}$ der einzige endliche Körper mit m Elementen.

Bemerkung 1.4.19. Es gibt (eindeutig bestimmte) endliche Körper zu jeder Primzahlpotenzordnung.

KAPITEL 2

Vektorräume

2.1. Definitionen und erste Beispiele

DEFINITION 2.1.1. Sei K ein Körper. Ein K-*Vektorraum* (oder *Vektorraum* über K) ist ein Tupel $(V, +, \cdot)$ bestehend aus einer Menge V, einer binären Verknüpfung (*Addition*)

$$+ : V \times V \to V$$

sowie einer *Skalarmultiplikation*

$$\cdot : K \times V \to V,$$

so dass gilt:

(i) $(V, +)$ ist eine kommutative Gruppe.
(ii) $\forall v, w \in V : \forall \lambda, \mu \in K :$
  (a) $(\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v)$
  (b) $\lambda \cdot (v + w) = (\lambda \cdot v) + (\mu \cdot w)$
  (c) $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$
  (d) $1 \cdot v = v$

BEMERKUNG 2.1.2. Die analoge Definition, bei der der Körper K durch einen (kommutativen) Ring R ersetzt wird, führt zum Begriff des *R-Moduls*.

BEISPIEL 2.1.3. Für $n \in \mathbb{N}$ ist der *Standard-K-Vektorrraum*

$$K^n := \underbrace{K \times \cdots \times K}_{n-mal} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, x_2, \ldots, x_n \in K \right\}$$

mit der Addition

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

und der Skalarmultiplikation

$$\lambda \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} \lambda y_1 \\ \vdots \\ \lambda y_n \end{pmatrix}$$

ein Vektorraum. Um dies zu zeigen, verifizieren wir die Axiome 2.1.1(i) und 2.1.1(ii):

Der *Nullvektor* $0 := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ ist wegen

$$0 + \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + 0$$

ein *Neutralelement* für die Vektoraddition. Das zu $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ additive Inverse ist $-\begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ wegen

$$\begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix} + \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -x_1 + x_1 \\ \vdots \\ -x_n + x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 .$$

Da die Addition in K assoziativ und kommutativ ist, ist auch die Addition in K^n assoziativ und kommutativ.
Seien $\lambda, \mu \in K$ und $v, w \in K^n$ beliebig mit

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{und} \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} .$$

Dann gilt

$$\begin{aligned} (\lambda + \mu) \cdot v &= (\lambda + \mu) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} (\lambda + \mu)\, v_1 \\ \vdots \\ (\lambda + \mu) \cdot v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 + \mu v_1 \\ \vdots \\ \lambda v_n + \mu v_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix} + \begin{pmatrix} \mu v_1 \\ \vdots \\ \mu v_n \end{pmatrix} = \left[\lambda \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right] + \left[\mu \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right] \\ &= (\lambda \cdot v) + (\mu \cdot v) . \end{aligned}$$

Die übrigen Eigenschaften lassen sich analog beweisen.

Beispiel 2.1.4. Sei M eine beliebige Menge (und K ein beliebiger Körper). Es ist

$$K^M = \{ f : f : M \to K \}$$

die Menge aller Abbildungen von M nach K. Mit der *punktweise* definierten Addition:

$$f + g : M \to K : m \mapsto f(m) + g(m)$$

und der ebenfalls punktweise definierten Skalarmultiplikation

$$\lambda \cdot f : M \to K : m \mapsto \lambda f(m)$$

ist $(K^M, +, \cdot)$ ein Vektorraum.

2.2. Linearkombinationen

Es sei $(v_1, v_2, \ldots, v_k)$ ein geordnetes k-*Tupel* (bzw. *Familie* der Länge k) von Vektoren aus einem K-Vektorraum V.

Definition 2.2.1. Ein Vektor $v \in V$ heißt

- ▷ *Linearkombination* von $(v_1, \ldots, v_k)$, falls existieren $\lambda_1, \ldots, \lambda_k \in K$, so dass gilt $v = \lambda_1 v_1 + \cdots + \lambda_k v_k$.
- ▷ *Affinkombination* von $(v_1, \ldots, v_k)$, falls existieren $\lambda_1, \ldots, \lambda_k \in K$, so dass gilt $v = \lambda_1 v_1 + \cdots + \lambda_k v_k$ und zusätzlich $\lambda_1 + \cdots + \lambda_k = 1$.

DEFINITION 2.2.2. Speziell für $K = \mathbb{R}$ (oder allgemeiner einen angeordneten Körper K) heißt eine Affinkombination

$$\lambda_1 v_1 + \cdots + \lambda_k v_k$$

Konvexkombination, falls zusätzlich $0 \leq \lambda_i \leq 1$ für alle $1 \leq i \leq k$.

DEFINITION 2.2.3. Das k-Tupel $(v_1, \ldots, v_k)$ heißt *linear unabhängig*, falls gilt:

$$\forall (\lambda_1, \ldots, \lambda_k) \in K^k : \lambda_1 v_1 + \cdots + \lambda_k v_k = 0 \Rightarrow \lambda_1 = \cdots = \lambda_k = 0.$$

Andernfalls *linear unabhängig*.

LEMMA 2.2.4. *Sei $(v_1, \ldots, v_k)$ linear unabhängig.*

(i) Für jede Permutation $\pi \in \mathrm{Sym}(\{1, \ldots, k\})$ ist auch $(v_{\pi(1)}, \ldots, v_{\pi(k)})$ linear unabhängig.
(ii) Jede Teilfamilie $(v_1, \ldots, v_i)$ für $i \leq k$ ist linear unabhängig.

BEWEIS. Seien $\lambda_1, \lambda_2, \ldots, \lambda_k \in K$ gegeben mit

$$0 \;=\; \lambda_1 v_{\pi(1)} + \cdots + \lambda_k v_{\pi(k)} \;=\; \lambda_{\pi^{-1}(1)} v_1 + \cdots + \lambda_{\pi^{-1}(k)} v_k \,.$$

Dann folgt $\lambda_{\pi^{-1}(1)} = \cdots = \lambda_{\pi^{-1}(k)} = 0$. Angenommen $(v_1, v_2, \ldots, v_i)$ wäre linear abhängig für ein $i \leq k$. Dann existieren $\lambda_1, \lambda_2 \ldots, \lambda_i \in K$ und $\lambda_h \neq 0$ für $1 \leq h \leq i$, so dass $\lambda_1 v_1 + \cdots + \lambda_i v_i = 0$ ist. Hieraus folgt nun

$$\lambda_1 v_1 + \cdots + \lambda_i v_i + 0 \cdot v_{i+1} + \cdots + 0 \cdot v_k \;=\; 0$$

mit $\lambda_h \neq 0$. Dies steht im Widerspruch zur Annahme, die Familie $(v_1, \ldots, v_h)$ sei linear unabhängig. □

BEISPIEL 2.2.5. Im Standard-K-Vektorraum K^n sind die *Einheitsvektoren*

$$e_1 \;:=\; \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_k \;:=\; \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n \;:=\; \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

linear unabhängig.

BEISPIEL 2.2.6. Sei V beliebiger K-Vektorraum und $V \ni x \neq 0$. Dann ist (x) linear unabhängig und $\underbrace{(x, x, \ldots, x)}_{k\text{-fach},\ k \geq 2}$ linear abhängig.

DEFINITION 2.2.7. Eine unendliche Familie von Vektoren aus K heißt *linear unabhängig* falls jede endliche Teilfamilie linear unabhängig ist.

BEMERKUNG 2.2.8.

- ▷ Wegen Lemma 2.2.4 lässt sich der Begriff der linearen (Un-)Abhängigkeit auch auf Mengen von Vektoren anwenden.
- ▷ Der Nullvektor $0 \in V$ ist linear abhängig.

PROPOSITION 2.2.9. *Für $n \geq 2$ sind $v_1, v_2, \ldots, v_n$ genau dann linear abhängig, falls einer dieser Vektoren eine Linearkombination der übrigen ist.*

BEWEIS. Gilt $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ mit $\lambda_i \neq 0$, so folgt $v_i = \sum_{j \neq i} -\frac{\lambda_j}{\lambda_i} v_j$. Ist umgekehrt zum Beispiel $v_1 = \mu_1 v_2 + \cdots + \mu_n v_n$, so folgt $(-1)v_1 + \mu_1 v_2 + \cdots + \mu_n v_n = 0$. □

2.3. Unterräume

Sei V ein K-Vektorraum.

DEFINITION 2.3.1. Eine nichtleere Teilmenge $U \subseteq V$ heißt *Teilraum* (oder *Unterraum*) von V, falls gilt $\forall \lambda, \mu \in K \; \forall u, v \in U$:

$$\lambda u + \mu v \;\in\; U\,.$$

BEMERKUNG 2.3.2.

- ▷ Wegen $U \neq \emptyset$ existiert $u \in U \Rightarrow 0 \cdot u = 0 \in U$.
- ▷ Falls U Unterraum von V, schreiben wir “$U \leq V$”.

BEISPIEL 2.3.3.

(i) Es ist $\{0\} \leq V$ und $V \leq V$.

(ii) Die Lösungen $x = (x_1, \dots, x_n) \in K^n$ des *homogenen linearen Gleichungssystems*

$$\begin{array}{ccccccc} \alpha_{11}\, x_1 & + & \cdots & + & \alpha_{1n}\, x_n & = & 0 \\ \vdots & & \ddots & & \vdots & & \vdots \\ \alpha_{n1}\, x_1 & + & \cdots & + & \alpha_{mn}\, x_n & = & 0 \end{array}$$

bilden einen Teilraum von K^n.

(iii) Betrachte den $\mathbb{R}$-Vektorraum $\mathbb{R}^{\mathbb{R}}$ aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ mit punktweise definierter Addition, vergleiche Beispiel 2.1.4. Dann bilden
- ▷ die stetigen Abbildungen $\mathcal{C}(\mathbb{R})$ einen Teilraum,
- ▷ die differenzierbaren Abbildungen einen Teilraum usw.

(iv) Betrachte die einmal stetig differenzierbaren Lösungen $u \in \mathcal{C}^1([0,1])$ der Differenzialgleichung

$$\dot{u}(t) \;=\; a(t) \cdot u(t)$$

für $a \in \mathcal{C}([0,1])$. Dann gilt für alle $\lambda, \mu \in \mathbb{R}$ und für alle Lösungen u, v, dass wegen

$$a(t)(\lambda u(t) + \mu v(t)) \;=\; \lambda a(t)u(t) + \mu a(t)v(t) \;=\; \lambda \dot{u}(t) + \mu \dot{v}(t)$$

die Funktion $\lambda\mu + \mu v \in \mathcal{C}^1([0,1])$ wieder eine Lösung ist. Das heißt die Lösungsmenge bildet einen Teilraum von $\mathcal{C}^1([0,1])$.

PROPOSITION 2.3.4. *Seien U, W Teilräume des K-Vektorraums V. Dann sind $U \cap W$ und die* Unterraumsumme

$$U + W \;:=\; \{u + w : \; u \in U, w \in W\}$$

Unterräume von V.

BEWEIS. Wir betrachten zunächst den Schnitt $U \cap W \;\leq\; V$. Seien $\lambda, \mu \in K$ und $x, y \in U \cap W$. Wegen $U \leq V$ ist $\lambda x + \mu y \in U$. Und wegen $W \leq V$ gilt $\lambda x + \mu y \in W$. Insgesamt ist also $\lambda x + \mu y \in U \cap W$.

Nun betrachten wir die Unterraumsumme $U{+}W \subseteq V$. Seien $\lambda\mu \in K$ und $u_1{+}w_1, u_2{+}w_2 \in U + W$. Dann ist

$$\lambda(\mu_1 + w_1) + \mu(u_2 + w_2) \;=\; \underbrace{\lambda u_1 + \mu u_2}_{\in U} + \underbrace{\lambda w_1 + \mu w_2}_{\in W} \;\in\; U + W\,.$$

□

Definition 2.3.5. Sei V ein K-Vektorraum und $M \subseteq V$. Die Menge

$$\mathrm{lin}_K(M) \ := \ \{\lambda_1 m_1 + \cdots + \lambda_m m_n : \ \lambda_i \in K, m_i \in M\}$$

heißt *lineare Hülle* (oder *linearer Aufspann*) von M in V. Falls $M = \emptyset$ setzen wir $\mathrm{lin}_K(M) := \{0\}$.

Proposition 2.3.6. *Der Aufspann* $\mathrm{lin}(M)$ *ist der kleinste Unterraum von* V*, der* M *enthält.*

Beweis. Zunächst zeigen wir, dass $\mathrm{lin}(M)$ ein Unterraum ist. Ohne Einschränkung können wir $M \neq \emptyset$ annehmen. Seien $\lambda, \mu \in K$ sowie $\lambda_1 m_1 + \cdots + \lambda_n m_n, \lambda'_1 m'_1 + \cdots + \lambda_{n'} m'_{n'} \in \mathrm{lin}(M)$. Dann gilt

$$\begin{aligned}\lambda(\lambda_1 m_1 + \cdots + \lambda_n m_n) + \mu(\lambda'_1 m'_1 + \cdots + \lambda_{n'} m'_{n'}) \\ = \ \lambda\lambda_1 m_1 + \cdots + \lambda\lambda_n m_n + \mu\lambda'_1 m'_1 + \cdots + \mu\lambda'_n m'_{n'} \ \in \ \mathrm{lin}(M) \, .\end{aligned}$$

Wir müssen nun noch die Minimalität beweisen. Hierzu betrachten wir einen Unterraum $U \leq V$ mit $U \supseteq M$. Es ist zu zeigen, dass gilt $U \supseteq \mathrm{lin}(M)$. Wähle $\lambda_1, \ldots, \lambda_n \in K$ und $m_1, \ldots, m_n \in M$. Wir zeigen $\lambda_1 m_1 + \cdots + \lambda_n m_n \in U$ durch Induktion nach n.

Anfang: Für $n = 1$ ist $\lambda_1 m_1 \in U$, weil $m_1 \in U$ und U ein Unterraum ist.

Voraussetzung: $\lambda_1 m_1 + \cdots + \lambda_{n-1} m_{n-1} \in U$.

Schluss: $\lambda_1 m_1 + \cdots + \lambda_n m_n = 1(\lambda_1 m_1 + \cdots + \lambda_{n-1} m_{n-1}) + \lambda_n m_n \in U$. □

2.4. Unterräume von $\mathbb{R}^2$

Wie sehen die Unterräume des reellen Standardvektorraums $\mathbb{R}^2$ aus? Offenbar gilt $\{0\} \leq \mathbb{R}^2$ und $\mathbb{R}^2 \leq \mathbb{R}^2$.

Sei $U \leq \mathbb{R}^2$ mit $U \neq 0$ und $u \in U \setminus \{0\}$. Dann gilt

$$\mathrm{lin}(u) \ = \ \{\lambda u : \ \lambda \in \mathbb{R}\} \ \subseteq \ U \, .$$

—BILD—

Falls existiert $v \in U \setminus \mathrm{lin}(u)$, dann gilt $U = \mathbb{R}^2$: Seien

$$u \ = \ \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{und} \quad v \ = \ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} .$$

Offenbar ist $u_1 v_2 - v_1 u_2 \neq 0$ wegen $v \notin \mathrm{lin}(u)$. Sonst gilt:

$$u_1 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} - v_1 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \ = \ \begin{pmatrix} u_1 v_1 & - & u_1 v_1 \\ u_1 v_2 & - & v_1 u_2 \end{pmatrix} \ = \ 0 \, ,$$

aber u und v linear unabhängig.

Sei $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ beliebig. Dann ergibt sich

$$\begin{aligned}(v_2 x_1 - v_1 x_2) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + (u_1 x_2 - u_2 x_1) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ = \ \begin{pmatrix} u_1 v_2 x_1 & - & u_1 v_1 x_2 \\ u_2 v_2 x_1 & - & u_2 v_2 x_2 \end{pmatrix} + \begin{pmatrix} u_1 v_1 x_2 & - & u_2 v_1 x_1 \\ u_1 v_2 x_2 & - & u_2 v_2 x_1 \end{pmatrix} \\ = \ (u_1 v_2 - u_2 v_1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\end{aligned}$$

Weil nun $u_1v_2 - u_2v_1 \neq 0$ ist, folgt weiter

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \operatorname{lin}(u,v)$$

Insgesamt also $\operatorname{lin}(u,v) = \mathbb{R}^2$.

Proposition 2.4.1. *Die nicht-trivialen Unterräume von $\mathbb{R}^2$ entsprechen genau den Geraden durch den Ursprung.*

2.5. Wie testet man lineare (Un-)abhängigkeit?

Beispiel 2.5.1. Drei Vektoren $(u,v,w) \in \mathbb{R}^4$ sind genau dann linear abhängig, wenn $\exists(\lambda,\mu,\nu) \in \mathbb{R}^3 \setminus \{0\} : \lambda\mu + \mu v + \nu w = 0$.
Betrachte

$$u = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{und } w = \begin{pmatrix} 3 \\ -1 \\ -4 \\ -2 \end{pmatrix}.$$

Diese Vektoren sind linear abhängig, wenn existieren $\lambda,\mu,\nu \in \mathbb{R}$, nicht all gleich 0, so dass

$$\left.\begin{array}{rrrcl} \lambda & & +3\nu & = & 0 \\ & \mu & +\,\nu & = & 0 \\ -\lambda & \mu & -4\nu & = & 0 \\ & 2\mu & -2\nu & = & 0 \end{array}\right\} . \tag{1}$$

Das heißt, die Vektoren sind linear abhängig genau dann, wenn das homogene LGS (1) nicht-triviale Lösung hat. Wegen

$$-3u + v + w = 0$$

existiert eine nicht-triviale Lösung, und u,v,w sind linear abhängig.

2.6. Basen und Erzeugendensysteme

Sei V ein K-Vektorraum.

Definition 2.6.1. (i) Eine Menge $M \subseteq V$ heißt *Erzeugendensystem* von V, falls $\operatorname{lin}(M) = V$.
(ii) Eine Familie in V heißt *Basis*, falls sie ein linear unabhängiges Erzeugendensystem bildet.

Beispiel 2.6.2. Im Standard-K-Vektorraum K^n ist die Familie der Einheitsvektoren

$$(e_1, e_2, \ldots, e_n)$$

eine Basis, die *Standardbasis* von K^n. In Beispiel 2.2.5 haben wir gesehen, dass die Einheitsvektoren linear unabhängig sind. Zu zeigen bleibt, dass sie den gesamten Raum K^n aufspannen:

Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. Dann ist

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n \in \operatorname{lin}(e_1, \ldots, e_n).$$

Beispiel 2.6.3. Betrachte den Unterraum $U = \operatorname{lin}(u, v, w) \leq \mathbb{R}^4$, wobei u, v, w wie in Beispiel 2.5.1 gewählt sind.

Wir zeigen, dass (u, v) eine Basis von U ist. Für $\lambda, \mu \in \mathbb{R}$ folgt aus

$$\begin{pmatrix} \lambda \\ 0 \\ -\lambda \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \mu \\ \mu \\ 2\mu \end{pmatrix} = \begin{pmatrix} \lambda \\ \mu \\ \mu - \lambda \\ 2\mu \end{pmatrix} = 0 ,$$

dass $\lambda = \mu = 0$ ist. Also ist das Paar (u, v) linear unabhängig. Außerdem ist $w = 3u - v$, das heißt $w \in \operatorname{lin}(u, v)$. Daher gilt $U = \operatorname{lin}(u, v)$, und (u, v) ist eine Basis von U.

Beispiel 2.6.4. Das Paar $(1, i)$ ist eine Basis des reellen Vektorraums $\mathbb{C}$.

Definition 2.6.5. Ein Vektorraum V heißt *endlich erzeugt*, falls er ein endliches Erzeugendensystem besitzt.

Satz 2.6.6. *Sei $V \neq \{0\}$ ein K-Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren aus V. Dann sind die folgenden Aussagen äquivalent.*

(i) *$(v_i)_{i\in I}$ ist eine Basis von V.*
(ii) *$(v_i)_{i\in I}$ ist ein* unverkürzbares Erzeugendensystem *von V, d.h. $\forall J \subsetneq I$ ist $(v_j)_{j\in J}$ kein Erzeugendensystem von V.*
(iii) *$(v_i)_{i\in I}$ ist eine* unverlängerbare linear unabhängige Familie*, d.h. $\forall J' \supsetneq I$ ist $(v_j)_{j\in J'}$ linear abhängig.*
(iv) *$(v_i)_{i\in I}$ ist ein Erzeugendensystem von V, aus dem sich jeder Vektor von V eindeutig linear kombinieren läßt.*

Beweis. $\boxed{\text{(i)} \Rightarrow \text{(ii)}}$ Sei $(v_i)_{i\in I}$ Basis von V. Angenommen, $(v_i)_{i\in I}$ ist kein unverkürzbares Erzeugendensystem von V. Dann existiert $J \subsetneq I$, so dass $\operatorname{lin}\{v_j : j \in J\} = V$. Sei $i_0 \in I \setminus J$. Dann existieren $j_1, \ldots, j_k \in J$ und $\lambda_1, \ldots, \lambda_k \in K$, so dass

$$v_{i_0} = \lambda_1 v_{j_1} + \cdots + \lambda_k v_{j_k} .$$

Aus Lemma 2.2.9 folgt dann, dass $(v_{i_0}, v_{j_1}, \ldots, v_{j_k})$ linear abhängig ist. Dies steht wegen Lemma 2.2.4 im Widerspruch dazu, dass $(v_i)_{i\in I}$ Basis ist.

$\boxed{\text{(ii)} \Rightarrow \text{(iii)}}$ Sei $(v_i)_{i\in I}$ unverkürzbares Erzeugendensystem von V.

<u>Behauptung:</u> $(v_i)_{i\in I}$ ist linear unabhängig.

Wegen $V \neq \{0\}$ ist $I \neq \emptyset$. Falls $I = \{i_0\}$ einelementig, so ist wegen $V \neq \{0\}$ auch $v_{i_0} \neq 0$, also (v_{i_0}) linear unabhängig. Wir nehmen nun also an $\#I \geq 2$. Wäre $(v_i)_{i\in I}$ linear abhängig, würden wegen Proposition 2.2.9 Indizes $i_0, i_1, \ldots, i_k \in I$ existieren mit $v_{i_0} = \lambda_1 v_{i_1} + \cdots + \lambda_k v_{i_k}$ und $\lambda_1, \ldots, \lambda_k \in K$. Dies steht im Widerspruch zur Unverkürzbarkeit.

<u>Behauptung:</u> Es existiert kein $J \supsetneq I$, so dass $(v_j)_{j\in J}$ linear unabhängig.

Angenommen, so ein $J \subsetneq I$ existiert doch. Dann existiert $j_0 \in J \setminus I$. Da $(v_i)_{i\in I}$ Erzeugendensystem ist, existieren $\lambda_1, \ldots, \lambda_k \in K$ und $i_1, \ldots, i_k \in I$, so dass $v_{j_0} = \lambda_1 v_{i_1} + \cdots + \lambda v_{i_k}$. Damit ist $(v_j)_{j\in I}$ linear unabhängig wegen Proposition 2.2.9.

$\boxed{\text{(iii)} \Rightarrow \text{(iv)}}$ Sei $(v_i)_{i\in I}$ unverlängerbare linear unabhängige Familie von Vektoren in V.

<u>Behauptung:</u> $(v_i)_{i\in I}$ erzeugt V.

Sei $v \in V$. Dann ist die um v verlängerte Familie linear abhängig. Daher existieren $\lambda, \lambda_1, \ldots, \lambda_k \in K$ und $i_1, \ldots, i_k \in I$, so dass

$$\lambda v + \lambda_1 v_1 + \cdots + \lambda_k v_k = 0$$

und $(\lambda, \lambda_1, \dots, \lambda_k) \neq (0, \dots, 0)$.
Angenommen $\lambda = 0$. Dann folgt $(\lambda_1, \dots, \lambda_k) \neq (0, \dots, 0)$ und deswegen ist $(v_1, \dots, v_k)$ linear abhängig. Hieraus schließen wir $\lambda \neq 0$ und $v = -\frac{\lambda_1}{\lambda} v_1 - \dots - \frac{\lambda_k}{\lambda} v_k$. Das heißt $\mathrm{lin}\{v_i : i \in I\} = V$.
Behauptung: $\forall v \in V \exists! \lambda_1, \dots, \lambda_k \in K \setminus \{0\}$, $i_1, \dots, i_k \in I$:

$$v = \lambda_1 v_1 + \dots + \lambda_k v_k .$$

Wir nehmen also an, dass existieren $\lambda_1, \dots, \lambda_k \in K$, $i_1, \dots, i_k \in I$ und $\lambda'_1, \dots, \lambda'_{k'}$, $i'_1, \dots, i'_{k'}$ mit

$$v = \lambda_1 v_{i_1} + \dots + \lambda_k v_{i_k} = \lambda'_1 v'_{i'_1} + \dots + \lambda'_{k'} v_{i'_{k'}} .$$

Setze $J = \{i_1, \dots, i_k, i'_1, \dots, i'_{k'}\}$. Für $j \in J \setminus \{i_{1'}, \dots, i_{k'}\}$ setze $\lambda_j = 0$. Ebenso für $i \in J \setminus \{i_1, \dots, i_k\}$ setze $\lambda'_j = 0$.

$$\begin{aligned} &\Rightarrow \sum_{j \in J} \lambda_j v_j = \sum_{j \in J} \lambda'_j v_j \\ &\Rightarrow \sum_{j \in J} (\lambda_j - \lambda'_j)\, v_j = 0 \\ &\Rightarrow \lambda_j = \lambda'_j \ \ \forall\, j \in J \end{aligned}$$

weil $(v_i)_{i \in I}$ linear unabhängig.

$\boxed{\text{(iv)} \Rightarrow \text{(i)}}$ Sei $(v_i)_{i \in I}$ ein Erzeugendensystem von V, aus dem sich jeder Vektor eindeutig linear kombinieren lässt.
Behauptung: $(v_i)_{i \in I}$ ist linear unabhängig.
Angenommen $(v_i)_{i \in I}$ ist linear abhängig. Dann folgt, dass existieren i_0, $i_1, \dots, i_k \in I$, $\lambda_1, \dots, \lambda_k \in K$ mit $v_{i_0} = \lambda_1 v_{i_1} + \dots + \lambda_k v_{i_k}$. Also lässt sich v_{i_0} auf zwei verschiedene Arten linear kombinieren. □

Korollar 2.6.7 (Basisauswahlsatz). *Sei V ein K-Vektorraum und $v_1, \dots, v_n$ ein (endliches) Erzeugendensystem von V. Dann existiert eine Teilmenge $J \subseteq \{1, \dots, n\}$, so dass $(v_j)_{j \in I}$ Basis von V ist.*

Bemerkung 2.6.8.

- ▷ Insbesondere besitzt jeder endlich erzeugte Vektorraum eine Basis.
- ▷ Es gilt: *Jeder* Vektorraum besitzt eine Basis. Der Beweis für diese stärkere Aussage benutzt das *Auswahlaxiom.*

2.7. Lineare Gleichungssysteme

Wir betrachten das *lineare Gleichungssystem (LGS)*

$$\begin{array}{ccccccc} \alpha_{11}\, x_1 & + \dots + & \alpha_{1n}\, x_n & = & \beta_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \alpha_{m1}\, x_1 & + \dots + & \alpha_{mn}\, x_n & = & \beta_n \end{array} \tag{2}$$

Fragen:

(i) Gibt es eine Lösung?
(ii) Falls ja: Ist die Lösung eindeutig?
(iii) Falls nein: Beschreibe alle Lösungen.
(iv) Gibt es einen Algorithmus?

Definition 2.7.1. Das lineare Gleichungssystem (2) heißt *homogen*, falls $\beta_1 = \dots = \beta_n = 0$. Andernfalls heißt (2) *inhomogen*.

Bemerkung 2.7.2. Aus Beispiel 2.3.3 wissen wir, dass die Lösungsmenge eines homogenen linearen Gleichungssystems in n Unbestimmten über K ein Unterraum von K^n ist. Konkretisierung von Frage (iii) im homogenen Fall: Bestimme eine Basis!

Wir betrachten besonders einfache Spezialfälle von (2):

Beispiel 2.7.3. $\boxed{m = n = 1}$ Gegeben $\alpha, \beta \in K$, gesucht $x \in K$ in

$$\alpha\, x = \beta.$$

Die folgenden Fälle treten auf:

- ▷ $\alpha \neq 0 \Rightarrow x = \beta/\alpha$ ist die eindeutige Lösung.
- ▷ $\alpha = 0$ und $\beta \neq 0 \Rightarrow$ es existiert keine Lösung.
- ▷ $\alpha = 0$ und $\beta = 0 \Rightarrow$ jedes $x \in K$ ist Lösung.

Im homogenen Fall (d.h. $\beta = 0$) gilt:

- ▷ $\alpha \neq 0 \Rightarrow$ Lösungsmenge $= \{0\}$
- ▷ $\alpha = 0 \Rightarrow$ Lösungsmenge $= K$.

Beispiel 2.7.4. $\boxed{m = 1,\, n = 2}$ Für α_i, $\beta \in K$:

$$\alpha_1\, x_1 + \alpha_2\, x_2 = \beta.$$

<u>Fallunterscheidung:</u>

- ▷ $\alpha_1 \neq 0 \Rightarrow$ Für beliebiges $x_2 = \lambda \in K$ setze $x_1 = -\frac{\alpha_2}{\alpha_1}\lambda + \frac{\beta}{\alpha_1}$. Dann ist die Lösungsmenge
$$\left\{ \begin{pmatrix} \frac{-\alpha_2}{\alpha_1} + \frac{\beta}{\alpha} \\ \lambda \end{pmatrix} \mid \lambda \in K \right\}$$
Falls zusätzlich $\beta = 0$ ist
$$\left\{ \begin{pmatrix} -\frac{\alpha_2}{\alpha_1} \\ 1 \end{pmatrix} \right\} \text{Basis des Lösungsraums.}$$
—BILD—
Lösungsmenge im homogenen Fall: Gerade durch 0; im inhomogenen Fall: dazu parallele Gerade.
- ▷ $\alpha_1 = 0$, $\alpha_2 \neq 0 \Rightarrow x_2 = \frac{\beta}{\alpha_2}$.
Lösungsmenge
$$\left\{ \begin{pmatrix} \lambda \\ \frac{\beta}{\alpha_2} \end{pmatrix} \mid \lambda \in K \right\}$$
—BILD—
Falls zusätzlich $\beta = 0$, so ist $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ eine Basis des Lösungsraums
- ▷ $\alpha_1 = \alpha_2 = 0$ & $\beta \neq 0 \Rightarrow$ es existiert keine Lösung.
- ▷ $\alpha_1 = \alpha_2 = \beta = 0 \Rightarrow K^2$ ist Lösungsmenge und $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ ist eine Basis.

2.8. Der Gauß-Jordan-Eliminationsalgorithmus

Carl Friedrich Gauß (1777 – 1855), Camille Jordan (1838 – 1922).

Beobachtung: Die Lösungsmenge des linearen LGS (2) ändert sich nicht unter den folgenden *elementaren Zeilenoperationen:*

(E1) Addiere zu einer Gleichung das λ-fache einer anderen Gleichung für beliebige $\lambda \in K$.
(E2) Tausche zwei Gleichungen.
(E3) Multipliziere eine Gleichung mit $\lambda \in K \setminus \{0\}$.

Falls $\alpha_{11} \neq 0$ subtrahiere das

$\frac{\alpha_{21}}{\alpha_{11}}$-fache der 1. Gleichung von der zweiten

$\frac{\alpha_{31}}{\alpha_{11}}$-fache der 1. Gleichung von der dritten

$\vdots$

$\frac{\alpha_{m1}}{\alpha_{11}}$-fache der 1. Gleichung von der m-ten.

Falls $\alpha_{11} = 0$ finde ein $\alpha_{i1} \neq 0$ und vertausche die 1. Gleichung mit der i-ten.
Falls kein solches α_{i1} existiert, (d.h. $\alpha_{11} = \alpha_{21} = \cdots = \alpha_{m1} = 0$), dann tue nichts.

Algorithm A: Erster Schritt des Gauß-Jordanschen Eliminationsalgorithmus.

Danach sieht das modifizierte LGS so aus:

$$\begin{array}{cccccc}
\alpha'_{11}x_1+ & \alpha'_{12}x_2+ & \dots & +\alpha'_{1n}x_n & = & \beta'_1 \\
 & \alpha'_{22}x_2+ & \dots & +\alpha'_{2n}x_n & = & \beta'_2 \\
 & \vdots & & & \vdots & \vdots \\
 & \alpha'_{m2}x_2+ & \dots & +\alpha'_{mn}x_n & = & \beta'_n
\end{array}$$

Die unteren $m-1$ Gleichungen des modifizierten LGS können nun wie in Algorithmus A behandelt werden.

Nach $m-1$ Schritten hat das dann entstandene LGS die folgende *Zeilenstufenform*:

$$(3)\qquad \begin{array}{ccccccccc}
\gamma_{1j_1}x_{j_1} & + & \gamma_{1j_1+1}x_{j_1+1} & + & \dots & + & \gamma_{1n}x_n & = & \delta_1 \\
 & & \gamma_{2j_2}x_{j_2} & + & \dots & + & \gamma_{2n}x_n & = & \delta_2 \\
 & & & \ddots & & & & \vdots & \\
 & & & & \gamma_{r,j_r}x_{j_r} & + \dots + & \gamma_{rn}x_n & = & \delta_r \\
 & & & & & & 0 & = & \delta_{r+1} \\
 & & & & & & \vdots & \vdots & \vdots \\
 & & & & & & 0 & = & \delta_n
\end{array}$$

Dabei gilt $\forall k \in \{1, \dots, r\} : \gamma_{k,j_k} \neq 0$ und $j_1 < j_2 < \cdots < j_r \leq n$ sowie $0 \leq r \leq m$. Die Unbestimmten x_{j_k} heißen *Pivotvariablen*.

DEFINITION 2.8.1. Die Zahl r heißt *Rang* von (3).

Wie bestimmt man die Lösungsmenge L des LGS (3) (und damit die von (2))?

Fall 1: $\boxed{\delta_{r+1} \neq 0 \text{ oder } \delta_{r+2} \neq 0 \text{ oder } \dots \text{ oder } \delta_m \neq 0} \Rightarrow L = \emptyset$.

Fall 2: $\boxed{\delta_{r+1} = \cdots = \delta_m = 0}$ Dann gehen wir wie folgt vor:

> Wähle beliebige Werte aus K für jede der $n-r$ Nicht-Pivotvariablen.
> Löse dann die r-te Gleichung auf:
> $$x_{j_r} = \gamma_{r,j_r}^{-1}(\delta_r - \gamma_{r,j_r+1}x_{j_r+1} - \cdots - \gamma_{rn}x_n)$$
> Danach die $(r-1)$-te Gleichung etc.
> Die Lösung ist genau dann eindeutig, wenn gilt $r = n$ und $\delta_{r+1} = \cdots = \delta_m = 0$.

Algorithm B: Bestimmung der Lösungsmenge eines LGS in Zeilenstufenform.

Beispiel 2.8.2. LGS über $\mathbb{R}$ in Zeilenstufenform mit $m = 3$ und $n = 4$ ohne Lösung:

$$\begin{array}{ccccccccc} \pi x_1 & + & x_2 & + & 2x_3 & & & = & 0 \\ & & & & \sqrt{2}x_3 & - & x_4 & = & \frac{1}{2} \\ & & & & & & 1 & = & 2 \end{array}$$

Es gilt $r = 2$, $j_1 = 1$ und $j_2 = 3$. Die Pivotvariablen sind also x_1 und x_3.

Beispiel 2.8.3. Das LGS

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & 2x_3 & = & 0 \\ x_1 & + & 2x_2 & & & = & 1 \\ 2x_1 & & & + & x_3 & = & 2 \end{array}$$

über $\mathbb{F}_3$ mit $m = n = 3$ lässt sich in einem Gauß-Schritt umformen zu

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & 2x_3 & = & 0 \\ & & x_2 & + & x_3 & = & 1 \\ & & x_2 & + & & = & 2 \end{array}$$

Nach einem weiteren Gauß-Schritt entsteht die Zeilenstufenform

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & 2x_3 & = & 0 \\ & & x_2 & + & x_3 & = & 1 \\ & & & & 2x_3 & = & 1 \end{array}$$

Es gilt $r = 3$, $j_1 = 1$, $j_2 = 2$ und $j_3 = 3$. Alle Variablen sind Pivotvariablen. Auflösen ergibt

$$x_3 = 2, \quad x_2 = 1 - 2 = 2, \quad x_1 = 0 - 2 - 1 = 0 .$$

Die eindeutige Lösung ist also

$$x = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \in \mathbb{F}_3^3 .$$

2.8.1. Der homogene Fall. Im homogenen Fall sind sämtliche $\delta_k = 0$, und, falls $r < n$, können wir $n - r$ verschiedene Lösungen $b_1, \ldots, b_{n-r}$ wie folgt konstruieren:

- ▷ b_k entsteht wie in Algorithmus B erläutert, wobei man für die k-te Nicht-Pivotvariable 1 und für alle anderen Nicht-Pivotvariablen 0 wählt und anschließend die Werte für die Pivotvariablen ausrechnet.

Proposition 2.8.4. $(b_1, \ldots, b_{n-r})$ *ist eine Basis des Lösungsraums.*

Beweis. Übungsaufgabe. □

Falls $r = n$, so besitzt das homogene System als einzige Lösung die Nulllösung.

2.8.2. Der inhomogene Fall. Im inhomogegen Fall wird (wie in Algorithmus B) eine beliebige *spezielle Lösung* $\tilde{x}$ von (3) berechnet. Jede andere Lösung ergibt sich als Summe einer Lösung des zugehörigen homogenen Systems und $\tilde{x}$.

BEISPIEL 2.8.5. $K = \mathbb{C}$, $m = 3$, $n = 4$

$$\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 0 \\ & & 2x_2 & + & 3ix_3 & & & = & 1+2i \\ x_1 & + & 3x_3 & + & (1+3_i)x_3 & + & x_4 & = & 1+2i \end{array}$$

besitzt die Zeilenstufenform

$$\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 0 \\ & & 2x_2 & + & 3ix_3 & & & = & 1+2i \end{array}$$

Hier ist $r = 2$, und die Nicht-Pivotvariablen sind x_3 und x_4. *Spezielle Lösung* (für $\tilde{x}_3 = \tilde{x}_4 = 0$):

$$\tilde{x}_2 = \frac{1}{2} + i, \quad \tilde{x}_1 = -\frac{1}{2} - i.$$

Das zugehörige homogene System

$$\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 0 \\ & & 2x_2 & + & 3ix_3 & & & = & 0 \end{array}$$

hat Basislösungen

$$b_1 = \begin{pmatrix} b_{11} \\ b_{12} \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad b_2 = \begin{pmatrix} b_{21} \\ b_{22} \\ 0 \\ 1 \end{pmatrix},$$

wobei gilt $2b_{12} + 3i = 0$, also $b_{12} = -\frac{3}{2}i$ und $b_{11} = -b_{12} - 1 = -1 + \frac{3}{2}i$. Außerdem ist $2b_{22} = 0$, also $b_{22} = 0$ sowie $b_{21} = -b_{22} - 1 = -1$. Insgesamt ist die Lösungsmenge

$$\begin{pmatrix} -\frac{1}{2} - i \\ \frac{1}{2} + i \\ 0 \\ 0 \end{pmatrix} + \operatorname{lin} \left\{ \begin{pmatrix} -1 + \frac{3}{2}i \\ -\frac{3}{2}i \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

2.9. Der Basisaustauschsatz

Sei V für im gesamten Abschnitt ein K-Vektorraum, der endlich erzeugt ist.

LEMMA 2.9.1 (Austauschlemma). *Sei $(v_1, \dots, v_r)$ eine Basis von V und $w = \lambda_1 v_1 + \dots + \lambda_r v_r \in V$. Ist $k \in \{1, \dots, r\}$ mit $\lambda_k \neq 0$, dann ist*

$$(v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r)$$

wieder eine Basis von V.

BEWEIS. Wir können ohne Einschränkung annehmen, dass $k = 1$ ist; andernfalls nummerieren wir um.

Behauptung: $(w, v_2, \dots, v_r)$ ist Erzeugendensystem von V.

Sei $v \in V$ mit $v = \mu_1 v_1 + \dots + \mu_r v_r$, für geeignete $\mu_i \in K$. Es gilt $v_1 = \frac{1}{\lambda_1} w - \frac{\lambda_2}{\lambda_1} v_2 - \dots - \frac{\lambda_1}{\lambda_1} v_r$, und damit ist $v = \frac{\mu_1}{\lambda_1} v_1 + (\mu_2 - \frac{\mu_1 \lambda_2}{\lambda_1}) v_2 + \dots + (\mu_r - \frac{\mu_1 \lambda_r}{\lambda_1}) v_r$.

Behauptung: $(w, v_2, \dots, v_r)$ ist linear unabhängig.

Seien $\mu, \mu_2, \ldots, \mu_r \in K$ und $\mu w + \mu_2 v_2 + \cdots + \mu_r w_r = 0$. Dann folgt

$$\begin{aligned} &\mu\lambda_1 v_1 + \mu\lambda_2 + \cdots + \mu\lambda_r v_r + \mu_2 v_2 + \cdots + \mu_r v_r \\ &= \mu\lambda_1 v_1 + (\mu\lambda_2 + \mu_2) + \cdots + (\mu\lambda_r + \mu_r) v_r = 0 . \end{aligned}$$

Aus der linearen Unabhängigkeit von $(v_1, \ldots, v_r)$ folgt dann also

$$\mu\lambda_1 = \mu\lambda_2 + \mu_2 = \ldots = \mu\lambda_r + \mu_r = 0 .$$

Da nun $\lambda_1 \neq 0$ ist, folgt $\mu = 0$ und hieraus wiederum $\mu_2 = \cdots = \mu_r = 0$. □

SATZ 2.9.2 (Basisaustauschsatz). *Sei $(v_1, \ldots, v_r)$ eine Basis von V, und sei $(w_1, \ldots, w_n)$ eine linear unabhängige Familie von Vektoren in V. Dann gilt $n \leq r$, und es gibt Indizes $i_1, \ldots, i_n \in \{1, \ldots, r\}$, so dass $(w_1, \ldots, w_n, v_{i_1}, \ldots, v_{i_{r-n}})$ wieder eine Basis von V ist.*

BEWEIS DURCH INDUKTION NACH n.
Induktionsanfang: Für $n = 0$ ist nichts zu beweisen.
Induktionsvoraussetzung: Sei also $n \geq 1$, und sei $(w_1, \ldots, w_{n-1}, v_{i_1}, \ldots, v_{i_{r-n+1}})$ eine Basis von V.
Induktionsschritt:

(i) Behauptung: $n \leq r$
Nach Induktionsvoraussetzung wissen wir, dass $n - 1 \leq r$ ist. Angenommen $n = r + 1$. Dann ist $(w_1, \ldots, w_{n-1})$ Basis von V, aber $(w_1, \ldots, w_n)$ linear unabhängig. Dies steht im Widerspruch zu Lemma 2.2.4(ii).

(ii) Behauptung: es existiert $k \in \{1, \ldots, r - n + 1\}$ mit der Eigenschaft, dass $(w_1, \ldots, w_n, v_{i_1}, \ldots, v_{i_{k-1}}, v_{i_{k+1}}, \ldots, v_{i_{r-n+1}})$ Basis von V ist.
Nach Induktionsvoraussetzung existieren $\lambda_1, \ldots, \lambda_{n-1}, \mu_1, \ldots, \mu_{r-n+1} \in K$, so dass $w_n = \lambda_1 w_1 + \cdots + \lambda_{n-1} w_{n-1} + \mu_1 v_{i_1} + \cdots + \mu_{r-n+1} v_{i_{r-n+1}}$ ist. Falls $\mu_1 = \mu_2 = \cdots = \mu_{r-n+1} = 0$ wäre, so folgte, dass $(w_1, \ldots, w_n)$ linear abhängig ist. Dies ist aber nicht der Fall, und so existiert $k \in \{1, \ldots, r-n+1\}$ mit $\mu_k \neq 0$. Die Behauptung folgt dann aus dem Austauschlemma.

□

KOROLLAR 2.9.3. *Jede Basis von V ist endlich.*

BEWEIS. Nach Voraussetzung besitzt V eine endliche Basis $(v_1, \ldots, v_r)$. Angenommen, es existiert eine zweite Basis, die unendlich ist. Dann existiert eine linear unabhängige Familie $(w_1, \ldots, w_{r+1})$ der Länge $r + 1$, was aber im Widerspruch zum Austauschsatz steht. □

KOROLLAR 2.9.4. *Jede Basis von V hat dieselbe Länge.*

KOROLLAR 2.9.5 (Basisergänzungssatz). *Jede linear unabhängige Familie in V lässt sich zu einer Basis fortsetzen. Insbesondere ist jede linear unabhängige Familie höchstens so lang wie eine Basis.*

2.10. Der Dimensionsbegriff

DEFINITION 2.10.1. Ist V ein K-Vektorraum, so bezeichnet

$$\dim_K V := \begin{cases} r\,, & \text{falls } V \text{ eine Basis der Länge } r \text{ besitzt} \\ \infty & \text{sonst} \end{cases}$$

die *Dimension* von V über K.

BEISPIEL 2.10.2.

(i) $\dim_K K^n = n$ für $n \geq 1$.
(ii) $\dim_K\{0\} = 0$. Setze daher $K^0 := \{0\}$.
(iii) $\dim_\mathbb{C} \mathbb{C} = 1$ und $\dim_\mathbb{R} \mathbb{C} = 2$.
(iv) $\dim_\mathbb{Q} \mathbb{R} = \infty$ (ohne Beweis).

PROPOSITION 2.10.3. *Sei V ein K-Vektorraum mit $\dim_K V < \infty$ und $U < V$ ein echter Teilraum. Dann gilt $\dim_K U < \dim_K V$.*

BEWEIS. Sei $(u_1, \ldots, u_k)$ Basis von U und $(v_1, \ldots, v_n)$ Basis von V. Da $U \subsetneq V$ existiert $v \in V \setminus U$ und $v \notin \operatorname{lin}(u_1, \ldots, u_n)$. Aus Proposition 2.2.9 folgt also, dass $(u_1, \ldots, u_k, v)$ linear unabhängig ist. Aus Korollar 2.9.5 folgt $k + 1 \leq n$ und damit die Behauptung. $\square$

2.11. Exkurs: Matroide

In diesem Abschnitt geht es um weitreichende kombinatorische Verallgemeinerungen der zuvor untersuchten Begriffe.

Zu einer Menge M bezeichnen wir die Menge aller k-elementigen Teilmengen von M mit $\binom{M}{k}$. Falls M endlich ist mit *Kardinalität* $\#M = n$, so gilt $\mathcal{P}(M) = \bigcup_{k=0}^{n} \binom{M}{k}$.

DEFINITION 2.11.1. Sei E eine endliche Menge, $r \in \mathbb{N}$ und $\mathcal{B} \subseteq \binom{E}{r}$. Das Paar $(E, \mathcal{B})$ heißt *Matroid* vom *Rang* r, falls für je zwei verschiedene $A, B \in \mathcal{B}$ gilt, dass für jedes Element $a \in A$ ein $b \in B$ existiert, so dass $(A \setminus \{a\}) \cup \{b\} \in \mathcal{B}$ ist.

DEFINITION 2.11.2. Sei $(E, \mathcal{B})$ ein Matroid.

- ▷ Die Elemente aus $\mathcal{B}$ heißen *Basen* des Matroids.
- ▷ Eine Teilmenge $X \subseteq E$ heißt *unabhängig*, falls es eine Basis gibt, die X enthält; andernfalls heißt X *abhängig*.
- ▷ Ein Element aus E, das in keiner Basis vorkommt, heißt *Schleife*.
- ▷ Eine Teilmenge $C \subseteq E$ heißt *Kreis*, falls jede echte Teilmenge von C unabhängig ist.

BEISPIEL 2.11.3. Sei E eine beliebige endliche Menge. Dann ist $(E, \{\{x\} \colon x \in E\})$ ein Matroid vom Rang 1.

BEISPIEL 2.11.4. Sei $E = \{1, 2, 3, 4, 5\}$ und $\mathcal{B}$ bestehe aus den Mengen

$$\begin{array}{ccccc} \{1,2,3\} & \{1,2,5\} & \{1,3,4\} & \{1,3,5\} & \{1,4,5\} \\ \{2,3,4\} & \{2,3,5\} & \{2,4,5\} & \{3,4,5\} & \end{array}$$

Dann ist $(E, \mathcal{B})$ ein Matroid vom Rang 3.

2.11.1. Matroide aus Vektorräumen.

Es sei $(v_1, v_2, \ldots, v_n)$ eine endliche Familie von Vektoren eines K-Vektorraums V. Wir setzen $E := \{1, 2, \ldots, n\}$ und $U := \operatorname{lin}_K\{v_1, v_2, \ldots, v_n\}$ sowie

$$\mathcal{B} := \{B \subseteq E : \{v_i \colon i \in B\} \text{ ist Basis von } U\} .$$

SATZ 2.11.5. *Das Paar $(E, \mathcal{B})$ ist ein Matroid vom Rang $\dim_K U$.*

BEWEIS. Dies ist genau die Aussage des Basisaustauschsatzes, angewendet auf den K-Vektorraum U. $\square$

Frage: Was sind die Schleifen und Kreise eines solchen Matroids?

Beispiel 2.11.6. Das unter Beispiel 2.11.4 angegebene Matroid wird von der Vektorkonfiguration

$$(1,0,0) \quad (1,1,0) \quad (1,0,2) \quad (1,2,0) \quad (1,3,4)$$

in $\mathbb{Q}^3$ erzeugt.

2.11.2. Matroide aus Graphen.

Definition 2.11.7. Das Paar (V,E) heißt *(endlicher) Graph*, falls V eine endliche Menge und E eine beliebige Teilmenge der Menge $\{\{x,y\}: x,y \in V\}$ ist. Die Elemente aus V heißen *Knoten*, die Elemente aus E *Kanten*.

Beispiel 2.11.8. Sei V eine beliebige endliche Menge.

- ▷ $(V,\emptyset)$ ist ein Graph.
- ▷ $(V,\binom{V}{2})$ ist ein Graph, der *vollständige Graph* mit Knotenmenge V.

Sei $\Gamma = (V,E)$ ein endlicher *Graph*, der zusammenhängend ist. Ferner sei $\mathcal{B} \subset \mathcal{P}(E)$ die Menge aller Teilmengen von Kanten, die aufspannenden Bäumen von Γ entsprechen.

Satz 2.11.9. *Das Paar $(E,\mathcal{B})$ ist ein Matroid vom Rang $\#V-1$.*

Frage: Was sind die Schleifen und Kreise eines solchen Matroids?

Zusätzliche Literatur zum Thema:

P. Läuchli: Matroide, Eine Einführung für Mathematiker und Informatiker. Hochschulverlag vdf, Zürich 1998

KAPITEL 3

Lineare Abbildungen und Matrizen

3.1. Definitionen und ein erstes Beispiel

Seien V und W Vektorräume über demselben Körper K.

DEFINITION 3.1.1. Eine Abbildung $f : V \to W$ heißt K-*linear*, falls gilt

$$f(\lambda u + \mu v) \;=\; \lambda f(u) + \mu f(v)$$

für alle $u, v \in V$ und $\lambda, \mu \in K$.

Sei $f : V \to W$ eine lineare Abbildung.

DEFINITION 3.1.2.

(i) Die Menge
$$\operatorname{img} f \;:=\; \{f(v) : \, v \in V\} \subseteq W$$
heißt *Bild* von f.

(ii) Die Menge
$$\ker f \;:=\; \{v \in V : \, f(v) = 0\} \subseteq V$$
heißt *Kern* von f.

Für jede lineare Abbildung $f : V \to W$ gilt $f(0_V) = 0_W$, also $0_V \in \ker f$ und $0_W \in \operatorname{img} f$. Insbesondere sind Kern und Bild nie leer.

BEISPIEL 3.1.3. Sei V beliebig und $v_1, \dots, v_n \in V$. Dann ist die Abbildung

$$f : K^n \to V : x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i \, v_i$$

linear. Es gilt

$$\operatorname{img} f \;=\; \operatorname{lin}_K(v_1, \dots, v_n)$$

und

$$\ker f \;=\; \{x \in K^n : \, x_1 v_1 + \dots + x_n v_n = 0\} \, .$$

Hieraus folgt: $\ker f = 0 \Leftrightarrow (v_1, \dots, v_n)$ linear unabhängig.

PROPOSITION 3.1.4. *Eine lineare Abbildung $f : V \to W$ ist injektiv genau dann, wenn $\ker f = 0$ ist. Außerdem gilt für alle $u, v \in V$:*

$$f(u) = f(v) \;\Leftrightarrow\; u - v \in \ker f \, .$$

BEWEIS. Seien $u, v \in V$ beliebig. Dann gilt

$$0 \;=\; f(u) - f(v) \;=\; f(u - v) \quad \Leftrightarrow \quad u - v \in \ker f \, . \tag{4}$$

Sei nun $f : V \to W$ injektiv. Sei $x \in \ker f$. Also $f(x) = 0 = f(0)$. da f injektiv ist, folgt $x = 0$.

Sei umgekehrt $\ker f = 0$. Wir wollen zeigen, dass f injektiv ist. Seien $u, v \in V$ mit $f(u) = f(v)$. Dann ist nach unserer Vorüberlegung (4) $u - v \in \ker f = \{0\}$, woraus mit $u = v$ die Behauptung folgt. □

3.2. Dimensionsformel und Rang

Für das Folgende erweitern wir die Addition von $\mathbb{N}$ auf $\mathbb{N} \cup \{\infty\}$ durch:

$$u + \infty \;:=\; \infty \;=:\; \infty + u \qquad \text{für alle } m \in \mathbb{N} \cup \{\infty\}\,.$$

Wiederum seien V und W Vektorräume über K.

Satz 3.2.1 (Dimensionsformel). *Sei $f : V \to W$ eine K-lineare Abbildung. Dann gilt* $\dim_K \ker f + \dim_K \operatorname{img} f = \dim_K V$.

Beweis. Falls $\dim V = \infty$, so ist zu zeigen

$$\dim \ker f = \infty \;\text{ oder }\; \dim \operatorname{img} f = \infty\,. \tag{5}$$

Wir betrachten zunächst den Fall, dass $k := \dim \ker f \in \mathbb{N}$ und $l := \dim \operatorname{img} f \in \mathbb{N}$ ist. Wir müssen zeigen, dass $\dim V = k + l$ ist. Sollte uns dies gelingen, folgt dann auch (5) für $\dim V = \infty$ durch Kontraposition.

Nach unserer Voraussetzung existiert eine Basis $(v_1, \dots, v_k)$ von $\ker f$ und eine Basis $(w_1, \dots, w_l)$ von $\operatorname{img} f$. Für jedes $j \in \{1, \dots, l\}$ existiert ein $v_{k+j} \in V$ mit $f(v_{k+j}) = w_j$. Wir wollen zeigen, dass $(v_1, \dots, v_{k+l})$ ganz V erzeugt.

Sei $v \in V$. Dann gibt es $\mu_1, \dots, \mu_l \in K$ mit

$$\begin{aligned} f(v) &= \mu_1 w_1 + \cdots + \mu_l w_l \;=\; \mu_1 f(v_{k+1}) + \cdots + \mu_l f(v_{k+l}) \\ &= f(\mu_1 v_{k+1} + \cdots + \mu_l v_{k+l})\,. \end{aligned}$$

Hieraus folgt

$$v - \mu_1 v_{k+1} - \cdots - \mu_l v_{k+l} \in \ker f = \lim(v_1, \dots, v_l)\,.$$

Nun bleibt zu zeigen, dass $(v_1, \dots, v_{k+l})$ linear unabhängig sind. Seien $\lambda_1, \dots, \lambda_{k+l} \in K$ mit

$$\lambda_1 v_1 + \cdots + \lambda_{k+l} v_{k+l} \;=\; 0\,. \tag{6}$$

Dann ist

$$0 \;=\; f\Big(\sum_{i=1}^{k+l} \lambda_i v_i\Big) \;=\; \sum_{i=1}^{k+l} \lambda_i f_i(v_i) \;=\; \lambda_{k+1} w_1 + \cdots + \lambda_{k+l} w_l\,.$$

Aus der linearen Unabhängigkeit der $(w_1, \dots, w_l)$ ergibt sich $\lambda_{k+1} = \cdots = \lambda_{k+l} = 0$. Nun ist aber $(v_1, \dots, v_k)$ ebenfalls linear unabhängig, und daher folgt aus der Gleichung (6) $\lambda_1 = \cdots = \lambda_k = 0$. □

Satz 3.2.2. *Sei $f : V \to V$ eine lineare Abbildung und* $\dim V < \infty$. *Äquivalent sind:*

(i) f ist injektiv,
(ii) f ist surjektiv,
(iii) f ist bijektiv.

Beweis. Sei $n = \dim V$. Die Dimensionsformel lautet dann $\dim \ker f + \dim \operatorname{img} f = n$. Hieraus folgt

$$\begin{aligned} f \text{ injektiv} \quad &\Leftrightarrow \quad \ker f \;=\; 0 \quad \Leftrightarrow \quad \dim \ker f = 0 \\ &\Leftrightarrow \quad \dim \operatorname{img} f = n \quad \Leftrightarrow \quad \operatorname{img} f \;=\; V \quad \Leftrightarrow \quad f \text{ surjektiv}\,. \end{aligned}$$

□

Bemerkung 3.2.3. Für unendlich-dimensionale Vektorräume existieren stets injektive lineare Abbildungen, die nicht surjektiv sind und surjektive lineare Abbildungen, die nicht injektiv sind. Siehe Übungsaufgaben.

DEFINITION 3.2.4. Sei $f : V \to W$ linear über K. Die Zahl

$$\operatorname{rank}_K f \;:=\; \dim_K f(V) \;=\; \dim_K \operatorname{img} f$$

heißt *Rang* von f über K.

Falls $\dim V < \infty$, so gilt wegen der Dimensionsformel 3.2.1

$$\operatorname{rank} f \;=\; \dim V - \dim \ker f\,.$$

BEISPIEL 3.2.5. Wir setzen das Beispiel 3.1.3 fort, wobei wir spezialisieren $V = K^n$. Wiederum wählen wir $v_1, \ldots, v_n \in V$. Die Abbildung

$$f : K^n \to K^n : x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_1\, v_i$$

ist linear. Dann gilt

$$\begin{aligned} f \text{ bijektiv} \quad &\Leftrightarrow \quad f \text{ injektiv} \quad \Leftrightarrow \quad \ker f \;=\; 0 \\ &\Leftrightarrow \quad (v_1, \ldots, v_n) \text{ linear unabhängig} \\ &\Leftrightarrow \quad (v_1, \ldots, v_n) \text{ ist eine Basis}\,. \end{aligned}$$

3.3. Der Hauptsatz über lineare Abbildungen

Wiederum sind V und W beides K-Vektorräume.

PROPOSITION 3.3.1. *Sei $(v_1, \ldots, v_n)$ eine Basis von V und $f : V \to W$ eine lineare Abbildung. Es gilt:*

(i) $\operatorname{lin}_K(f(v_1), \ldots, f(v_n)) = \operatorname{img} f$.
(ii) $\operatorname{rank} f$ *ist genau die maximale Länge einer linear unabhängigen Teilfamilie von* $(f(v_1), \ldots, f(v_n))$.
(iii) f *surjektiv* $\Leftrightarrow \operatorname{rank} f = \dim W$.
(iv) f *injektiv* $\Leftrightarrow (f(v_1), \ldots, f(v_n))$ *linear unabhängig.*
(v) f *bijektiv* $\Leftrightarrow (f(v_1), \ldots, f(v_n))$ *Basis von* W.

BEWEIS.

(i) Sei $w \in \operatorname{img} f$. Dann existiert $v \in V$ mit $w = f(v)$, und existiert $\lambda_1, \ldots, \lambda_n \in K$, so dass $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Also ist $w = f(v) = f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n)$.
(ii) Übungsaufgabe.
(iii) Falls f surjektiv ist, also $\operatorname{img} f = W$, so gilt $\operatorname{rank} f = \dim \operatorname{img} f = \dim W$. Die Umkehrung folgt analog.
(iv) Falls f injektiv ist, also $\ker f = 0$, so gilt $\operatorname{rank} f = \dim V$. Also ist $(f(v_1), \ldots, f(v_n))$ linear unabhängig. Wiederum folgt die Umkehrung analog.
(v) Dies folgt aus (iii) und (iv).

□

SATZ 3.3.2 (Hauptsatz über lineare Abbildungen). *Sei $(v_1, \ldots, v_n)$ eine Basis von V, und seien $w_1, \ldots, w_n \in W$ beliebig. Dann existiert eine eindeutig bestimmte lineare Abbildung $f : V \to W$ mit der Eigenschaft*

$$f(v_i) \;=\; w_i \quad \textit{für alle } i \in \{1, 2, \ldots, n\}\,.$$

Beweis. Jeder Vektor $v \in V$ hat eine eindeutige Darstellung als Linearkombination

$$v \;=\; \lambda_1 v_1 + \cdots + \lambda_n v_n$$

mit $\lambda_i \in K$. Wir setzen $f(v) := \lambda_1 w_1 + \cdots + \lambda_n w_n$. Wegen $v_i = 0 \cdot v_1 + \cdots + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + \cdots + 0 \cdot v_n$ ist $f(v_i) = w_i$. Offenbar ist f linear. Angenommen es existiert eine weitere lineare Abbildung $g : V \to W$ mit $g(v_i) = w_i$ für alle $i \in \{1, 2, \ldots, n\}$. Dann gilt für beliebiges $v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in V$, dass

$$g(v) \;=\; \sum_{i=1}^{n} \lambda_i g(v_i) \;=\; \sum_{i=1}^{n} \lambda_i w_i \;=\; \sum_{i=1}^{n} \lambda_i f(v_i) \;=\; f(v)$$

ist. Insgesamt folgt $g = f$. □

Sei V ein K-Vektorraum mit $\dim V = n < \infty$. Wähle eine Basis $(v_1, \ldots, v_n)$ von V. Der Standard-K-Vektorraum K^n hat $(e_1, \ldots, e_n)$ als Basis. Wegen Satz 3.3.2 existiert eine eindeutig lineare Abbildung $f : K^n \to V$ mit $f(e_i) = v_i$ für alle i. Wegen Proposition 3.3.1(v) ist f bijektiv.

Definition 3.3.3. Eine bijektive K-lineare Abbildung heißt *K-Vektorraum-Isomorphismus*. Zwei K-Vektorräume V und W heißen *isomorph* über K, falls ein K-Vektorraum-Isomorphismus von V nach W existiert.

Korollar 3.3.4. *Sei V ein beliebiger n-dimensionaler K-Vektorraum mit $n < \infty$. Dann ist V isomorph zu K^n.*

Beispiel 3.3.5. Der Vektorraum der Polynome vom Grad $\leq d$ über dem Körper K ist als K-Vektorraum isomorph zu K^{d+1}.

3.4. Matrizen

3.4.1. Definition und Notation.

Definition 3.4.1. Sei X eine Menge und $m, n \in \mathbb{N} \setminus \{0\}$. Eine *$m \times n$-Matrix* M mit *Koeffizienten* in X ist eine Abbildung

$$M : \{1, \ldots, m\} \times \{1, \ldots, n\} \to X\,.$$

Bemerkung 3.4.2. Wir hatten die Elemente aus K^n als n-Tupel von Elementen aus K eingeführt. Bachten Sie, dass sich das n-Tupel $(v_1, v_2, \ldots, v_n)$ mit der Abbildung $1 \mapsto v_1, 2 \mapsto v_2, \ldots, n \mapsto v_n$ identifizieren lässt. Ein Zeilenvektor der Länge n ist somit auch eine $1 \times n$-Matrix, und ein Spaltenvektor der Länge m ist auch eine $m \times 1$-Matrix.

Üblicherweise schreibt man eine Matrix $M : \{1, \ldots, m\} \times \{1, \ldots, n\} \to X$ als rechteckiges Schema:

$$M \;=\; \begin{pmatrix} M(1,1) & \ldots & M(1,n) \\ \vdots & \ddots & \vdots \\ M(m,1) & \ldots & M(m,n) \end{pmatrix}$$

3.4.2. Matrizen aus linearen Abbildungen, Teil I. Seien V und W Vektorräume über K und $f : V \to W$ eine lineare Abbildung. Seien $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_m)$ Basen von V bzw. W. Für jedes $j \in \{1, \ldots, n\}$ existiert eindeutig bestimmte $\mu_{1j}, \ldots, \mu_{mj} \in K$ mit

$$f(v_j) \;=\; \mu_{1j} w_1 + \cdots + \mu_{mj} w_m\,.$$

Definition 3.4.3. Die Matrix

$$\begin{aligned}[f]^{\mathcal{C}}_{\mathcal{B}} &: \{1,\dots,m\}\times\{1,\dots,n\}\to K\\ &: (i,j)\mapsto \mu_{ij}\end{aligned}$$

heißt *(darstellende) Matrix von f* bezüglich $\mathcal{B}$ und $\mathcal{C}$.

Beispiel 3.4.4. Sei $\mathbb{F}_2=\{0,1\}$ der Körper mit 2 Elementen und

$$V = \mathbb{F}_2^2 = \left\{ \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\}.$$

Die Vektoren $e_1=\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ und $e_2=\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ bilden die Standardbasis von V. Die Abbildung

$$\varphi : \mathbb{F}_2^2\to\mathbb{F}_2^2 : \begin{pmatrix}x_1\\x_2\end{pmatrix}\mapsto\begin{pmatrix}0\\x_2\end{pmatrix}$$

ist linear; vergleiche Beispiel 3.2.5. Bezüglich der Standardbasis $\mathcal{B}=(e_1,e_2)$ gilt

$$[\varphi]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

3.5. Vektorräume von linearen Abbildungen

Seien V und W Vektorräume über K. Dann ist die Menge W^V aller Abbildungen von V nach W mit der punktweisen Addition und punktweisen Skalarmultiplikation ein K-Vektorraum.

Definition 3.5.1.

(i) $\mathrm{Hom}_K(V,W) := \{\varphi : V\to W : \varphi \text{ linear über } K\}$. Die Elemente von $\mathrm{Hom}_K(V,W)$ heißen auch *K-Vektorraum-Homomorphismen*.
(ii) $\mathrm{End}_K(V) := \mathrm{Hom}_K(V,V)$. Die Elemente aus $\mathrm{End}_K(V)$ heißen auch *K-Vektorraum-Endomorphismen*.

Proposition 3.5.2. *Die Menge $\mathrm{Hom}_K(V,W)$ ist ein K-Unterraum von W^V.*

Beweis. Übungsaufgabe. □

Im Folgenden seien U,V,W stets K-Vektorräume.

Proposition 3.5.3. *Es gilt:*

(i) Die Identität $\mathrm{id}_V : V\to V : v\mapsto v$ ist linear.
(ii) Ist $\varphi : U\to V$ linear und bijektiv, dann ist auch die Umkehrabbildung $\varphi^{-1} : V\to U$ linear.
(iii) Sind $\varphi : U\to V$ und $\psi : V\to W$ linear, so ist auch $\psi\circ\varphi : U\to W$ linear.

Beweis.

(i) Offensichtlich ist die Identität auf V eine lineare Abbildung.
(ii) Seien $v_1,v_2\in V$ und $\lambda_1,\lambda_2\in K$. Da $\varphi : U\to V$ bijektiv ist, existieren eindeutig u_1, $u_2\in U$ mit $\varphi(u_i)=v_i$ für $i=1,2$. Es ergibt sich

$$\begin{aligned}\varphi^{-1}(\lambda_1v_1+\lambda_2v_2) &= \varphi^{-1}(\lambda_1\varphi(u_1)+\lambda_2\varphi(u_2))\\ &= \varphi^{-1}(\varphi(\lambda_1u_1+\lambda_2u_2)) = \lambda_1u_1+\lambda_2u_2\\ &= \lambda_1\varphi^{-1}(v_1)+\lambda_2\varphi^{-1}(v_2),\end{aligned}$$

und φ^{-1} ist linear.
(iii) Übungsaufgabe.

□

Proposition 3.5.4. *Seien $\varphi_1, \varphi_2 : U \to V$ sowie $\psi_1, \psi_2 : V \to W$ lineare Abbildungen. Dann gilt:*

(i) $\psi_1 \circ (\varphi_1 + \varphi_2) = \psi_1 \circ \varphi_1 + \psi_1 \circ \varphi_2$,
(ii) $(\psi_1 + \psi_2) \circ \varphi_1 = \psi_1 \circ \varphi_1 + \psi_2 \circ \varphi_1$,
(iii) $(\lambda\psi_1) \circ \varphi_1 = \lambda(\psi_1 \circ \varphi_1)$ *für alle* $\lambda \in K$.

Beweis. Wir beweisen die erste Behauptung. Sei $u \in U$. Dann ist

$$\begin{aligned}[]
[\psi_1 \circ (\varphi_1 + \varphi_2)](u) &= \psi_1(\varphi_1(u) + \varphi_2(u)) \\
&= \psi_1(\varphi_1(u)) + \psi_1(\varphi_2(u)) \;=\; [\psi_1 \circ \varphi_1 + \psi_1 \circ \varphi_2](u)\,.
\end{aligned}$$

Die beiden anderen Aussagen zeigt man analog. □

Satz 3.5.5. *Das Tripel* $(\mathrm{End}_K(V), +, \circ)$ *ist ein Ring mit Eins. Das multiplikative Neutralelement ist die Identität* id_V.

Beweis.

- ▷ $(\mathrm{End}(V), +)$ ist eine abelsche Gruppe.
- ▷ Die Verkettung beliebiger Abbildungen ist assoziativ, siehe Satz 1.3.13.
- ▷ Die Distributivgesetze gelten wegen Proposition 3.5.4.

□

Bemerkung 3.5.6. Da $\mathrm{End}_K(V)$ zusätzlich ein K-Vektorraum ist und außerdem Proposition 3.5.4(iii) gilt, ist $\mathrm{End}_K(V)$ sogar eine *K-Algebra*.

3.6. Die allgemeine lineare Gruppe

Definition 3.6.1. $\mathrm{GL}_K(V) := \{\varphi : V \to V : \ \varphi \text{ linear und bijektiv}\}$.

Bemerkung 3.6.2. $(\mathrm{GL}_K(V), \circ)$ ist eine Gruppe, genannt die *allgemeine lineare Gruppe* auf V.

Beispiel 3.6.3. Wir setzen das Beispiel 3.4.4 mit $K = \mathbb{F}_2$ und $V = \mathbb{F}_2^2$ fort. Als Basis von V wählen wir wiederum die Standardbasis aus $e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ und $e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. Die Menge $\mathrm{End}_{\mathbb{F}_2}(\mathbb{F}_2^2)$ besteht aus 16 linearen Abbildungen, deren Matrizen bezüglich der Standarsbasis folgendermaßen aussehen:

$$\begin{matrix}
\begin{pmatrix}0&0\\0&0\end{pmatrix} & \begin{pmatrix}0&0\\0&1\end{pmatrix} & \begin{pmatrix}0&0\\1&0\end{pmatrix} & \begin{pmatrix}0&0\\1&1\end{pmatrix} \\
\begin{pmatrix}0&1\\0&0\end{pmatrix} & \begin{pmatrix}0&1\\0&1\end{pmatrix} & \begin{bmatrix}0&1\\1&0\end{bmatrix} & \begin{bmatrix}0&1\\1&1\end{bmatrix} \\
\begin{pmatrix}1&0\\0&0\end{pmatrix} & \begin{bmatrix}1&0\\0&1\end{bmatrix} & \begin{pmatrix}1&0\\1&0\end{pmatrix} & \begin{bmatrix}1&0\\1&1\end{bmatrix} \\
\begin{pmatrix}1&1\\0&0\end{pmatrix} & \begin{bmatrix}1&1\\0&1\end{bmatrix} & \begin{bmatrix}1&1\\1&0\end{bmatrix} & \begin{pmatrix}1&1\\1&1\end{pmatrix}
\end{matrix}$$

Die mit eckigen Klammern markierten Matrizen entsprechen dabei genau den invertierbaren linearen Abbildungen. Es gilt also $\#\,\mathrm{GL}_2\,\mathbb{F}_2 = 6$.

3.7. Vektorräume von Matrizen

Sei K im Folgenden stets ein Körper.

Definition 3.7.1. Seien $m, n \in \mathbb{N}$. Dann ist

$$K^{m\times n} \ := \ K^{\{1,\dots,m\}\times\{1,\dots,n\}}$$

die Menge aller $m \times n$-Matrizen über K.

Notation:

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \ =: \ (\alpha_{ij})_{1\le i\le m,\, 1\le j\le n}.$$

Proposition 3.7.2. *Mit komponentenweiser Addition und Skalarmultiplikation ist $K^{m\times n}$ ein K-Vektorraum der Dimension mn.*

Beweis. Übungsaufgabe. □

3.7.1. Matrixmultiplikation. Wir definieren eine weitere Verknüpfung: Seien $(\alpha_{ij}) \in K^{l\times m}$ und $(\beta_{ij}) \in K^{m\times n}$. Dann ist

$$(\alpha_{ij}) \cdot (\beta_{ij}) \ := \ (\gamma_{ij}) \in K^{l\times n}$$

mit

$$\gamma_{ik} \ := \ \sum_{j=1}^{m} \alpha_{ij} \cdot \beta_{jk}$$

die *Produktmatrix* von (α_{ij}) und (β_{ij}). Die Abbildung $\cdot : K^{l\times m} \times K^{m\times n} \to K^{l\times n}$ heißt *Matrixmultiplikation.*

Proposition 3.7.3. *Die Matrixmultiplikation $\cdot$ ist assoziativ.*

Beweis. Übungsaufgabe. □

Beispiel 3.7.4. $K = \mathbb{Q}$ und $l = 2$, $m = 3$, $n = 2$.

$$\begin{pmatrix} 2 & 0 & 1 \\ -1/2 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1/3 \\ 0 & 2 \\ -1 & -1 \end{pmatrix} \ = \ \begin{pmatrix} 1 & -1/3 \\ -1/2 & 35/6 \end{pmatrix}$$

Beispiel 3.7.5. $K = \mathbb{F}_2$ und $l = m = n = 2$.

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

Die folgende Aussage entspricht Proposition 3.5.4.

Proposition 3.7.6. *Seien $A_1, A_2 \in K^{l\times m}$ und $B_1, B_2 \in K^{m\times n}$ Matrizen sowie $\lambda \in K$. Dann gilt:*

(i) $A_1(B_1 + B_2) = A_1B_1 + A_1B_2$,
(ii) $(A_1 + A_2)B_1 = A_1B_1 + A_2B_1$,
(iii) $(\lambda A_1)B_1 = \lambda(A_1B_1)$.

Beweis. Wir zeigen die zweite Behauptung. Seien $A_1 = (\alpha_{ij}^{(1)})$, $A_2 = (\alpha_{ij}^{(2)})$, $B_1 = (\beta_{ij})$ und $(A_1 + A_2)B_1 = (\gamma_{ij})$. Dann gilt

$$\gamma_{ik} \;=\; \sum_{j=1}^{m}(\alpha_{ij}^{(1)} + \alpha_{ij}^{(2)})\beta_{jk} \;=\; \sum_{j=1}^{m}\alpha_{ij}^{(1)}\beta_{jk} + \sum_{j=1}^{m}\alpha_{ij}^{(2)}\beta_{jk}\,.$$

Ferner ist

$$A_1B_1 \;=\; (\sum_{j=1}^{m}\alpha_{ij}^{(1)}\beta_{jk})_{i,k} \quad \text{und} \quad A_2B_1 \;=\; (\sum_{j=1}^{m}\alpha_{ij}^{(2)}\beta_{jk})_{i,k}\,,$$

und hieraus folgt die Behauptung. Analog für die anderen beiden Aussagen. □

3.7.2. Die K-Algebra der quadratischen Matrizen. Die Matrix

$$E_n \;=\; \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \;\in\; K^{n\times n}$$

heißt *n-reihige Einheitsmatrix*. Eine Matrix $A \in K^{n\times n}$ heißt *invertierbar*, falls $B \in K^{n\times n}$ existiert mit $AB = E_n$. Wegen $E_nE_n = E_n$ ist insbesondere E_n invertierbar.

Definition 3.7.7. $\mathrm{GL}_n\,K := \{A \in K^{n\times n} :\ A \text{ invertierbar}\}$.

Proposition 3.7.8. $(\mathrm{GL}_n\,K, \cdot)$ *ist eine Gruppe.*

Beweis.

- ▷ Seien $A, C \in \mathrm{GL}_n\,K$. Dann existieren $B, D \in K^{n\times n}$ mit $AB = CD = E_n$. Wegen der Assoziativität der Matrixmultiplikation ist $(AC)(DB) = A(CD)B = AB = E_n$, und es folgt $AC \in \mathrm{GL}_n\,K$.
- ▷ Die Einheitsmatrix E_n ist das Neutralelement.
- ▷ Die Matrixmultiplikation ist assoziativ.
- ▷ Invertierbarkeit sichert die Existenz von Inversen per Definition. Man rechnet nach, dass aus $AB = E_n$ auch $BA = E_n$ folgt.[1] Dies bedeutet, dass die Inversen auch tatsächlich wieder in $\mathrm{GL}_n\,K$ liegen.

□

Beispiel 3.7.9. Die allgemeine lineare Gruppe $\mathrm{GL}_1\,K$ ist (als Gruppe) isomorph zur multiplikativen Gruppe des Körpers K.

Beispiel 3.7.10.

$$\mathrm{GL}_2\,\mathbb{F}_2 \;=\; \{\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)\}$$

Eine Matrix $(\alpha_{ij})_{i,j} \in K^{n\times n}$ heißt *Diagonalmatrix*, falls $\alpha_{ij} = 0$ für alle $i \neq j$. Zu $\delta_1, \delta_2, \dots, \delta_n \in K$ sei $\mathrm{diag}(\delta_1, \delta_2, \dots, \delta_n) := (\alpha_{ij})_{i,j})$ mit

$$\alpha_{ij} \;:=\; \begin{cases} \delta_i & \text{für } i = j \\ 0 & \text{sonst} \end{cases}$$

die zugehörige Diagonalmatrix in $K^{n\times n}$. Aus der Definition der Matrixmultiplikation ergibt sich unmittelbar, dass

$$\mathrm{diag}(\delta_1, \delta_2, \dots, \delta_n) \cdot \mathrm{diag}(\eta_1, \eta_2, \dots, \eta_n) \;=\; \mathrm{diag}(\delta_1\eta_1, \delta_2\eta_2, \dots, \delta_n\eta_n)\,. \tag{7}$$

Insbesondere ergibt sich die folgende Proposition.

[1] Dies gilt, obwohl die Matrixmultiplikation im Allgemeinen nicht kommutativ ist.

PROPOSITION 3.7.11. *Diagonalmatrizen sind genau dann invertierbar, falls alle Diagonaleinträge von $0 \in K$ verschieden sind.*

Analog zu Satz 3.5.5 gilt das Folgende.

SATZ 3.7.12. *Das Tripel $(K^{n\times n}, +, \cdot)$ ist ein Ring mit Eins. Das multiplikative Neutralelement ist die Einheitsmatrix E_n.*

BEMERKUNG 3.7.13. Wiederum analog zu Bemerkung 3.5.6 ist $K^{n\times n}$ wegen Proposition 3.7.6(iii) eine K-Algebra. Ihre *Einheitengruppe* ist $\mathrm{GL}_n K$.

3.8. Lineare Abbildungen aus Matrizen

Sei K im Folgenden wieder ein Körper.

Wir nehmen eine Matrix $A = (\alpha_{ij}) \in K^{m\times n}$ und einen Vektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. Wenn wir K^n mit $K^{n\times 1}$ identifizieren, erhalten wir durch Spezialisierung der Matrixmultiplikation

$$Ax = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_1 & +\dots+ & \alpha_{1n}x_n \\ \vdots & \ddots & \vdots \\ \alpha_{m1}x_1 & +\dots+ & \alpha_{mn}x_n \end{pmatrix} \in K^m .$$

BEMERKUNG 3.8.1. Das lineare Gleichungssystem

$$\begin{array}{ccccc} \alpha_{11}\,x_1 & +\dots+ & \alpha_{1n}\,x_n & = & \beta_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \alpha_{m1}\,x_1 & +\dots+ & \alpha_{mn}\,x_n & = & \beta_n \end{array}$$

lässt sich nun schreiben als $Ax = b$, wobei $A = (\alpha_{ij})_{i,j} \in K^{m\times n}$ und $b = (\beta_i)_i \in K^m$. Hierbei ist

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

ein Spaltenvektor von Unbestimmten. Die Matrix $(A|b) \in K^{m\times(n+1)}$, die dadurch entsteht, dass man die Spalte b and die Matrix A rechts anfügt, heißt auch *erweiterte Koeeffizientenmatrix* des linearen Gleichungssystems $Ax = b$.

LEMMA 3.8.2. *Die Abbildung*

$$\varphi_A : K^n \to K^m : x \mapsto Ax .$$

ist linear.

BEWEIS. Für $x, y \in K^n$ und $\lambda, \mu \in K$ ist

$$\begin{aligned} \varphi_A(\lambda x + \mu y) &= A(\lambda x + \mu y) \overset{3.7.6(\mathrm{i})}{=} A(\lambda x) + A(\mu y) \overset{3.7.6(\mathrm{iii})}{=} (\lambda A)x + (\mu A)y \\ &= \lambda\varphi_A(x) + \mu\varphi_A(y) . \end{aligned}$$

□

Die Spalten von A sind genau die Bilder der Standardbasisvektoren unter der linearen Abbildung φ_A:

LEMMA 3.8.3. *Es gilt für alle $i \in \{1, \dots, n\}$, dass*

$$\varphi_A(e_i) \;=\; Ae_i \;=\; \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{mi} \end{pmatrix} \;=\; i\text{-te Spalte von } A\,.$$

Hieraus folgt unmittelbar

$$\begin{aligned} \operatorname{img} \varphi_A &= \varphi_A(K^n) \;=\; \{b \in K^m : \exists x \in K^n : Ax = b\} \\ &= \operatorname{lin}_K(\varphi_A(e_1), \dots, \varphi_A(e_n)) \\ &= \text{Unterraum von } K^m\text{, der von den Spalten von } A \text{ aufgespannt wird} \\ &=: \textit{Spaltenraum} \text{ von } A\,. \end{aligned}$$

Nochmals anders ausgedrückt: Im Bild der Abbildung φ_A liegen genau diejenigen Vektoren $b \in K^m$, für die das lineare Gleichungssystem $Ax = b$ eine Lösung besitzt.

3.8.1. Der Rang einer Matrix.

DEFINITION 3.8.4. Der *Rang* der Matrix A is definiert als

$$\operatorname{rank}_K A \;:=\; \operatorname{rank}_K \varphi_A \;=\; \dim_K \operatorname{lin}_K(\varphi_A(e_1), \dots, \varphi_A(e_n))\,.$$

BEISPIEL 3.8.5. Betrachte

$$A \;=\; \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 2 & 0 & 2 \end{pmatrix} \;\in\; \mathbb{F}_3^{3\times 3}\,.$$

Wende Gauß-Jordan-Elimination an auf das GLS $Ax = 0$, und wir erhalten eine 3×3-Matrix über $\mathbb{F}_3$ in Zeilenstufenform:

$$B \;=\; \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}\,.$$

Der Rang des zugehörigen linearen Gleichungssystems in Zeilenstufenform ist zwei. Außerdem gilt:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}\,,$$

und hieraus folgt, dass $\operatorname{rank}_{\mathbb{F}_3} B = \operatorname{rank}_{\mathbb{F}_3} A = 2$ ist.

SATZ 3.8.6. *Der Rang einer Matrix $A \in K^{m\times n}$ entspricht stets dem Rang einer Zeilenstufenform aus dem Gauß-Jordan-Algorithmus, angewendet auf das lineare Gleichungssystem $Ax = b$ für $b \in K^m$ beliebig.*

BEWEIS. Wir zeigen: Die elementaren Zeilenoperationen (E1), (E2) und (E3) verändern den Rang der Matrix A nicht. Beachten Sie, dass die rechte Seite b für die Definition des Rangs einer Zeilenstufenform eines linearen Gleichungssystems unerheblich ist.

(E1) Addiere zu einer Zeile i das λ-fache der Zeile j: Dann ist die transformierte Matrix $A' = LA$, wobei $L = E_m + \lambda E_{ij}$ und

$$E_{ij} \;=\; (\eta_{kl})_{k,l} \quad \text{mit} \quad \eta_{kl} \;=\; \begin{cases} 1 & \text{falls } k = i \text{ und } j = l \\ 0 & \text{sonst} \end{cases}$$

Wegen $i \neq j$ folgt $E_{ij}^2 = 0$. Also ist $(E_m + \lambda E_{ij})(E_m - \lambda E_{ij}) = E_m$, und insbesondere ist L invertierbar. Diese Matrix L ist eine *Transvektion*.

(E2) Tausche zwei Reihen i und j: Dann ist die transformierte Matrix $A' = LA$ mit

$$L = (\pi_{kl})_{k,l} \quad \text{mit} \quad \pi_{kl} = \begin{cases} 1 & \text{falls } k = l \not\in \{i,j\} \\ 1 & \text{falls } (k,l) \in \{(i,j),(j,i)\} \\ 0 & \text{sonst} \end{cases}$$

Es folgt $L^2 = E_m$, insbesondere also ist $L \in \mathrm{GL}_m(K)$. Diese Matrix L ist eine *Permutationsmatrix*.

(E3) Multipliziere die i-te Reihe mit $\lambda \in K \setminus \{0\}$: Dann ist die transformierte Matrix $A' = LA$, wobei

$$L = \mathrm{diag}(1,\ldots,1, \underbrace{\lambda}_{i\text{-te Stelle}}, 1,\ldots,1)$$

Die Proposition 3.7.11 besagt, dass wiederum L invertierbar ist.

Zusammenfassend stellt sich heraus, dass sich alle drei elementaren Zeilenoperationen als Multiplikation der (erweiterten) Koeffizientenmatrix mit einer invertierbaren $m \times m$-Matrix von links schreiben lassen.

Sei nun A'' Zeilenstufenform zu A. Dann existieren $L_1, \ldots, L_k \in \mathrm{GL}_m(K)$, so dass $A'' = L_k L_{k-1} \ldots L_1 A$ bzw. $A'' = L'' A$ für $L'' = L_k L_{k-1} \ldots L_1 \in \mathrm{GL}_m K$ ist. In Satz 3.8.11 werden wir zeigen, dass gilt:

$$\varphi_{A''} = \varphi_{L''} \circ \varphi_A .$$

Außerdem ist $\varphi_{L''}$ invertierbar und daher $\mathrm{lin}_K \varphi_{A''} = \varphi_{L''}(\mathrm{lin}_K \varphi_A)$, also

$$\mathrm{rank}_K A'' = \dim_K \mathrm{img}\, \varphi_{A''} = \dim_K \mathrm{img}\, \varphi_A = \mathrm{rank}_K A .$$

□

3.8.2. Äquivalenz von linearen Abbildungen und Matrizen.

Satz 3.8.7. *Die Abbildung*

$$\Phi : K^{m\times n} \to \mathrm{Hom}_K(K^n, K^m) : A \mapsto \varphi_A$$

ist ein K-linearer Isomorphismus. Die inverse Abbildung Φ^{-1} ordnet einer linearen Abbildung $\varphi : K^n \to K^m$ die Matrix $[\varphi]$ von φ bezüglich der Standardbasen von K^n und K^m zu.

Beweis.

Behauptung: Φ ist bijektiv.

Seien $A, B \in K^{m\times n}$ mit $\varphi_A = \varphi_B$. Offenbar gilt $[\varphi_A] = A$. Damit ergibt sich für $\Phi(A) = \Phi(B)$, also $\varphi_A = \varphi_B$, dass dann $A = [\varphi_A] = [\varphi_B] = B$ ist. Dies besagt, dass Φ injektiv ist. Sei $\varphi \in \mathrm{Hom}(K^m, K^n)$ beliebig. Dann ist $\Phi([\varphi]) = \varphi$, das heißt Φ ist surjektiv.

Behauptung: Φ ist linear.

Seien $A, B \in K^{m\times n}$ und $\lambda, \mu \in K$ sowie $x \in K^n$.

$$\begin{aligned} \Phi(\lambda A + \mu B)(x) &= \varphi_{\lambda A + \mu B}(x) = (\lambda A + \mu B)(x) \\ &= \lambda A x + \mu B x = \lambda \varphi_A(x) + \mu \varphi_B(x) = (\lambda \Phi(A) + \mu \Phi(B))(x) . \end{aligned}$$

□

Bemerkung 3.8.8. In Proposition 3.5.3 wurde gezeigt, dass die Umkehrableitung einer bijektiven linearen Abbildung selbst wieder linear ist, angewendet auf Φ bedeutet dies:

$$[\lambda\varphi + \mu\psi] \;=\; \lambda[\varphi] + \mu[\psi]$$

für alle $\varphi, \psi \in \operatorname{Hom}(K^m, K^n)$ und $\lambda, \mu \in K$.

Korollar 3.8.9. $\dim_K \operatorname{Hom}_K(K^n, K^m) = mn.$

Beispiel 3.8.10. Sei $K = \mathbb{Q}$ und $m = n = 2$. Betrachte

$$A \;=\; \begin{pmatrix} 2 & -1 \\ 1/2 & 0 \end{pmatrix} \quad \text{und} \quad B \;=\; \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Sei $x = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$ beliebig. Dann ist

$$\begin{aligned} \Phi(A+B)(x) \;&=\; \begin{pmatrix} 0 & -1 \\ 1/2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; \begin{pmatrix} -x_2 \\ 1/2x_1 + 2x_2 \end{pmatrix} \\ &=\; \begin{pmatrix} 2x_1 - x_2 \\ 1/2x_1 \end{pmatrix} + \begin{pmatrix} -2x_1 \\ 2x_2 \end{pmatrix} \;=\; \Phi(A)(x) + \Phi(B)(x)\,. \end{aligned}$$

Für die Abbildungen heißt dies $\Phi(A+B) = \varphi_{A+B} = \varphi_A + \varphi_B = \Phi(A) + \varphi(B)$.

Satz 3.8.11.

(i) Für $A \in K^{l\times m}$ und $B \in K^{m\times n}$ gilt $\varphi_A \circ \varphi_B = \varphi_{AB}$.

(ii) Für $\psi \in \operatorname{Hom}(K^l, K^m)$ und $\varphi \in \operatorname{Hom}(K^m, K^n)$ gilt $[\varphi \circ \psi] = [\varphi] \cdot [\psi]$.

(iii) Die Abbildung

$$\Phi \;:\; K^{n\times n} \to \operatorname{End}(K^n) \;:\; A \mapsto \varphi_A$$

ist (ein K-Vektorraumisomorphismus und) ein Ringisomorphismus. Insgesamt ist Φ ein Isomorphismus von K-Algebren.

Beweis. Wir beweisen zunächst die erste Aussage. Hierzu seien $A \in K^{l\times m}$ und $B \in K^{m\times n}$. Dann gilt für alle $x \in K^n$, dass

$$(\varphi_A \circ \varphi_B)(x) \;=\; \varphi_A(\varphi_B(x)) \;=\; \varphi_A(Bx) \;=\; A(Bx) \;=\; (AB)x \;=\; \varphi_{AB}(x)\,.$$

Die zweite Aussage beweist man analog, und die dritte Aussage ist ein Spezialfall der ersten. □

3.9. Die Transponierte einer Matrix

Sei $A = (\alpha_{ij})_{i,j} \in K^{m\times n}$.

Definition 3.9.1. Die Matrix

$$A^{\mathrm{tr}} \;:=\; (\alpha_{ji})_{i,j} \;=\; \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & & \alpha_{m2} \\ \vdots & & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{mn} \end{pmatrix} \;\in\; K^{n\times m}$$

heißt *Transponierte* von A.

Offenbar gilt $A^{\mathrm{tr}\,\mathrm{tr}} = A$, und Diagonalmatrizen ändern sich nicht unter Transposition.

Proposition 3.9.2. *Seien $A \in K^{l\times m}$ und $B \in K^{m\times n}$. Dann gilt $(AB)^{\mathrm{tr}} \;=\; B^{\mathrm{tr}} \cdot A^{\mathrm{tr}}$.*

Beweis. Übungsaufgabe. □

Korollar 3.9.3. *Wenn $A \in \mathrm{GL}_n\, K$ ist, so auch A^{tr}, und es gilt $(A^{-1})^{\mathrm{tr}} = (A^{\mathrm{tr}})^{-1}$.*

Beweis. Es gilt $A^{\mathrm{tr}}(A^{-1})^{\mathrm{tr}} = (A^{-1}A)^{\mathrm{tr}} = {E_n}^{\mathrm{tr}} = E_n$. □

Bemerkung 3.9.4. Der *Zeilenrang* von A ist definiert als die Anzahl linear unabhängiger Zeilen von A. Dies ist dasselbe wie der (Spalten-)rang der Transponierten A^{tr}.

Korollar 3.9.5. *Für $A \in K^{m\times n}$ gilt:* $\operatorname{rank}_K A^{\mathrm{tr}} = \operatorname{rank}_K A$.

Beweis. Sei $A'' = L''A$ Zeilenstufenform von A wie im Beweis zu Satz 3.8.6. Man sieht direkt: Der Rang von A'' ist dasselbe wie der Zeilenrang von A''. Dies bedeutet

$$\operatorname{rank} A \;=\; \operatorname{rank} A'' \;=\; \operatorname{rank} A''^{\mathrm{tr}} \;=\; \operatorname{rank}(L''A)^{\mathrm{tr}} \;=\; \operatorname{rank}(A^{\mathrm{tr}}L''^{\mathrm{tr}}) \;=\; \operatorname{rank} A^{\mathrm{tr}}\,.$$

□

3.10. Matrizen aus linearen Abbildungen, Teil II

Sei V ein K-Vektorraum mit $\dim V = n$ und $\mathcal{B} = (b_1,\dots,b_n)$ eine Basis von V. Dann lässt sich $v \in V$ eindeutig schreiben als $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$. Setze

$$[v]_{\mathcal{B}} \;:=\; \begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix} \;\in\; K^n\,.$$

Die Abbildung

$$\kappa_{\mathcal{B}} : V \to K^n : v \mapsto [v]_{\mathcal{B}}$$

ist ein K-Vektorraum-Isomorphismus, weil das Bild von $\mathcal{B}$ wegen $[b_i]_{\mathcal{B}} = \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} = e_i$ eine Basis von K^n ist.

Seien V und W Vektorräume über K und und $f : V \to W$ eine lineare Abbildung. Zu Basen $\mathcal{B} = (b_1,\dots,b_n)$ und $\mathcal{C} = (c_1,\dots,c_m)$ von V bzw. W ist

$$[f]^{\mathcal{C}}_{\mathcal{B}} \;=\; ([f(b_1)]_{\mathcal{C}}, [f(b_2)]_{\mathcal{C}}, \dots, [f(b_n)]_{\mathcal{C}})$$

die Matrix von f bezüglich $\mathcal{B}$ und $\mathcal{C}$.

Proposition 3.10.1. *Für alle $v \in V$ gilt*

$$[f]^{\mathcal{C}}_{\mathcal{B}} \cdot [v]_{\mathcal{B}} = [f(v)]_{\mathcal{C}}\,.$$

Dies ist gleichbedeutend mit $[f]^{\mathcal{C}}_{\mathcal{B}} \cdot \kappa_{\mathcal{B}}(v) = (\kappa_{\mathcal{C}} \circ f)(v)$ bzw. $[f]^{\mathcal{C}}_{\mathcal{B}} = [\kappa_{\mathcal{C}} \circ f \circ \kappa_{\mathcal{B}}^{-1}]$.

Beweis. Da die Abbildungen von V nach K^m

$$v \mapsto [f]^{\mathcal{C}}_{\mathcal{B}} \cdot [v]_{\mathcal{B}} \quad \text{und} \quad v \mapsto [f(v)]_{\mathcal{C}}$$

beide linear sind, genügt es, sie auf einer Basis zu vergleichen, beispielsweise auf der Basis $\mathcal{B}$:

$$[f]^{\mathcal{C}}_{\mathcal{B}} \cdot [b_i]_{\mathcal{B}} \;=\; [f]^{\mathcal{C}}_{\mathcal{B}} \cdot e_i \;=\; i\text{-te Spalte von } [f]^{\mathcal{C}}_{\mathcal{B}} \;=\; [f(b_i)]_{\mathcal{C}}\,.$$

□

Für jede lineare Abbildung $f : V \to W$ existiert zu gegebenen Basen $\mathcal{B}$ bzw. $\mathcal{C}$ eine Matrix $[f]^{\mathcal{C}}_{\mathcal{B}}$. Wir definieren die Abbildung

$$\Psi^{\mathcal{C}}_{\mathcal{B}} \;:\; \operatorname{Hom}_K(V,W) \to K^{m\times n} \;:\; f \mapsto [f]^{\mathcal{C}}_{\mathcal{B}}$$

analog zu Satz 3.8.11.

Satz 3.10.2. *Die Abbildung $\Psi^{\mathcal{C}}_{\mathcal{B}} : \operatorname{Hom}_K(V,W) \to K^{m\times n}$ ist ein K-linearer Isomorphismus.*

Bemerkung 3.10.3. Wir betrachten den Spezialfall $V = K^n$ und $W = K^m$. Sind $\mathcal{B}$ und $\mathcal{C}$ die Standardbasen von K^n bzw. K^m, so gilt $\Psi_{\mathcal{B}}^{\mathcal{C}} = \Phi^{-1}$.

Proposition 3.10.4. *Seien U, V, W Vektorräume über K mit Basen $\mathcal{A}$, $\mathcal{B}$ bzw. $\mathcal{C}$. Für lineare Abbildungen $g : U \to V$ und $f : V \to W$ gilt dann*

$$[f \circ g]_{\mathcal{A}}^{\mathcal{C}} = [f]_{\mathcal{B}}^{\mathcal{C}} \cdot [g]_{\mathcal{A}}^{\mathcal{B}} .$$

Genauer: Das Diagramm

$$\begin{tikzcd}
U \arrow[r, "g"] \arrow[d, "\kappa_{\mathcal{A}}"] & V \arrow[r, "f"] \arrow[d, "\kappa_{\mathcal{B}}"] & W \arrow[d, "\kappa_{\mathcal{C}}"] \\
K^p \arrow[r, "{[g]_{\mathcal{A}}^{\mathcal{B}}}"'] & K^n \arrow[r, "{[f]_{\mathcal{B}}^{\mathcal{C}}}"'] & K^m
\end{tikzcd}$$

ist kommutativ.

Beweis.

$$\begin{aligned}
[f \circ g]_{\mathcal{A}}^{\mathcal{C}} &= [\kappa_{\mathcal{C}} \circ f \circ g \circ \kappa_{\mathcal{A}}^{-1}] = [\kappa_{\mathcal{C}} \circ f \circ \kappa_{\mathcal{B}}^{-1} \circ \kappa_{\mathcal{B}} \circ g \circ \kappa_{\mathcal{A}}^{-1}] \\
&= [\kappa_{\mathcal{C}} \circ f \circ \kappa_{\mathcal{B}}^{-1}] \cdot [\kappa_{\mathcal{B}} \circ g \circ \kappa_{\mathcal{A}}^{-1}] = [f]_{\mathcal{B}}^{\mathcal{C}} \cdot [g]_{\mathcal{A}}^{\mathcal{B}} .
\end{aligned}$$

□

3.11. Basiswechsel

Seien $\mathcal{B} = (b_1, \dots, b_n)$ und $\mathcal{B}' = (b_1', \dots, b_n')$ Basen des K-Vektorraums V. Jedes $v \in V$ lässt sich nun bezüglich $\mathcal{B}$ und bezüglich $\mathcal{B}'$ darstellen:

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \quad \text{und} \quad [v]_{\mathcal{B}'} = \begin{pmatrix} \lambda_1' \\ \vdots \\ \lambda_n' \end{pmatrix} .$$

Insbesondere gilt für die Vektoren $b_1', \dots, b_n'$:

$$[b_i']_{\mathcal{B}} = \begin{pmatrix} s_{1i} \\ \vdots \\ s_{ni} \end{pmatrix} \quad \text{und} \quad [b_i']_{\mathcal{B}'} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = e_i .$$

Es gilt

$$S := \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{ni} \end{pmatrix} = [\mathrm{id}_V]_{\mathcal{B}'}^{\mathcal{B}} .$$

Dies ist die *Transformationsmatrix des Basiswechsels* von $\mathcal{B}'$ nach $\mathcal{B}$. Weil die Spalten von S eine Basis von K^n bilden, ist S invertierbar und es gilt $S^{-1} = [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}$.

Beispiel 3.11.1. Seien $V = \mathbb{R}^2$, $\mathcal{B} = (\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right))$ und $\mathcal{B}' = (\left(\begin{smallmatrix}2\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1\\2\end{smallmatrix}\right))$. Dann ist

$$S = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \quad \text{und} \quad S^{-1} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

Dies ergibt sich aus der folgenden Rechnung:

$$\begin{aligned}
\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} \lambda_1' \\ \lambda_2' \end{pmatrix} &\Leftrightarrow \left\{ \begin{array}{lcl} \lambda_1 &=& 2\lambda_1' - \lambda_2' \\ \lambda_2 &=& \lambda_1' + 2\lambda_2' \end{array} \right. \\
&\Leftrightarrow \begin{pmatrix} \lambda_1' \\ \lambda_2' \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \\
&\Leftrightarrow \left\{ \begin{array}{lcl} \lambda_1' &=& \frac{2}{5}\lambda_1 + \frac{1}{5}\lambda_2 \\ \lambda_2' &=& -\frac{1}{5}\lambda_1 + \frac{2}{5}\lambda_2 \end{array} \right.
\end{aligned}$$

—Bild—

Proposition 3.11.2. *Sei $f : V \to W$ linear. Ferner seien $\mathcal{B}$, $\mathcal{B}'$ Basen von V und $\mathcal{C}$, $\mathcal{C}'$ Basen von W. Setze $S = [\mathrm{id}_V]^{\mathcal{B}}_{\mathcal{B}'}$ und $R = [\mathrm{id}_W]^{\mathcal{C}}_{\mathcal{C}'}$. Dann gilt*

$$\begin{aligned}
[f]^{\mathcal{C}'}_{\mathcal{B}'} &= R^{-1} \cdot [f]^{\mathcal{C}}_{\mathcal{B}} \cdot S \quad \textit{und} \\
[f]^{\mathcal{C}}_{\mathcal{B}} &= R \cdot [f]^{\mathcal{C}'}_{\mathcal{B}'} \cdot S^{-1}.
\end{aligned}$$

Beweis. Folgt aus Proposition 3.10.4. □

3.12. Nochmals lineare Gleichungssysteme

Betrachte das lineare Gleichungssystem

$$\begin{array}{ccccccc}
\alpha_{11}\,x_1 & + \cdots + & \alpha_{1n}\,x_n & = & \beta_1 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
\alpha_{m1}\,x_1 & + \cdots + & \alpha_{mn}\,x_n & = & \beta_n
\end{array} \tag{8}$$

über dem Körper K, bzw. in Matrixschreibweise: $Ax = b$ für $A = (\alpha_{ij})_{i,j} \in K^{m\times n}$, $b = (\beta_i)_i \in K^n$ und einen Spaltenvektor $x = (x_j)_j$ von n Unbestimmten. Das zugehörige homogene System ist

$$Ax = 0. \tag{9}$$

Es sei $(A|b) \in K^{m\times(n+1)}$ die erweiterte Koeffizientenmatrix von (8)

3.12.1. Basislösungen des homogenen Systems. Die Lösungen von (9) bilden einen Unterraum U von K^n. Dabei ist $\dim_K U = n - \mathrm{rank}_K A =: k$. Eine Basis von $(u_1, \ldots, u_k)$ von U heißt *System von Basislösungen* von (9). Jede Lösung von (9) ist Linearkombination der Basislösungen.

3.12.2. Existenz von Lösungen. Das homogene System (9) hat stets die triviale Lösung $x = 0$.

Proposition 3.12.1. *Äquivalent sind:*

(i) Das inhomogene System $Ax = b$ hat mindestens eine Lösung.
(ii) Die rechte Seite b liegt im Spaltenraum von $A = \mathrm{img}\,\varphi_A$.
(iii) $\mathrm{rank}\, A = \mathrm{rank}(A|b)$.

Proposition 3.12.2. *Äquivalent sind:*

(i) Das lineare Gleichungssystem $Ax = b$ besitzt für jedes $b \in K^m$ eine Lösung.
(ii) Die lineare Abbildung φ_A ist surjektiv.

(iii) $\operatorname{rank} A = m$.

3.12.3. Die Menge aller Lösungen des inhomogenen Systems. Angenommen, das inhomogene System (8) hat mindestens eine Lösung $x_0 \in K^n$. Dann ist

$$x_0 + U \;=\; \{x_0 + \lambda_1 u_1 + \cdots + \lambda_k u_k : \; \lambda_i \in K\}$$

die Menge aller Lösungen von (8)

3.12.4. Lineare Gleichungssysteme mit quadratischer Koeffizientenmatrix. Sei nun $m = n$, das heißt die Koeffizientenmatrix $A \in K^{n\times n}$ ist quadratisch.

Proposition 3.12.3. *Äquivalent sind:*

(i) Das lineare Gleichungssystem $Ax = b$ hat eine eindeutige Lösung.
(ii) Das lineare Gleichungssystem $Ax = b$ besitzt für jedes $b \in K^m$ eine Lösung.
(iii) $\operatorname{rank} A = n$.
(iv) Die Matrix A ist invertierbar.

3.13. Matrixgleichungen und Matrixinversion

Wir betrachten Matrizen $A \in K^{m\times n}$ und $B \in K^{m\times p}$. Wenn wir die k-te Spalte von B mit b_k bezeichnen, können wir insgesamt p lineare Gleichungssysteme

$$Ax \;=\; b_k \quad \text{für } 1 \leq k \leq p \tag{10}$$

betrachten. Sind alle diese Gleichungssysteme lösbar, können wir jeweils Lösungen auswählen und zu einer Matrix $X \in K^{n\times p}$ zusammenstellen. Auf diese Weise erhalten wir eine Lösung der *Matrixgleichung*

$$AX \;=\; B\,. \tag{11}$$

Umgekehrt liefert jede Lösung der Matrixgleichung (11) simultan je eine Lösung für die p Gleichungen (10).

Speziell für $m = n = p$ und $B = E_n$ erhalten wir für $A \in \mathrm{GL}_n K$ invertierbar die Inverse als Lösung der Matrixgleichung

$$AX \;=\; E_n\,.$$

Umgekehrt gilt: Falls eine solche Lösung nicht existiert, ist die Matrix A nicht invertierbar.

KAPITEL 4

Affine Geometrie

4.1. Quotientenräume

Sei V ein K-Vektorraum und U ein Unterraum.

Definition 4.1.1. Sei $x \in V$. Dann heißt die Menge

$$x + U \;:=\; \{x + u : \; u \in U\}$$

Nebenklasse von U in V.

Proposition 4.1.2. *Sei $U \leq V$. Dann ist*

$$V/U \;:=\; \{x + U : \; x \in V\}$$

eine Partition von V.

Beweis. Offensichtlich überdeckt V/U die Menge V. Sei $z \in (x + U) \cap (y + U)$, das heißt, es existieren $u, u' \in U$, so dass $x + u = z = y + u'$. Hieraus folgt $x - y = u' - u \in U$. Sei $u'' \in U$ beliebig. Dann gilt $x+u'' = y+u'-u+u'' \in y+U$, insgesamt also $x+U \subseteq y+U$. Aus Symmetrie folgt $x + U = y + U$, und damit gilt die Behauptung. $\square$

Auf der Menge V/U lassen sich wiederum eine Addition und eine Skalarmultiplikation definieren:

$$(x + U) + (y + U) \;:=\; (x + y) + U \quad \text{und} \quad \lambda \cdot (x + U) \;:= (\lambda x) + U\,.$$

Man rechnet nach, dass diese Operationen wohldefiniert sind, also nicht von der Wahl der jeweiligen Repräsentanten der Nebenklassen abhängen.

Proposition 4.1.3. *Das Tripel $(V/U, +, \cdot)$ ist ein K-Vektorraum.*

Definition 4.1.4. Der K-Vektorraum $(V/U, +, \cdot)$ heißt *Quotientenraum* von V nach U.

Proposition 4.1.5. *Ist V endlichdimensional, so gilt* $\dim U + \dim(V/U) = \dim V$.

Beweis. Sei $(u_1, u_2, \ldots, u_k)$ eine Basis von U. Dann existieren Vektoren $v_1, v_2, \ldots, v_m \in V$, so dass $(u_1, \ldots, u_k, v_1, \ldots, v_k)$ eine Basis von V ist. Wir definieren die lineare Abbildung

$$\pi_U \;:\; V \to V \;:\; \lambda_1 u_1 + \cdots + \lambda_k u_k + \mu_1 v_1 + \cdots + \mu_m v_m \mapsto \lambda_1 u_1 + \cdots + \lambda_k u_k\,. \tag{12}$$

Offensichtlich gilt $\operatorname{img} \pi_U = U$, und die Behauptung folgt aus der Dimensionsformel. $\square$

Die Abbildung π_U aus (12) heißt *kanonische Projektion* auf U.

Proposition 4.1.6. *Sei $\varphi \in \operatorname{End}_K(V)$ mit $\varphi(U) \subseteq U$. Dann definiert*

$$\varphi \;:\; V/U \to V/U \;:\; x + U \mapsto \varphi(x) + U$$

eine lineare Abbildung.

Beweis. Es ist wiederum die Wohldefiniertheit zu zeigen. Seien $x, y \in V$ mit $x + U = y + U$, also $x - y \in U$. Dann folgt $\varphi(x) - \varphi(y) = \varphi(x - y) \in U$. $\square$

Die in Proposition 4.1.6 definierte Abbildung $\varphi : V/U \to V/U$ heißt die von φ auf dem Quotienten V/U *induzierte* lineare Abbildung.

4.2. Affine Räume

Sei wieder V ein K-Vektorraum.

Definition 4.2.1. Ein *affiner Unterraum* von V ist eine Nebenklasse $x + U$ eines beliebigen Unterraums U in V. Die Menge aller affinen Unterräume von V wird mit $\mathrm{AG}(V)$ bezeichnet. Wir setzen $\dim_K(x + U) := \dim_K(U)$.

Definition 4.2.2. Für $v_1, v_2, \ldots, v_m \in V$ und $\lambda_1, \lambda_2, \ldots, \lambda_m \in K$ heißt die Linearkombination $\lambda_1 v_1 \lambda_2 v_2 + \cdots + \lambda_m vm$ eine *Affinkombination*, falls $\sum \lambda_i = 1$. Die Menge aller Affinkombinationen aus $M \subset V$ heißt *affine Hülle* von M und wird mit $\operatorname{aff} M$ bezeichnet.

Proposition 4.2.3. *Für $M \subset V$ ist $\operatorname{aff} M$ der kleinste affine Unterraum, der M enthält.*

Beweis. Sei $x \in M$ und $m_1, m_2, \ldots, m_k \in M$ sowie $\lambda_1, \lambda_2, \ldots, \lambda_k \in K$. Dann ist

$$x + \sum_{i=1}^{k} \lambda_i (m_i - x) \;=\; \sum_{i=1}^{k} \lambda_i m_i + (1 - \sum_{i=1}^{k} \lambda_i) x \,.$$

Offensichtlich gilt $\sum_{i=1}^{k} \lambda_i + (1 - \sum_{i=1}^{k} \lambda_i) = 1$, und deshalb ist

$$\operatorname{aff} M \;=\; x + \operatorname{lin}\{y - x : \, y \in M\} \,.$$

Die Behauptung folgt damit aus Proposition 2.3.6. □

Eine endliche Menge $M \subset V$ heißt *affin unabhhängig*, falls $\dim \operatorname{aff} M = \# M - 1$.

Bemerkung 4.2.4. Sei $M \subset V$ endlich. Dann bilden die affin unabhängigen Teilmengen von M ein Matroid auf der Grundmenge M.

Im Folgenden betrachten wir stets die Situation $V = K^n$; insbesondere ist V endlichdimensional. Statt $\mathrm{AG}(K^n)$ schreiben wir auch $\mathrm{AG}_n K$.

Die affinen Unterräume von K^n sind genau die Lösungsmengen von Systemen endlich vieler linearer Gleichungen in n Unbestimmten über K. Die linearen Unterräume sind genau die Lösungen der homogenenen linearen Gleichungssysteme.

Definition 4.2.5. Die affinen Unterräume der Dimension 0, 1, 2, $n - 1$ heißen *(affine) Punkte, Geraden, Ebenen* bzw. *Hyperebenen*.

Proposition 4.2.6. *Für jede affine Hyperebene H existiert eine lineare Abbildung $\varphi_H : K^n \to K$ vom Rang 1 und ein $\alpha \in K$, so dass*

$$H \;=\; \{x \in K^n : \, \varphi_H(x) = \alpha\} \,.$$

Umgekehrt ist jede solche Menge eine affine Hyperebene.

Beweis. Sei $H = x + U$ mit $x \in K^n$ und $U \leq K^n$, wobei $\dim U = n - 1$. Dann existiert eine lineare Abbildung $\varphi_H : K^n \to K$ mit der Eigenschaft, dass $U = \ker \varphi_H$ ist. Aus der Linearität von φ_H folgt die Behauptung. □

Definition 4.2.7. Zwei affine Unterräume $x + U$ und $y + W$ mit $1 \leq \dim U \leq \dim W$ heißen *parallel*, falls $U \leq W$ ist. Zwei nicht-parallele disjunkte affine Unterräume heißen *windschief*.

4.3. Exkurs: Affine Ebenen

DEFINITION 4.3.1. Es sei P eine Menge und $\mathcal{L} \subset \mathcal{P}(P)$. Das Paar $(P, \mathcal{L})$ heißt *abstrakte affine Ebene* mit *Punktmenge* P und *Geradenmenge* $\mathcal{L}$, falls gilt:

(i) Es gibt ein Viereck, das heißt es existieren vier Punkte, von denen keine drei auf einer Geraden liegen.
(ii) Je zwei Punkte $p, q \in P$ liegen auf genau einer gemeinsamen Geraden $p \vee q$.
(iii) Für jeden Punkt $p \in P$ und jede Gerade $L \in \mathcal{L}$ existiert genau eine Gerade $L' \in \mathcal{L}$, die durch p geht und *parallel* zu L ist.

Ist K ein Körper so ist $\mathrm{AG}_2\, K$ eine abstrakte affine Ebene. Die Umkehrung gilt aber nicht. Es gibt zum Beispiel eine *nicht-desarguessche* affine Ebenen mit 64 Punkten und 72 Geraden.[1]

PROPOSITION 4.3.2. *Sei $\mathbb{A} = (P, \mathcal{L})$ eine endliche affine Ebene. Dann existiert eine Zahl $n \in \mathbb{N}$, so dass jede Gerade genau n Punkte enthält, und jeder Punkt liegt auf genau $n+1$ Geraden. Außerdem gilt $\#P = n^2$ und $\#cL = n^2 + n$.*

BEWEIS. Sei (p_1, p_2, p_3, p_4) ein Viereck mit $L_1 = p_1 \vee p_2$, $L_2 = p_2 \vee p_2$, $L_3 = p_3 \vee p_4$ und $L_4 = p_4 \vee p_1$. Ferner sei n die Anzahl der Punkte auf L_1. Es ist $p_3 \notin L_1$, und die Abbildung

$$[L_1, p_3] : L \to \mathcal{L}_{p_3} : p \mapsto p \vee p_3$$

ist wegen der Eindeutigkeit der Verbindungsgerade injektiv. Zusammen mit dem Parallelenaxiom folgt $\#\mathcal{L}_{p_3} = n + 1$. Da $p_3 \notin L_4$ folgt analog $\#L_4 = \#\mathcal{L}_{p_3} = n$ etc. □

Die Zahl n in Proposition 4.3.2 heißt *Ordnung* von $\mathbb{A}$.

VERMUTUNG 4.3.3. Die Ordnung einer endlichen abstrakten affinene Ebene ist eine Primzahlpotenz.

SATZ 4.3.4 (Bruck-Ryser 1949). *Wenn n der Form $4k+1$ oder $4k+2$ ist für $k \in \mathbb{N}$ und n ist nicht Summe zweier ganzzahliger Quadrate, dann tritt n nicht auf als Ordnung einer endlichen affinen Ebene.*

SATZ 4.3.5 (Lam 1991). *Es gibt keine affine Ebene der Ordnung 10.*

4.4. Affine Abbildungen

DEFINITION 4.4.1. Eine Abbildung $\varphi : K^n \to K^m$ der Form

$$\varphi(x) \;=\; v + \psi(x)$$

mit $v \in K^m$ und $\psi \in \mathrm{Hom}(K^n, K^m)$ heißt *affin*. Der *Rang* von φ ist definiert als der Rang von ψ. Gilt $m = n$ und $\operatorname{rank} \varphi = n$, so heißt φ affine Transformation.

Das Bild eines affinen Unterraums unter einer affinen Abbildung ist wiederum ein affiner Unterraum.

BEISPIEL 4.4.2.

- Die Abbildung $x \mapsto v + x$ ist eine *Translation*.
- Invertierbare lineare Abbildungen (auf K^n) heißen auch *lineare Transformationen*.

[1] D.R. Hughes und F.C. Piper: *Projective planes*, Graduate Texts in Mathematics, Vol. 6. Springer-Verlag, New York-Berlin, 1973.

Die affinen Transformationen von K^n bilden eine Gruppe $\mathrm{AGL}_n\, K$, die von den linearen Transformationen und den Translationen erzeugt wird.

4.5. Exkurs: Polytope

In diesem Abschnitt betrachten wir ausschließlich den Körper der reellen Zahlen. Es bezeichnet $\mathbb{R}_{\geq 0}$ die Menge der nicht-negativen reellen Zahlen.

Definition 4.5.1. Sei $M \subset \mathbb{R}^n$. Dann heißt die Menge

$$\operatorname{pos} M \ := \ \{\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k : \ \lambda_i \in \mathbb{R}_{\geq 0},\, x_i \in M\}$$

die *positive Hülle* von M und die Menge

$$\operatorname{conv} M \ := \ \operatorname{aff} M \cap \operatorname{pos} M$$

die konvexe Hülle von M.

Definition 4.5.2. Ein *(konvexes) Polytop* ist die konvexe Hülle endlich vieler Punkte in $\mathbb{R}^n$.

Beispiel 4.5.3. Wir setzen $[n] := \{1, 2, \ldots, n\}$ und $e_I := \sum_{i \in I} e_i$ für $I \subseteq [n]$. Das Polytop

$$\operatorname{conv}\{e_I : \ I \subseteq [n]\} \ = \ [0,1]^n$$

heißt n-dimensionaler *Einheitswürfel*.

Definition 4.5.4. Sei $M = ([n], \mathcal{B})$ ein Matroid auf der Menge $[n]$. Dann heißt

$$\operatorname{conv}\{e_B : \ B \in \mathcal{B}\}$$

das *Matroidpolytop* zu M.

Lemma 4.5.5. *Das Bild eines Polytops unter einer affinen Abbildung ist wieder ein Polytop.*

Definition 4.5.6. Ein *Zonotop* ist ein affines Bild eines Einheitswürfels.

KAPITEL 5

Determinanten

5.1. Vorüberlegungen

Ziel: Einführung einer Abbildung

$$\det : K^{n\times n} \to K$$

mit den folgenden Eigenschaften:

(i) $\det(E_n) = 1$
(ii) $\det(AB) = (\det A)(\det B)$
(iii) $\det(A^{\mathrm{tr}}) = \det A$
(iv) Falls $A \in \mathrm{GL}_n\, K$, so ist $\det A \neq 0$.
(v) Speziell für $K = \mathbb{R}$, soll $\det A$ das Volumen des *Parallelepipeds* der Spalten (oder Zeilen) von A ergeben.

In Abschnitt 2.4 wurden sämtliche Unterräume von $\mathbb{R}^2$ bestimmt. Dabei kam auch heraus, dass die Vektoren

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

in $\mathbb{R}^2$ genau dann linear unabhängig sind, wenn gilt

$$0 \neq u_1v_2 - u_2v_1 \overset{\text{zu zeigen}}{=} \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}.$$

Sei K im folgenden stets ein beliebiger Körper. Wir identifizieren ab jetzt oft $K^{n\times n} =_{\uparrow} \underbrace{K^n \times \cdots \times K^n}_{n\text{ -mal}}$. Das heißt wir fassen eine Matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in K^{n\times n}$$

auf als das n-Tupel ihrer Spalten

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$$

5.2. Multilinearformen

Definition 5.2.1. Eine Abbildung $F : \underbrace{K^n \times \cdots \times K^n}_{n\text{-mal}} \to K$ heißt:

(i) *Multilinearform* auf K^n (oder n-*Form*), falls für alle $i \in \{1, 2, \dots, n\}$, für alle $v_1, \dots, v_n \in K^n$, für alle $\lambda, \mu \in K$ und für alle $x, y \in K^n$ gilt:

$$F(v_1, \dots, v_{i-1}, \lambda x + \mu y, v_{i+1}, \dots, v_n) = \lambda F(v_1, \dots, x, \dots, v_n) + \mu F(v_1, \dots, y, \dots, v_n).$$

Das heißt, die Abbildung F ist linear in jeder ihrer n Argumente.

(ii) *alternierende* Multilinearform auf K^n, falls (i) und zusätzlich gilt:

$$F(v_1, \dots, v_n) = 0\,, \quad \text{falls } v_i = v_j \text{ für } i \neq j\,.$$

(iii) *normierte* alternierende Multilinearform (oder *Determinantenform*), falls außer (i) und (ii) zusätzlich:

$$F(E_n) \;=\; F(e_1, \dots, e_n) \;=\; 1\,.$$

Bemerkung 5.2.2. In der Literatur findet man statt der Eigenschaft (ii) oft die Eigenschaft

$$F(\dots, \underset{\substack{\uparrow\\ i}}{x}, \dots, \underset{\substack{\uparrow\\ j}}{y}, \dots) \;=\; -F(\dots, \underset{\substack{\uparrow\\ i}}{y}, \dots, \underset{\substack{\uparrow\\ j}}{x}, \dots) \tag{13}$$

▷ Es gilt „(ii) $\Rightarrow$ (13)“. Denn: Ohne Einschränkung sei $n = 2$. Dann gilt:

$$\begin{aligned} F(x,y) + F(y,x) &\overset{\text{(ii)}}{=} F(x,y) + F(y,x) + F(x-y, x-y) \\ &\overset{\text{(i)}}{=} F(x,y) + F(y,x) + F(x, x-y) - F(y, x-y) \\ &\overset{\text{(i)}}{=} F(x,x) + F(y,y) \\ &\overset{\text{(ii)}}{=} 0\,. \end{aligned}$$

▷ Die Umkehrung gilt nur für den Fall, dass in K die Ungleichung $2 := 1 + 1 \neq 0$ gilt (das heißt die *Charakteristik* von K ist ungleich zwei):

$$2F(x,x) \;=\; F(x,x) + F(x,x) \overset{(13)}{=} F(x,x) - F(x,x) \;=\; 0\,.$$

Beispiel 5.2.3. Für $K = \mathbb{F}_2$ existiert auf $\mathbb{F}_2^2$ eine nicht-alternierende 2-Form, die (13) aber nicht (ii) erfüllt:

$$F\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \;:=\; x_1 y_1 + x_2 y_2\,.$$

Wir wollen Beispiele über einem beliebigen Körper K aber in niedrigen Dimensionen studieren.

Beispiel 5.2.4. Die Abbildung $\mathrm{id}_K : K \to K$ ist eine Determinantenform auf $K = K^1$. Denn:

▷ für $n = 1$: multilinear = linear
▷ für $n = 1$: alternierend keine zusätzliche Bedingung
▷ $\mathrm{id}_K(1) = 1$, also ist die alternierende Multilinearform normiert.

Beispiel 5.2.5. Die Abbildung

$$F : K^2 \times K^2 : \left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) \mapsto u_1 v_2 - u_2 v_1$$

ist eine Determinantenform auf K^2. Denn:

$$\begin{aligned} F\left(\lambda \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \mu \begin{pmatrix} u_1' \\ u_2' \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) &= (\lambda u_1 + \mu u_1')v_2 - (\lambda u_2 + \mu u_2')v_1 \\ &= \lambda(u_1 v_2 - u_2 v_1) + \mu(u_1' v_2 - u_2' v_1) \\ &= \lambda F\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) + \mu\left(\begin{pmatrix} u_1' \\ u_2' \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) \end{aligned}$$

Damit ist F eine Multilinearform. Außerdem gilt

$$F\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = u_1u_2 - u_2u_1 = 0$$

sowie

$$F\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 1\,.$$

LEMMA 5.2.6. *Sei $F : K^n \times \cdots \times K^n \to K$ eine alternierende Multilinearform und $v_1, \ldots, v_n \in K^n$. Dann gilt $F(v_1, \ldots, v_n) = 0$, falls $(v_1, \ldots, v_n)$ linear abhängig.*

BEWEIS. Ohne Einschränkung sei $v_1 = \lambda_2 v_2 + \cdots + \lambda_n v_n$. Dann gilt

$$\begin{aligned} F(v_1, \ldots, v_n) &= F(\lambda_2 v_2 + \cdots + \lambda_n v_n, v_2, \ldots, v_n) \\ &= \sum_{k=2}^{n} \lambda_i F(v_i, v_2, \ldots, v_n) = 0\,. \end{aligned}$$

□

SATZ 5.2.7. *Sei $F : K^n \times \cdots \times K^n \to K$ eine alternierende Multilinearform mit der Eigenschaft $F(e_1, \ldots, e_n) = 0$. Dann ist F die Nullabbildung.*

BEWEIS. Wir führen den Beweis per Induktion nach n. Für den Induktionsanfang $n = 1$ gilt:

$$F(e_1) = 0 \Rightarrow \forall \lambda \in K : F(\lambda e_1) = \lambda F(e_1) = 0\,.$$

Sei $n \geq 2$ und die Behauptung bereits beweisen für alle alternierende $(n-1)$-Formen. Wir zeigen zunächst

$$\forall \sigma \in \mathrm{Sym}(\{1, \ldots, n\}) : F(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = 0\,. \tag{14}$$

Es sei σ eine beliebige Permutation. Dann gilt

$$F(e_{\sigma(1)}, \ldots, e_1, \ldots, e_{\sigma(n)}) = -F(e_1, \ldots, e_{\sigma(1)}, \ldots, e_{\sigma(n)})\,.$$

Die Abbildung

$$F_1 : \underbrace{K^{n-1} \times \cdots \times K^{n-1}}_{n-1} \to K : (y_2, \ldots, y_n) \mapsto F(e_1, \begin{pmatrix} 0 \\ y_2 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ y_n \end{pmatrix})$$

ist eine alternierende $(n-1)$-Form. Nach Induktionsvoraussetzung gilt also $F_1 \equiv 0$. Wir erhalten (für $\sigma(1) \neq 1$)

$$\begin{aligned} F(e_{\sigma(1)}, \ldots, e_1, \ldots, e_{\sigma(n)}) &= -F(e_1, \ldots, e_{\sigma(1)}, \ldots, e_{\sigma(n)} \\ &= -F_1(e'_{\sigma(2)}, \ldots, e'_{\sigma(1)}, \ldots, e'_{\sigma(n)}) = 0\,. \end{aligned}$$

Hier bezeichnet ausnahmsweise e'_k den $(k-1)$-ten Standardbasisvektor in K^{n-1}, um eine Verwechslung mit den Standardbasisvektoren von K^n zu vermeiden. Damit gilt (14).

Im nächsten Schritt zeigen wir

$$\forall x_2, \ldots, x_n \in K^n : F(e_1, x_1, \ldots, x_n) = 0 \tag{15}$$

Wir schreiben x_i als $\lambda_i e_1 + \Lambda_i$ mit $\lambda_i \in K$ und $\Lambda_i \in \mathrm{lin}(e_2, \ldots, e_n)$. Dann ergibt sich

$$\begin{aligned} F(e_1, x_2, \ldots, x_n) &= F(e_1, \lambda_2 e_1 + \Lambda_2, \ldots, \lambda_n e_1 + \Lambda_n) \\ &= F(e_1, \lambda_2 e_1, \lambda_3 e_1 + \Lambda_3, \ldots, \lambda_n e_1 + \Lambda_n) \\ &\qquad + F(e_1, \Lambda_2, \lambda_3 e_1 + \Lambda_3, \ldots, \lambda_n e_1 + \Lambda_n) \\ &= F(e_1, \Lambda_2, \lambda_3 e_1 + \Lambda_3, \ldots, \lambda_n e_1 + \Lambda_n) \\ &= \ldots = F(e_1, \Lambda_2, \ldots, \Lambda_n) = F_1(\Lambda_2, \ldots, \Lambda_n) = 0\,. \end{aligned}$$

Schließlich seien $x_1, \ldots, x_n \in K^n$ beliebig. Wiederum schreiben wir $x_i = \lambda_i e_1 + \Lambda_i$ mit $\lambda_i \in K$ und $\Lambda_i \in \mathrm{lin}(e_2, \ldots, e_n)$.

$$\begin{aligned} F(x_1, \ldots, x_n) &= F(\lambda_1 e_1 + \Lambda_1, x_2, \ldots, x_n) \\ &= F(\lambda_1 e_1, x_2, \ldots, x_n) + F(\Lambda_1, x_2, \ldots, x_n) \\ &= F(\Lambda_1, x_2, \ldots, x_n)\,, \end{aligned}$$

weil $(\Lambda_1, x_2, \ldots, x_n)$ linear abhängig ist. Damit folgt weiter

$$F(x_1, \ldots, x_n) \;=\; F(\Lambda_1, x_2, \ldots, x_n) \;=\; \ldots \;=\; F(\Lambda_1, \Lambda_2, \ldots, \Lambda_n) \;=\; 0\,,$$

weil außerdem auch $(\Lambda_1, \Lambda_2, \ldots, \Lambda_n)$ linear abhängig ist. □

KOROLLAR 5.2.8. *Seien $F, G : K^n \times \cdots \times K^n \to K$ alternierende Linearformen mit $F(e_1, \ldots, e_n) = G(e_1, \ldots, e_n)$. Dann gilt $F = G$. Insbesondere gibt es höchstens eine Determinantenform auf K^n.*

BEWEIS. Übungsaufgabe. □

5.3. Konstruktion der Determinantenform auf K^n

In den Beispielen 5.2.4 und 5.2.5 hatten wir bereits Determinantenformen auf $K = K^1$ und K^2 konstruiert:

$$D_1 : K \to K : x \mapsto x$$

$$D_2 : K^2 \times K^2 : (\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}) \mapsto u_1\, v_2 - u_2\, v_1.$$

Für $n \geq 3$ gehen wir induktiv vor, d.h. wir gehen davon aus, dass wir eine Determinantenform

$$D_{n-1} : \underbrace{K^{n-1} \times \cdots \times K^{n-1}}_{n-1} \to K$$

bereits konstruiert haben. Wegen Korollar 5.2.8 ist D_{n-1} dann auch eindeutig bestimmt.

Wir setzen

$$D_n(A) \;:=\; \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} D_{n-1}(A_{ij}) \tag{16}$$

für $A = (\alpha_{ij}) \in \underbrace{K^n \times \cdots \times K^n}_{n} =_{\uparrow} K^{n \times n}$ und $i \in \{1, \ldots, n\}$. Dabei ist

$$A_{ij} \;=\; \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{ij} & \ldots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{1i} & \ldots & \alpha_{ij} & \ldots & \alpha_{in} \\ \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \ldots & \alpha_{nj} & \ldots & \alpha_{nn} \end{pmatrix}_{i,j} \;=\; \left(\begin{array}{cc|cc} \alpha_{11} & \ldots & \ldots & \alpha_{1n} \\ \vdots & & & \vdots \\ \hline \vdots & & & \vdots \\ \alpha_{n1} & \ldots & \ldots & \alpha_{nn} \end{array}\right) \;\in\; K^{(n-1)\times(n-1)}$$

durch Streichung der i-ten Zeile und der j-ten Spalte aus A hervorgegangen. Die Formel (16) heißt *Laplace-Entwicklung nach der i-ten Zeile*.

Statt $D_n(A) = D_n((\alpha_{ij}))$ schreiben wir auch $|\alpha_{ij}|$ bzw. $D_n(a_1, \ldots, a_n)$ für $A = (a_1, \ldots, a_n)$.

Beispiel 5.3.1. Über dem Körper $K = \mathbb{F}_3$ entwicklen wir eine 3×3-Matrix nach der ersten Zeile:

$$\begin{aligned}\begin{vmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 2 \end{vmatrix} &= 0 \cdot \begin{vmatrix} 0 & 0 \\ 0 & 2 \end{vmatrix} - 1 \begin{vmatrix} 1 & 0 \\ 2 & 2 \end{vmatrix} + 2 \begin{vmatrix} 1 & 0 \\ 2 & 0 \end{vmatrix} \\ &= 2(1 \cdot 2 - 2 \cdot 0) + 2(1 \cdot 0 - 2 \cdot 0) \\ &= 2 \cdot 2 \;=\; 1\,. \end{aligned}$$

Nun entwickeln wir dieselbe Matrix noch einmal nach der zweiten Zeile:

$$\begin{aligned}\begin{vmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 2 \end{vmatrix} &= -1 \cdot \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} + 0 \begin{vmatrix} 0 & 2 \\ 2 & 2 \end{vmatrix} - 0 \begin{vmatrix} 0 & 1 \\ 2 & 0 \end{vmatrix} \\ &= -(1 \cdot 2 - 2 \cdot 0) \;=\; -2 \;=\; 1\,. \end{aligned}$$

In beiden Fällen kommt dasselbe heraus.

Proposition 5.3.2. *Die Abbildung $D_n : K^{n\times n} \to K$ ist eine Determinantenform.*

Beweis. Die Abbildung $a_j \mapsto \alpha_{ij}\, D_{n-1}(A_{ij}) : K^n \to K$ ist linear, weil die j-te Spalte in der Streichmatrix A_{ij} gar nicht auftaucht. Hieraus folgt, dass

$$D_n(A) \;=\; \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} D_{n-1}(A_{ij})$$

multilinear ist.

Sei $A = (\alpha_{ij}) \in K^{n\times n}$ mit benachbarten Spalten $a_k = a_{k+1}$, die gleich sind. Es gilt

$$\begin{aligned} D_n(a_1, \dots, a_k, a_{k+1}, \dots, a_n) &= \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} D_{n-1}(A_{ij}) \\ &= (-1)^{i+k} \alpha_{ik} D_{n-1}(A_{ik}) + (-1)^{i+k+1} \alpha_{i,k+1} D_{n-1}(A_{i,k+1})\,, \end{aligned}$$

denn in allen anderen Fällen hat A_{ij} zwei gleiche Spalten. daher ist D_n alternierend. Die Normiertheit ist klar. □

Wegen Korollar 5.2.8 ist D_n die einzige alternierende normierte n-Form über dem Körper K.

Definition 5.3.3. Die Abbildung $D_n : K^{n\times n} \to K$ heißt die *Determinante* auf K^n. Wir schreiben $\det := D_n$.

5.4. Eigenschaften der Determinante

Proposition 5.4.1. *Für $A \in K^{n\times n}$ gilt* $\det(A^{\mathrm{tr}}) = \det(A)$.

Beweis. Betrachte die Abbildungen

$$D^n : \underbrace{K^n \times \cdots \times K^n}_{n} =_{\uparrow} K^{n\times n} \to K$$

(Familie von Zeilenvektoren)

mit

$$D^n(A) \;:=\; \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} D^{n-1}(A_{ij}) \tag{17}$$

mit $A = (\alpha_{ij})$ und $D^1(\alpha) = \alpha$. Dies ist die *Laplace-Entwicklung nach der j-ten Spalte*. Wie in Proposition 5.3.2 zeigt man, dass D^n eine normierte alternierende Multilinearform auf K^n ist. Aus Korollar 5.2.8 folgt dann $\det = D^n$, also $\det(A^{\mathrm{tr}}) = D^n(A) = \det(A)$. □

PROPOSITION 5.4.2. *Seien $A, B \in K^{n\times n}$. Dann gilt*

(i) $\det(AB) = \det(A)\det(B)$.
(ii) Falls A invertierbar ist, gilt $\det(A) \neq 0$ *und* $\det(A^{-1}) = [\det(A)]^{-1}$.

BEWEIS. Seien $b_1, \dots, b_n$ die Spalten von B. Dann sind $Ab_1, \dots, Ab_n$ die Spalten von AB. Betrachte die Abbildung $F_A : K^{n\times n} \to K$ mit $F_A(B) = F_A(b_1, \dots, b_n) := \det(Ab_1, \dots, Ab_n) = \det(AB)$. Die Abbildung F_A ist multilinear und alternierend. Seien $a_1, \dots, a_n$ die Spalten von A. Dann ist

$$F_A(E_n) \;=\; F_A(e_1, \dots, e_n) \;=\; \det(a_1, \dots, a_n) \;=\; \det(A).$$

Analog ist durch $G_A(b_1, \dots, b_n) := \det(A)\det(b_1, \dots, b_n)$ eine weitere alternierende Multilinearform auf K^n gegeben, die $G_A(e_1, \dots, e_n) = \det(A)$ erfüllt. Aus Korollar 5.2.8 folgt $F_A = G_A$; insbesondere ergibt sich

$$\det(AB) \;=\; F_A(B) \;=\; G_A(B) \;=\; \det(A)\det(B)\,.$$

Die zweite Behauptung folgt aus der ersten: Für $a \in \mathrm{GL}_n K$ gilt nämlich $\det(A)\det(A^{-1}) = \det(E_n) = 1$. □

SATZ 5.4.3. *Sei $A \in K^{n\times n}$. Äquivalent sind:*

(i) $A \in \mathrm{GL}_n(K)$, das heißt A ist invertierbar.
(ii) $\det(A) \neq 0$.
(iii) Für alle $b \in K^n$ ist das lineare Gleichungssystem $Ax = b$ eindeutig lösbar.
(iv) Die Spalten von A sind linear unabhängig.
(v) Die Zeilen von A sind linear unabhängig.
(vi) $\operatorname{rank} A = n$.

BEWEIS. Falls $A \in \mathrm{GL}_n(K)$, so ist nach Proposition 5.4.2 $\det(A) \neq 0$. Ist dagegen $A \notin GL_n(K)$, so sind die Spalten von A linear abhängig nach Proposition 3.12.2, und wegen Lemma 5.2.6 ist dann $\det(A) = 0$. Die übrigen Aussagen wiederholen Proposition 3.12.3. □

SATZ 5.4.4 (Leibnizformel). *Sei $A = (\alpha_{ij}) \in K^{n\times n}$. Es gilt:*

$$\det(A) \;=\; \sum_{\sigma\in\mathrm{Sym}\{1,\dots,n\}} \mathrm{sgn}(\sigma)\alpha_{1,\sigma(1)}\alpha_{2,\sigma(2)}\cdots\alpha_{d,\sigma(n)}$$

Hierbei bezeichnet $\mathrm{sgn}(\sigma) \in \{\pm 1\}$ das *Signum* der Permutation σ. Dies ist 1, falls σ *gerade* ist, also als Produkt einer geraden Anzahl von *Transpositionen* geschrieben werden kann. Falls σ *ungerade* ist, ist es -1.

BEWEIS. Übungsaufgabe. □

5.5. Ähnlichkeit von Matrizen

DEFINITION 5.5.1. Zwei quadratische Matrizen $A, B \in K^{n\times n}$ heißen *ähnlich*, falls existiert $S \in \mathrm{GL}_n(K)$ mit $B = S^{-1}AS$.

Ähnlichkeit ist eine Äquivalenzrelation auf der Menge $K^{n\times n}$. Die folgende Aussage ist eine Konsequenz aus Proposition 5.4.2.

PROPOSITION 5.5.2. *Ähnliche Matrizen haben dieselbe Determinante.*

Diese Beobachtung erlaubt es, einem beliebigen linearen Endomorphismus

$$\varphi : K^n \to K^n$$

seine *Determinante* zuzuordnen: Seien $\mathcal{B}$, $\mathcal{B}'$ beliebige Basen von K^n. Dann existiert $S \in \mathrm{GL}_n(K)$ mit

$$[\varphi]_{\mathcal{B}'} \;=\; S^{-1}[\varphi]_{\mathcal{B}}S\,.$$

Die Matrix S ist die Matrix des Basiswechsels von $\mathcal{B}$ nach $\mathcal{B}'$ aus Abschnitt 3.11. Es folgt, dass $\det(\varphi) := \det[\varphi]_B$ wohldefiniert ist.

5.6. Determinanten und der Gauß-Jordansche Eliminationsalgorithmus

PROPOSITION 5.6.1. *Sei* $A = (\alpha_{ij}) \in K^{n\times n}$ *eine* obere Dreiecksmatrix*, das heißt* $\alpha_{ij} = 0$ *für alle* $i > j$*. Dann gilt* $\det(A) = \alpha_{11}\alpha_{22}\cdots\alpha_{nn}$*.*

BEWEIS. Beweis per Induktion nach n. Der Induktionsanfang $n = 1$ klar. Sei also $n \geq 2$ und

$$A \;=\; \begin{pmatrix} \alpha_{11} & & * \\ & \ddots & \\ 0 & & \alpha_{nn} \end{pmatrix}$$

Durch Streichen der ersten Zeile und der der ersten Spalte ergibt sich

$$A_{11} \;=\; \begin{pmatrix} \alpha_{22} & & * \\ & \ddots & \\ 0 & & \alpha_{nn} \end{pmatrix}\,,$$

und die Laplace-Entwicklung nach der ersten Zeile liefert $\det(A) = \alpha_{11}\det(A_{11})$. Die Streichmatrix A_{11} ist wieder eine obere Dreiecksmatrix, hat aber eine Zeile und eine Spalte weniger. Nach Induktionsannahme ist damit $\det(A_{11}) = \alpha_{22}\cdots\alpha_{nn}$, und es folgt die Behauptung. □

Wegen $\det(A^{\mathrm{tr}}) = \det(A)$ gilt dieselbe Aussage auch für untere Dreiecksmatrizen.

SATZ 5.6.2. *Die Determinante einer Matrix lässt sich mit dem Gauß-Jordanschen Eliminationsalgorithmus berechnen.*

BEWEIS. Sei $A \in K^{n\times n}$ beliebig. Der Gauß-Jordan-Algorithmus formt A durch endlich viele elementare Zeilenoperationen zu einer Matrix B in Zeilenstufenform um. Eine Matrix in Zeilenstufenform ist eine obere Dreiecksmatrix. Die elementaren Zeilenoperationen entsprechen der Multiplikation mit Matrizen von links, vgl. Beweis zu Satz 3.8.6.

(E1) $A' = LA$, wobei $L = E_m + \lambda E_{ij}$ für $i \neq j$. Es gilt $\det(L) = 1$.

(E2) $A' = LA$, wobei

$$L \;=\; (\pi_{kl})_{k,l} \quad \text{mit} \quad \pi_{kl} \;=\; \begin{cases} 1 & \text{falls } k = l \notin \{i,j\} \\ 1 & \text{falls } (k,l) \in \{(i,j),(j,i)\} \\ 0 & \text{sonst} \end{cases}$$

Es gilt $\det(L) = \det\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = -1$.

(E3) $A' = LA$, wobei

$$L \;=\; \mathrm{diag}(1,\ldots,1,\lambda,1,\ldots,1) \quad \text{für } \lambda \neq 0 \quad .$$

Es gilt $\det(L) = \lambda$.

Um eine Zeilenstufenform zu erreichen, genügen die Operationen (E1) und (E2). Hieraus folgt

$$\det(A) \;=\; \det(B)\cdot(-1)^t\,, \tag{18}$$

wobei t die Anzahl der Anwendungen von (E2) ist. Die Determinante von B lässt sich aus Proposition 5.6.1 berechnen, und aus dem Determinantenmultiplikationssatz 5.4.2 erhält man die Determinante von A aus (18). $\square$

Beispiel 5.6.3. Sei $K = \mathbb{C}$ und

$$A \;=\; \begin{pmatrix} \boxed{1} & 0 & 1+i & 0 \\ i & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & -1 & 0 & i \end{pmatrix} \xrightarrow{\text{(E1)}} \begin{pmatrix} 1 & 0 & 1+i & 0 \\ 0 & \boxed{1} & 1-i & 0 \\ 0 & 0 & 1 & 2 \\ 0 & -1 & 0 & i \end{pmatrix}$$

$$\xrightarrow{\text{(E1)}} \begin{pmatrix} 1 & 0 & 1+i & 0 \\ 0 & 1 & 1-i & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1-i & i \end{pmatrix} \xrightarrow{\text{(E2)}} \begin{pmatrix} 1 & 0 & 1+i & 0 \\ 0 & 1 & 1-i & 0 \\ 0 & 0 & \boxed{\text{1 - i}} & i \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\xrightarrow{\text{(E1)}} \begin{pmatrix} 1 & 0 & 1+i & 0 \\ 0 & 1 & 1-i & 0 \\ 0 & 0 & 1-i & i \\ 0 & 0 & 0 & \frac{1}{2}(5-i) \end{pmatrix}$$

Für den letzten Schritt beachte, dass $(1-i)^{-1} = \frac{1}{2}(1+i)$ ist, also

$$2 - \frac{1}{2}(1+i)i \;=\; 2 - \frac{1}{2}(i-1) \;=\; \frac{5}{2} - \frac{i}{2}$$

Da einmal (E2) angewendet wurde, ergibt sich

$$\det A \;=\; 1\cdot 1\cdot(1-i)\cdot\frac{1}{2}(5-i)\cdot(-1) \;=\; 3i-2\,.$$

5.7. Matrixinversion und Adjunkte

Für $A = (\alpha_{ij}) \in K^{n\times n}$ setze

$$\alpha^{\#}_{ik} \;:=\; (-1)^{i+k}\det A_{ki} \quad \text{für } 1 \le i,k \le n$$

und $A^{\#} := (\alpha^{\#}_{ik}) \in K^{n\times n}$. Die Matrix $A^{\#}$ heißt *Adjunkte* von A.

Proposition 5.7.1. *Es gilt $A\cdot A^{\#} = A^{\#}A = \det A\cdot E_n$.*

Beweis. Der Koeffizient der Matrix $AA^{\#}$ an der Stelle (i,k) lautet

$$\sum_{j=1}^{n}\alpha_{ij}\alpha^{\#}_{jk} \;=\; \sum_{j=1}^{n}(-1)^{j+k}\,\alpha_{ij}\,\det\,A_{kj} \;=\; \det A'\,,$$

wobei A' aus A entsteht, wenn man die k-te Zeile durch die i-te ersetzt. Es ergibt sich sofort

$$\sum_{j=1}^{n}\alpha_{ij}\,\alpha^{\#}_{jk} \;=\; \begin{cases} 0, & \text{falls } i \ne k \\ \det A, & \text{falls } i = k\,. \end{cases}$$

$\square$

Korollar 5.7.2. *Falls $A \in \mathrm{GL}_n K$ ist, dann gilt*

$$A^{-1} \;=\; \frac{1}{\det A}\cdot A^{\#}\,.$$

5.8. Die Determinantenabbildung aus dem Blickwinkel der Analysis

Wir betrachten den Fall $K = \mathbb{R}$.

PROPOSITION 5.8.1. *Die Abbildung*

$$\det : \mathbb{R}^{n\times n} =_{\uparrow} \mathbb{R}^{n^2} \to \mathbb{R}$$

ist eine differenzierbare Abbildung.

Dies folgt aus der Konstruktion von det per Laplaceentwicklung. Insbesondere ist det stetig. Dies impliziert, dass die Menge

$$\mathrm{GL}_n \mathbb{R} \;=\; \det^{-1}(\mathbb{R} \setminus \{0\})$$

eine offene Teilmenge von $\mathbb{R}^{n\times n}$ ist.
Aus Korollar 5.7.2 wissen wir, dass für $A \in \mathrm{GL}_n \mathbb{R}$ gilt, dass

$$A^{-1} \;=\; \frac{1}{\det A} \cdot A^{\#}$$

ist mit $A^{\#} = (\alpha_{ik}^{\#})$ und

$$\alpha_{ik}^{\#} \;=\; (-1)^{i+k} \det A_{ki} \,.$$

Außerdem ist für ein festes Paar (i, k) die Abbildung

$$A \mapsto \frac{1}{\det A} \cdot \alpha_{ik}^{\#} \;:\; \mathrm{GL}_n \mathbb{R} \to \mathbb{R}$$

differenzierbar. Hieraus ergibt sich unmittelbar:

PROPOSITION 5.8.2. *Die Abbildung*

$$A \mapsto A^{-1} \;:\; \mathrm{GL}_n \mathbb{R} \to \mathrm{GL}_n \mathbb{R}$$

ist differenzierbar.

Die Situation für $K = \mathbb{C}$ ist entsprechend.

BEMERKUNG 5.8.3. Der Betrag der Determinante der reellen Matrix A ist genau das Volumen des Parallelotops, das von den Spalten (oder den Zeilen) von A aufgespannt wird.

5.9. Cramersche Regel

Betrachte das lineare Gleichungssystem

$$Ax \;=\; b \tag{19}$$

für $A \in K^{n\times n}$ und $b \in K^n$ über einem beliebigen Körper K. Bekanntermaßen ist (19) genau dann eindeutig lösbar, wenn $\det A \neq 0$, also $A \in \mathrm{GL}_n K$ ist. Das heißt, falls A invertierbar, ist

$$x_0 \;=\; A^{-1}b$$

die eindeutige Lösung von (19).
Sei nun $A \in \mathrm{GL}_n K$ mit Spalten $(a_1, \dots, a_n)$. Dann gilt mit Korollar 5.7.2 $A^{-1} = \frac{1}{\det A} \cdot A^{\#}$ mit $\alpha_{ik}^{\#} = (-1)^{i+k} \cdot \det A_{ki}$. Oder anders ausgedrückt: $A^{-1} = (\gamma_{ik})$ mit

$$\gamma_{ik} \;=\; \frac{(-1)^{i+k}}{\det A} \cdot \det A_{ki} \;=\; \frac{1}{\det A} \det(a_1, \dots, a_{i-1}, e_k, a_{i+1}, \dots, a_n)$$

Sei $x_0 = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ die eindeutige Lösung von (19) für $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Dann ist

$$\xi_i = \sum_{k=1}^{n} \frac{(-1)^{i+k}}{\det A} \cdot \det A_{ki} \cdot b_k = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A}. \tag{20}$$

Die Formel (20) heißt *Cramersche Regel.*

Wir spezialisieren wieder $K = \mathbb{R}$. Seien $A : U \to \mathrm{GL}_n\,\mathbb{R} : t \mapsto A_t$ und $b : U \to \mathbb{R}^n : t \mapsto b_t$ stetige (differenzierbare) Abbildungen für ein (offenes/geschlossenes/halboffenes/uneigentliches) Intervall $U \subseteq \mathbb{R}$. Sei ferner $x_t \in \mathbb{R}^n$ die eindeutige Lösung des linearen Gleichungssystems $A_t x = b_t$. Dann ist die Abbildung

$$t \mapsto x_t : \mathbb{R} \to \mathbb{R}^n$$

ebenfalls stetig (differenzierbar).

Beispiel 5.9.1. Wir betrachten

$$A : [0,5] \to \mathrm{GL}_2\,\mathbb{R} : t \mapsto \begin{pmatrix} \sin t & \frac{1}{10}\sin(t^2) - \frac{11}{10} \\ 1 + t^2 & \cos t \end{pmatrix}$$

und

$$b : [0,5] \to \mathrm{GL}_2\,\mathbb{R} : t \mapsto \begin{pmatrix} 2t \\ 1 \end{pmatrix}$$

Die durch $A_t x_t = b_t$ implizit definierte Lösungskurve $x : [0,5] \to \mathbb{R}^2$ ist in Abbildung 5.9.1 skizziert.

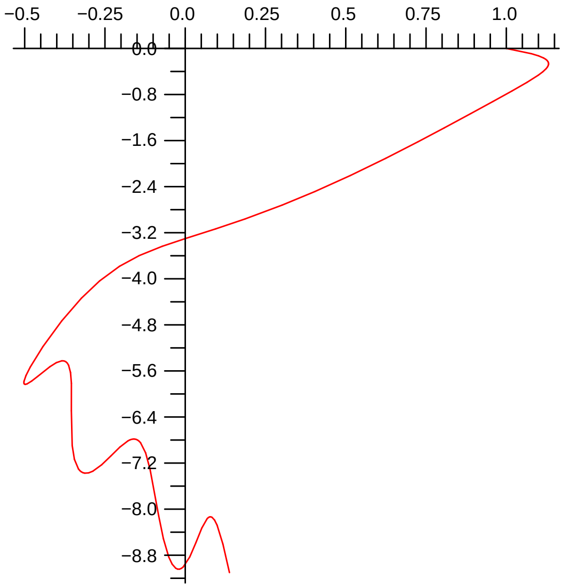

Abbildung 1. Lösungskurve aus Beispiel 5.9.1.

5.10. Exkurs: Algorithmische Aspekte

5.10.1. Numerische Stabilität. Naive Implementierung des Eliminationsverfahrens ergibt schnell Probleme mit beschränkter Rechengenauigkeit. Selbst wenn man sorgfältig implementiert, stößt man aber an Grenzen, wenn die beteiligten Matrizen ungünstig *konditioniert* sind.

Die folgende Rechnung wurde mit Maple (Version 10) ausgeführt.

```
with(LinearAlgebra):
A:=proc(k)
  Matrix([[1-10^(-k),1-2*10^(-k)],
  [1-2*10^(-k),1-3*10^(-k)]]);
end:
```

Jede der so definierten Matrizen A_k ist invertierbar:

```
A(k);
```

$$\begin{bmatrix} 1-10^{-k} & 1-2\cdot 10^{-k} \\ 1-2\cdot 10^{-k} & 1-3\cdot 10^{-k} \end{bmatrix}$$

```
A(1);
```

$$\begin{bmatrix} \frac{9}{10} & 4/5 \\ 4/5 & \frac{7}{10} \end{bmatrix}$$

Wir wählen eine rechte Seite in Abhängigkeit von k so, dass $(1,0)$ stets die eindeutige Lösung des Gleichungssystems $A_k x = b_k$ ist.

```
b:=proc(k)
  A(k).Vector[column]([1,0]);
end:
evalf(A(1)),evalf(b(1));
```

$$\begin{bmatrix} 0.9000000000 & 0.8000000000 \\ 0.8000000000 & 0.7000000000 \end{bmatrix}, \begin{bmatrix} 0.9000000000 \\ 0.8000000000 \end{bmatrix}$$

```
GaussianElimination(A(1));
```

$$\begin{bmatrix} \frac{9}{10} & 4/5 \\ 0 & -\frac{1}{90} \end{bmatrix}$$

Mit exakter Arithmetik lässt sich das Gleichungssystem für beliebig hohes k korrekt lösen:

```
LinearSolve(A(10000),b(10000));
```

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

```
LUDecomposition(A(1));
```

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \frac{8}{9} & 1 \end{bmatrix}, \begin{bmatrix} \frac{9}{10} & 4/5 \\ 0 & -\frac{1}{90} \end{bmatrix}$$

```
A(1).MatrixInverse(GaussianElimination(A(1)));
```

$$\begin{bmatrix} 1 & 0 \\ \frac{8}{9} & 1 \end{bmatrix}$$

Jetzt probieren wir das eingebaute numerische Verfahren für $k \in \{1,2,\dots,8\}$.

```
for k from 1 to 8 do
  LinearSolve(evalf(A(k)),evalf(b(k)))
od;
```

$$\begin{bmatrix} 1.00000000000001754 \\ 0.99999999999998002 \end{bmatrix}$$

$$\begin{bmatrix} 1.0 \\ -0.0 \end{bmatrix}$$

$$\begin{bmatrix} 1.0 \\ -0.0 \end{bmatrix}$$

$$\begin{bmatrix} 0.99999999988919974 \\ 0.000000000110911280150349909 \end{bmatrix}$$

$$\begin{bmatrix} 1.0 \\ -0.0 \end{bmatrix}$$

$$\begin{bmatrix} 1.00000111020196991 \\ -0.00000111021307209279602 \end{bmatrix}$$

$$\begin{bmatrix} 1.0 \\ -0.0 \end{bmatrix}$$

$$\begin{bmatrix} 1.0 \\ -0.0 \end{bmatrix}$$

Error, (in LinearAlgebra:-LA_Main:-BackwardSubstitute) inconsistent system

Das numerische Verfahren bricht bereits bei $k = 8$ mit einer Fehlermeldung ab, die korrekt darüber informiert, dass die Genauigkeit hier verloren gegangen ist. Die zuvor berechneten Lösungen (für $k \leq 7$) weichen nur in geringem Maß von der richtigen Lösung ab. Eine naive Implementierung hat bereits bei diesen simplen Beispielen kaum eine Chance auch nur in die Nähe der korrekten Lösung zu kommen.

Interessant ist nun $k = 9$. Hier gibt Maple eine falsche Lösung aus, ohne dies zu erkennen. Überdies wird nicht-existente Genauigkeit suggeriert, da sich scheinbar ein glattes Ergebnis bei $(1/2, 1/2)$ ergibt:

LinearSolve(evalf(A(9)), evalf(b(9)));

$$\begin{bmatrix} 0.500000000500000041 \\ 0.500000000000000000 \end{bmatrix}$$

Um sicher zu gehen, hier noch einmal die Eingabe in obiges Verfahren. Dies soll zeigen, dass das Problem nicht daran liegt, dass die Eingabegenauigkeit bereits überschritten wäre (wie etwa bei $k = 20$).

evalf(A(9)), evalf(b(9));

$$\begin{bmatrix} 0.9999999990 & 0.9999999980 \\ 0.9999999980 & 0.9999999970 \end{bmatrix}, \begin{bmatrix} 0.9999999990 \\ 0.9999999980 \end{bmatrix}$$

Der Grund dafür, dass die Numerik hier so kompliziert ist, liegt darin, dass die beiden *Eigenwerte* der Matrix A_k auf sehr unterschiedlichen Skalen leben. Die Matrizen A_k besitzen eine sehr hohe *Konditionszahl*. Dies äußert sich beispielsweise darin, dass die Determinante fast verschwindet.

for k from 1 to 10 do evalf(Eigenvalues(A(k))) od;

$$\begin{bmatrix} 1.606225775 \\ -0.0062257748 \end{bmatrix}$$

$$\begin{bmatrix} 1.960051019 \\ -0.0000510191 \end{bmatrix}$$

$$\begin{bmatrix} 1.996000501 \\ -0.0000005010 \end{bmatrix}$$

$$\begin{bmatrix} 1.999600005 \\ -0.0000000050 \end{bmatrix}$$

$$\begin{bmatrix} 1.999960000 \\ -0.0000000001 \end{bmatrix}$$

$$\begin{bmatrix} 1.999996000 \\ 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 1.999999600 \\ 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 1.999999960 \\ 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 1.999999996 \\ 0.0 \end{bmatrix}$$

$$\begin{bmatrix} 2.0 \\ 0.0 \end{bmatrix}$$

5.10.2. Komplexitätsbetrachtungen.

Bemerkung 5.10.1. Der Aufwand, die Determinante einer $n \times n$-Matrix auszurechnen beträgt im ungünstigsten Fall:

(i) $O(n^{\log n})$ via Leibnizformel,
(ii) $O(n^{\log n})$ via rekursiver Laplace-Elimination,
(iii) $O(n^3)$ via Gauß-Jordan-Elimination.

Gezählt werden jeweils arithmetische Operationen (Addition, Multiplikation, Subtraktion, Division) im Grundkörper K.

KAPITEL 6

Polynome

6.1. Arithmetik

Sei R ein kommutativer Ring mit Eins.

Definition 6.1.1.

(i) Ein (abstraktes) *Polynom* mit *Koeffizienten* in R ist eine Abbildung

$$p : \mathbb{N} \to R$$

mit $\#\{i \in \mathbb{N} : p(i) \neq 0\} < \infty$. Das heißt die abstrakten Polynome entsprechen genau den endlichen Folgen in R (beliebiger Länge).

(ii) Die Zahl $\deg p := \max\{i \in \mathbb{N} : p(i) \neq 0\}$ für $p \neq 0$ heißt *Grad* von p. Wir setzen $\deg 0 := -\infty$.

(iii) Für $p \neq 0$ heißt der Koeffizient $\mathrm{lc}(p) := p(\deg p)$ der *Leitkoeffizient* von p.

Bemerkung 6.1.2. Das *Nullpolynom* $0 = (0, 0, 0, \dots)$ besitzt keinen Leitkoeffizienten. Für $p \neq 0$ gilt stets $\mathrm{lc}(p) \neq 0$.

Üblicherweise schreibt man das Polynom $(\alpha_0, \alpha_1, \dots, \alpha_k, 0, 0, 0, \dots)$ als formale Summe

$$\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_k t^k ,$$

wobei t ein beliebiges nicht mit einer anderen Bedeutung belegtes Symbol (*Unbestimmte*) ist. Die Menge aller Polynome mit Koeffizienten in R über der Unbestimmten t wird mit $R[t]$ bezeichnet. Via der Einbettung $r \mapsto (r, 0, 0, 0, \dots)$ gilt $R \subset R[t]$.

Auf $R[t]$ lassen sich eine Addition $+$ und eine Multiplikation $\cdot$ definieren. Seien $p, q \in R[t]$.

(i) $p + q : \mathbb{N} \to R : i \mapsto p(i) + q(i)$, das heißt die Addition von Polynomen ist die punktweise Addition von R-wertigen Funktionen.

(ii) $p \cdot q : \mathbb{N} \to R : k \mapsto \sum_{i=0}^{k} p(i) \cdot q(k - i)$.

In der üblichen Notation sieht das dann für $p = \sum_{i=0}^{k} \alpha_i t^i$ und $q = \sum_{i=0}^{l} \beta_i t^i$ so aus:

$$p + q \;=\; \sum_{i=0}^{\max(k,l)} (\alpha_i + \beta_i) t^i \quad \text{und} \quad p \cdot q \;=\; \sum_{i=0}^{k+l} \Big(\sum_{j=0}^{i} \alpha_j \cdot \beta_{i-j} \Big) t^i .$$

Es gilt $\deg(p + q) \leq \max(\deg(p), \deg(q))$ und $\deg(p \cdot q) = \deg p + \deg q$.

Proposition 6.1.3. *Das Tripel* $(R[t], +, \cdot)$ *ist ein Ring mit additivem Neutralelement* 0 *und multiplikativem Neutralelement* 1.

6.2. Polynome mit Koeffizienten in einem Körper

Sei K ein Körper und t eine Unbestimmte.

PROPOSITION 6.2.1. *Mit der Polynomaddition und der Skalarmultiplikation*

$$\lambda \cdot p \ := (i \mapsto \lambda \cdot p(i) \ : \ \mathbb{N} \to K)$$

für $\lambda \in K$ *und* $p \in K[t]$ *ist* $K[t]$ *ein* K*-Vektorraum.*

Da das Einselement des Körpers K mit dem Einselement des Rings $K[t]$ übereinstimmt, ist $K[t]$ insgesamt eine K-Algebra.

PROPOSITION 6.2.2. *Die Polynome* $1, t, t^2, \ldots$ *bilden eine Basis von* $K[t]$*. Insbesondere gilt* $\dim_K K[t] = \infty$*.*

BEWEIS. Jedes Polynom $p = \sum_{i=0}^{d} \alpha_i t^i$ mit $d = \deg p$ lässt sich eindeutig darstellen als

$$\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_d t^d \, .$$

□

PROPOSITION 6.2.3 (Division mit Rest). *Zu* $a, b \in K[t]$ *mit* $b \neq 0$ *existieren eindeutig bestimmte Polynome* $q, r \in K[t]$ *mit* $a = q \cdot b + r$ *und* $\deg r < \deg b$*.*

BEWEIS. Der Beweis ist konstruktiv. Seien $a, b \in K[t]$ mit $b \neq 0$.

Wir setzen $q_1 := 0$ und $r_1 := a$. Wir nehmen an, dass ein $k \geq 1$ gilt $\deg r_k \geq \deg b$. Dann definieren wir induktiv:

$$q_{k+1} \ := \ \operatorname{lc}(r_k) / \operatorname{lc}(b) t^{\deg r_k - \deg b} \tag{21}$$

$$r_{k+1} \ := \ r_k - q_{k+1} \cdot b \tag{22}$$

Anschließend wird k um eins erhöht und die Induktion in einer Schleife fortgesetzt. Das Verfahren terminiert, weil der Grad von r_k sich in jedem Schleifendurchlauf vermindert.

Für $k = 1$ gilt $a = 0 \cdot b + a = q_1 \cdot b + r_1$. Induktiv gilt dann für $k \geq 1$, dass

$$\begin{aligned}
\Big(\sum_{i=1}^{k+1} q_k\Big) \cdot b + r_{k+1} \ &= \ \Big(\sum_{i=1}^{k} q_k\Big) \cdot b + q_{k+1} \cdot b + r_k - q_{k+1} \cdot b \\
&= \ a - r_k + q_{k+1} \cdot b + r_k - q_{k+1} \cdot b \ = \ a \, .
\end{aligned}$$

Sei nun k minimal, so dass $\deg r_k < \deg b$. Dann erfüllen $q := \sum_{i=1}^{k} q_k$ und $r := r_k$ die geforderten Bedingungen.

Der Beweis der Eindeutigkeit ist eine Übungsaufgabe. □

Die q_{k+1} definierende Gleichung (21) ist die einzige Stelle, an der wir benutzt haben, dass K ein Körper ist. Genauer gesagt: Wir haben verwendet, dass $\operatorname{lc}(b)$ in K stets invertierbar ist. Betrachten wir statt K einen beliebigen kommutativen Ring, so gilt die Proposition 6.2.3 für jedes Polynom b, dessen Leitkoeffizient $\operatorname{lc}(b)$ invertierbar ist.

6.3. Einsetzungsabbildung und Nullstellen

Sei nun wieder R ein beliebiger kommutativer Ring mit Eins. Zu jedem Polynom $p = \sum \alpha_i t^i \in R[t]$ existiert eine Abbildung

$$p^* \,:\, R \to R \,:\, x \mapsto \alpha_0 + \alpha_1 x + \cdots + \alpha_d x^d\,,$$

die *Einsetzungsabbildung* von p.

Beispiel 6.3.1. Betrachte $p = t^2 - 1 \in \mathbb{R}[t]$. Dann ist

$$p^*(0) = -1,\ p^*(1) = p^*(-1) = 0\,.$$

Beispiel 6.3.2. Betrachte $p = t^2 + t \in \mathbb{F}_2[t]$. Dann ist

$$p^*(0) = 0^2 + 0 = 0,\ p^*(1) = 1^2 + 1 = 0\,.$$

Beachte, dass p^* identisch verschwindet, obwohl $p \neq 0$ gilt.

Beispiel 6.3.3. Betrachte $p = t^2 + t + 1 \in \mathbb{Q}^{2\times 2}[t]$. Dann ist

$$\begin{aligned} p^*\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}^2 + \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 2 & 3 \end{pmatrix}. \end{aligned}$$

Definition 6.3.4. Ein Element $x \in R$ heißt *Nullstelle* des Polynoms $p \in R[t]$, falls $p^*(x) = 0$ ist.

Proposition 6.3.5. *Sei x eine Nullstelle von $p \in R[t]$. Dann gilt, dass $t - x$ das Polynom p teilt.*

Die *Teilbarkeit* in einem beliebigen Ring ist definiert wie für ganze Zahlen.

Beweis. Wegen $\mathrm{lc}(t - x) = 1$ ist die Polynomdivision in $R[t]$ wie in Proposition 6.2.3 möglich. Ohne Einschränkung sei $d := \deg p \geq 1$. Für $p = \sum_{i=0}^d \alpha_i t^i$ ergibt die Division mit Rest durch $(t - x)$ dann

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_d x^d \;=\; p^*(x) \;=\; 0\,.$$

□

Korollar 6.3.6. *Jedes Polynom vom Grad d hat höchstens d (verschiedene) Nullstellen.*

Für das Folgende vergleiche mit den Beispielen 6.3.1 und 6.3.2.

Beispiel 6.3.7. Die Nullstellen von $t^2 - 1 \in \mathbb{R}[t]$ sind genau 1 und -1.

Beispiel 6.3.8. Die Nullstellen von $t^2 + t \in \mathbb{F}_2[t]$ sind genau 0 und 1.

Proposition 6.3.9. *Sei K ein Körper und $A \in K^{n\times n}$. Dann ist die Abbildung*

$$\Phi_A \,:\, K[t] \to K^{n\times n} \,:\, p \mapsto p^*(A)$$

ein Homomorphismus von K-Algebren, das heißt, eine K-lineare Abbildung, die gleichzeitig ein Ringhomomorphismus ist.

Beweis. Übungsaufgabe. □

6.4. Die Einsetzungsabbildung für Matrizen

Sei K ein Körper.

PROPOSITION 6.4.1. *Für jede invertierbare Matrix* $A \in \mathrm{GL}_n K$ *existiert ein Polynom* $p = \sum \alpha_i t^i \in K[t] \subset K^{n\times n}[t]$ *vom Grad* $d < n^2$ *mit*

$$A^{-1} \;=\; p^*(A) \;=\; \alpha_0 E_n + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_d A^d$$

und $\alpha_0 \neq 0$.

BEWEIS. Wegen $A(\lambda B + \mu C) = \lambda AB + \mu AC$ ist die Abbildung

$$L_A \,:\, K^{n\times n} \to K^{n\times n} \,:\, B \mapsto AB$$

linear. Da $A \in \mathrm{GL}_n K$, ist L_A bijektiv; das heißt, $L_A \in \mathrm{GL}(K^{n\times n})$.
Sei $k = \min\left\{i \in \mathbb{N} :\, (E_n = A^0, A = A^1, A^2, \ldots, A^k) \text{ linear abhängig}\right\}$. Dann existieren $\beta_0, \beta_1, \ldots, \beta_{k-1} \in K$, so dass

$$A^k \;=\; \beta_0 E_n + \beta_1 A + \beta_2 A^2 + \cdots + \beta_{k-1} A^{k-1}\,. \tag{23}$$

Dann gilt

$$\begin{aligned} A^{k+1} \;&=\; A(\beta_0 E_n + \beta_1 A + \beta_2 A^2 + \cdots + \beta_{k-1} A^{k-1}) \;=\; \beta_0 A + \beta_1 A^2 + \cdots + \beta_{k-1} A^k \\ &= \beta_0 A + \beta_1 A^2 + \cdots + +\beta_{k-2} A^{k-1} + \beta_{k-1}(\beta_0 E_n + \beta_1 A + \beta_2 A^2 + \cdots + \beta_{k-1} A^{k-1}) \\ &= \beta_0\beta_{k-1} E_n + (\beta_0 + \beta_1\beta_{k-1})A + (\beta_1 + \beta_2\beta_{k-1})A^2 + \cdots + (\beta_{k-2} + \beta_{k-1}\beta_{k-1})A^{k-1}\,. \end{aligned}$$

Dies bedeutet insbesondere, dass der Unterraum

$$U \;=\; \mathrm{lin}\left\{A^k :\, k \geq 1\right\} \;\leq\; K^{n\times n}\,.$$

die Basis $\mathcal{B} := (E_n = A^0, A^1 = A, A^2, \ldots, A^{k-1})$ besitzt. Es gilt $\dim U = k$, und $k-1 < n^2$, weil U ein Teilraum von $K^{n\times n}$ ist. Nach Konstruktion ist $L_A(U) \subseteq U$. Da aber L_A bijektiv und U endlichdimensional ist, folgt $L_A(U) = U$. Also ist ${L_A}^{-1}(U) = L_{A^{-1}}(U) = U$. Damit ist

$$A^{-1}(E_n, A, A^2, \ldots, A^{k-1}) \;=\; (A^{-1}, E_n, A, A^2, \ldots, A^{k-2})$$

ebenfalls eine Basis von U. Insbesondere ist $A^{-1} \in U$, und es existieren eindeutig bestimmte Koeffizienten $\alpha_0, \alpha_1, \ldots, \alpha_{k-1} \in K$ mit

$$A^{-1} \;=\; \alpha_0 E_n + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_{k-1} A^{k-1}\,.$$

Aus der Gleichung (23) folgt dann

$$\begin{aligned} E_n \;&=\; AA^{-1} \;=\; \alpha_0 A + \alpha_1 A^2 + \cdots + \alpha_{k-1} A^k \\ &=\; \alpha_0 A + \alpha_1 A^2 + \cdots + \alpha_{k-2} A^{k-1} + \alpha_{k-1}(\beta_0 E_n + \beta_1 A + \beta_2 A^2 + \cdots + \beta_{k-1} A^{k-1}) \\ &=\; \alpha_{k-1}\beta_0 E_n + (\alpha_0 + \alpha_{k-1}\beta_1)A + (\alpha_1 + \alpha_{k-1}\beta_2)A^2 + \cdots + (\alpha_{k-2} + \alpha_{k-1}\alpha_{k-1})A^{k-1} \end{aligned}$$

Aus der eindeutigen linearen Darstellbarkeit von E_n bezüglich der Basis $\mathcal{B}$ folgt, dass $\alpha_{k-1}\beta_0 = 1$ ist. Dies bedeutet $\alpha_{k-1} \neq 0$ und $\beta_0 \neq 0$. Induktiv zeigt man durch Betrachten von $A^i = A^{i+1}A^{-1}$, dass $\beta_i \neq 0$ für alle $i \geq 0$ gilt. Wiederum wegen der eindeutigen linearen Darstellbarkeit von E_n folgt dann $\alpha_0 + \alpha_{k-1}\beta_1 = 0$, und hieraus $\alpha_0 \neq 0$. □

BEISPIEL 6.4.2. Betrachte

$$A \;=\; \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \;\in\; \mathbb{Q}^{2\times 2}$$

wie in Beispiel 6.3.3. Dann ist

$$A^2 \;=\; \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$$

Wegen

$$\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = -2 \neq 0$$

ist $A \in \mathrm{GL}_2\,\mathbb{Q}$. Es gilt

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

Damit ist der Unterraum U aus dem Beweis zu Proposition 6.4.1

$$U = \operatorname{lin}\{E_n, A, A^2, \dots\}$$

zweidimensional mit

$$(E_n, A)$$

als einer Basis. Wir erhalten

$$\begin{aligned} A^{-1} &= A^{-1}E_n = A^{-1}(\frac{1}{2}(A^2 - A)) \\ &= \frac{1}{2}(A - E_n) = -\frac{1}{2}E_n - \frac{1}{2}A\,. \end{aligned}$$

Das heißt $p = -1/2 - 1/2t \in \mathbb{Q}^{2\times 2}[t]$ ist ein Polynom mit $p^*(A) = A^{-1}$.

6.5. Interpolation

Sei K stets ein Körper.

Satz 6.5.1 (Lagrange-Interpolation). *Gegeben seien $d+1$ paarweise verschiedene* Stützstellen $x_0, x_1, \dots, x_d \in K$ *mit* $d+1$ Stützwerten $y_0, y_1, \dots, y_d \in K$. *Dann existiert ein eindeutig bestimmtes Polynom $p \in K[t]$ vom Grad höchstens d mit der Eigenschaft $p^*(x_i) = y_i$ für alle $i \in \{0, 1, \dots, d\}$.*

Beweis. Betrachte die *Lagrange-Polynome*

$$L_i = \prod_{k=0}^{d} \frac{t - x_k}{x_i - x_k}$$

für $i \in \{0, 1, \dots, d\}$. Es gilt $\deg L_i = d$ sowie

$$L_i^*(x_k) = \delta_{ik} = \begin{cases} 1 & \text{falls } i = k \\ 0 & \text{sonst.} \end{cases}$$

Hieraus folgt unmittelbar, dass $p := \sum_{i=0}^{d} y_0 L_i$ die gewünschte Interpolationseigenschaft besitzt. Offenbar gilt $\deg p \leq d$.

Angenommen, es gäbe ein zweites Polynom q vom Grad höchstens d mit $q^*(x_i) = y_i$. Dann gilt für die Differenz $r := p - q$, dass $r^*(x_i) = p^*(x_i) - q^*(x_i) = y_i - y_i = 0$ ist für alle $i \in \{0, 1, \dots, d\}$. Also hat r mindestens $d+1$ paarweise verschiedene Nullstellen. Wäre $r \neq 0$ ergäbe sich hieraus ein Widerspruch zu Korollar 6.3.6. □

Satz 6.5.2 (Viète-Formeln). *Für $p \in K[t]$ vom Grad $d \geq 1$ und $k \in \{1, 2, \dots, d\}$ mit*

$$p = \sum_{i=0}^{d} \alpha_i t^i = \alpha_d \prod_{i=1}^{d} (t - \beta_i)\,.$$

gilt

$$\sum_{i_1 < i_2 < \dots < i_k} \beta_{i_1}\beta_{i_2}\dots\beta_{i_k} = (-1)^{d-k}\alpha_{d-k}/\alpha_d\,.$$

Beweis. Beweis per Induktion nach d. Für $d = 1$ gilt $\alpha_1(t - \beta_1) = \alpha_1 t - \alpha_1\beta_1 = \alpha_1 t + \alpha_0$. Durch Koeffizientenvergleich erhält man $-\alpha_1\beta_1 = \alpha_0$ oder $\beta_1 = -\alpha_0/\alpha_1$. Im Induktionsschritt gilt dann

$$\begin{aligned}
\alpha_d \prod_{i=1}^{d}(t-\beta_i) &= \alpha_d(t-\beta_d)\cdot\prod_{i=1}^{d-1}(t-\beta_i)\\
&= \alpha_d(t-\beta_d)\Big(t^{d-1} + \sum_{k=1}^{d-1}(-1)^{d-1-k}\sum_{i_1<i_2<\cdots<i_k\le d-1}\beta_{i_1}\beta_{i_2}\cdots\beta_{i_k}\,t^{d-1-k}\Big)\\
&= \alpha_d\Big(t^d - \sum_{k=1}^{d-1}(-1)^{d-k}\sum_{i_1<i_2<\cdots<i_k\le d-1}\beta_{i_1}\beta_{i_2}\cdots\beta_{i_k}\,t^{d-k}\\
&\qquad - \beta_d t^{d-1} - \sum_{k=2}^{d}(-1)^{d-k}\sum_{i_1<i_2<\cdots<i_{k-1}\le d-1}\beta_{i_1}\beta_{i_2}\cdots\beta_{i_{k-1}}\beta_d\,t^{d-k}\Big)
\end{aligned}$$

Die beiden Summen lassen sich nun zusammen fassen, wenn man die Randterme extra betrachtet. Wir erhalten

$$\begin{aligned}
\alpha_d \prod_{i=1}^{d}(t-\beta_i) &= \alpha_d\Big(t^d + (-1)^{d-1}\beta_d t^{d-1} + (-1)^{d-1}\sum_{i_1\le d-1}\beta_{i_1}t^{d-1} - \beta_1\beta_2\cdots\beta_{d-1}\beta_d\\
&\qquad - \sum_{k=2}^{d-2}(-1)^{d-k}\sum_{i_1<i_2<\cdots<i_k\le d}\beta_{i_1}\beta_{i_2}\cdots\beta_{i_k}\,t^{d-k}\Big)\\
&= \alpha_d\Big(t^d + \sum_{k=1}^{d-1}(-1)^{d-k}\sum_{i_1<i_2<\cdots<i_k\le d}\beta_{i_1}\beta_{i_2}\cdots\beta_{i_k}\,t^{d-k}\Big)
\end{aligned}$$

□

6.6. Fundamentalsatz der Algebra

Satz 6.6.1 (Fundamentalsatz der Algebra). *Jedes nicht-konstante Polynom $p \in \mathbb{C}[t]$ zerfällt in Linearfaktoren.*

Beweis. Der Beweis dieses Satzes geht über die Möglichkeiten einer Anfängervorlesung hinaus. Es gehen wesentlich analytische bzw. topologische bzw. funktionentheoretische Methoden ein. Siehe `en.wikipedia.org` für einen Überblick über verschiedene Beweisstrategien. □

Korollar 6.6.2. *Jedes nicht-konstante Polynom $p \in \mathbb{R}[t]$ zerfällt in lineare und quadratische Faktoren.*

Beweis. Sei $p \in \mathbb{R}[t]$ mit $\deg p = d \ge 1$. Aufgefasst als Polynom mit komplexen Koeffizienten zerfällt p in Linearfaktoren, also

$$p \;=\; (t-z_1)(t-z_2)\cdots(t-z_d)$$

für $z_i \in \mathbb{C}$. Für $z \in \mathbb{C}$ gilt $p^*(\overline{z}) = \overline{p^*(z)}$, weil die Koeffizienten von p reell sind. Insbsondere folgt: Wenn z Nullstelle von p ist, so auch $\overline{z}$. Für $z \in \mathbb{C}\setminus\mathbb{R}$ teilt $(t-z)(t-\overline{z}) \in \mathbb{R}[t]$ das Polynom p. Die komplexen Nullstellen $z_1, z_2, \ldots, z_d$ sind entweder reell, oder sie kommen in konjugiert-komplexen Paaren. Hieraus folgt die Behauptung. □

6.7. Quotientenkörper

DEFINITION 6.7.1. Ein Ring R heißt *nullteilerfrei*, falls für alle $a, b \in R \setminus \{0\}$ gilt $ab \neq 0$.

Sei im folgenden stets R ein kommutativer nullteilerfreier Ring mit Eins. Auf der Menge $R \times (R \setminus \{0\})$ führen wir die folgende Relation ein:

$$(a, b) \sim (c, d) \;:\Leftrightarrow\; ad = bc$$

Man rechnet nach, dass $\sim$ ein Äquivalenzrelation ist.

PROPOSITION 6.7.2. *Die Menge* $\mathrm{Quot}(R) := (R \times (R \setminus \{0\}))/\sim$ *mit der induzierten Addition und Multiplikation ist ein Körper.*

BEISPIEL 6.7.3. $\mathrm{Quot}(\mathbb{Z}) = \mathbb{Q}$.

BEISPIEL 6.7.4. Der Quotientenkörper des Polynomrings $K[t]$ ist der Körper der *rationalen Funktionen* $K(t)$.

Wenn Q der Quotientenkörper von R ist, so lässt sich via der Einbettung

$$x \mapsto (x, 1) \,:\, R \to Q$$

der Ring R als Teilring des Körpers Q auffassen. Typischerweise schreiben wir die Äquivalenzklasse des Paars $(a, b) \in R \times (R \setminus \{0\})$ als *Bruch* a/b. Der Quotientenkörper Q ist also die Menge aller Brüche in R.

6.8. Exkurs: Multivariate Polynome

Sei K ein Körper. Für paarweise verschiedene Unbestimmte $t_1, \ldots, t_n$ sei

$$K[t_1, \ldots, t_n] \;:=\; \underbrace{(K[t_1, \ldots, t_{n-1}])}_{\text{Koeffizienten}}[t_n]$$

induktiv definiert. Dies ist die Algebra der *multivariaten Polynome* in den Unbestimmten $t_1, \ldots, t_n$ mit Koeffizienten in K.
Die folgende Notation ist üblich. Für $m \in \mathbb{N}^n$ sei

$$t^m \;:=\; t^{(m_1, m_2, \ldots, m_n)} \;=\; t_1^{m_1} t_2^{m_2} \cdots t_n^{m_n} \;\in\; K[t_1, \ldots, t_n]$$

das *Monom* mit dem *Exponenten* m.

LEMMA 6.8.1. *Die Monome bilden eine K-Basis von $K[t_1, \ldots, t_n]$.*

Damit hat also $p \in K[t_1, \ldots, t_n]$ eine eindeutige Darstellung

$$p \;=\; \sum_{m \in \mathbb{N}^n} \alpha_m t^m$$

als Linearkombination von Monomen. Klarerweise ist diese Summe endlich, also

$$\#\{m \in \mathbb{N}^n : \alpha_m \neq 0\} \;<\; \infty\,.$$

Zu $p = \sum_{m \in \mathbb{N}^n} \alpha_m t^m \in K[t_1, \ldots, t_n]$ existiert wiederum eine Einsetzungsabbildung

$$p^* \;:\; K^n \to K \;:\; \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{m \in \mathbb{N}^n} \alpha_m x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}\,.$$

Definition 6.8.2. Das Nullstellengebilde

$$V(p) \ := \ \{x \in K^n : \ p^*(x) = 0\}$$

eines multivariaten Polynoms heißt *(affine) algebraische Hyperfläche.*

Oft schreiben wir statt der durchnummerierten Unbestimmten „t_1“, „t_2“, ... auch unterschiedliche Symbole wie „x“, „y“ etc.

Beispiel 6.8.3. Betrachte $p = x^2+y^2-1 \in \mathbb{R}[x,y]$. Dann ist $V(p) \subset \mathbb{R}^2$ der Einheitskreis. Die Hyperfläche $V(y-x^2) \subset \mathbb{R}^2$ ist die Normalparabel, und $V(xy-1) \subset \mathbb{R}^2$ ist die Standardhyperbel.

Definition 6.8.4. Das multivariate Polynom

$$e_k \ := \sum_{i_1<i_2<\cdots<i_k} t_{i_1}t_{i_2}\cdots t_{i_k} \ \in \ K[t_1,t_2,\ldots,t_n]$$

heißt k-tes *elementar-symmetrisches Polynom* in n Unbestimmten.

Die Viète-Formeln aus Satz 6.5.2 lauten dann so: Für ein univariates (monisches) Polynom $p = \sum_{k=0} \alpha_k t^k \in K[t]$ vom Grad $d \geq 1$, das also $\alpha_d = 1$ erfüllt, und das in Linearfaktoern $p = (t-\beta_1)(t-\beta_2)\cdots(t-\beta_d)$ zerfällt, gilt

$$e_k^*(\beta_1,\beta_2,\ldots,\beta_d) \ = \ (-1)^{d-k}\alpha_{d-k} \,.$$

Beispiel 6.8.5 (Steele). Betrachte $p = (t-u)(t-v)(t-w) \in K[u,v,w,t] \subset K(u,v,w)[t]$. Die Viète-Formeln besagen dann

$$p \ = \ t^3 - e_1t^2 + e_2t - e_3$$

wobei e_1, e_2, e_3 die drei elementarsymmetrischen Polynome in $K[u,v,w]$ sind. Dann ist

$$0 \ = \ p^*(u)+p^*(v)+p^*(w) \ = \ (u^3+v^3+w^3) - e_1(u^2+v^2+w^2) + e_2e_1 - 3e_3$$

oder, in expandierter Form

$$(u^3+v^3+w^3) - 3uvw \ = \ (u+v+w)(u^2+v^2+w^2-uv-uw-vw)\,.$$

Die letzte Gleichung gilt für Polynome im Ring $K[u,v,w]$. Substituiert man nun wiederum Elemente aus K, so ergibt sich beispielsweise im Spezialfall $u+v+w=0$ die Konsequenz $u^3+v^3+w^3 = 3uvw$, also etwa $2^3+3^3-5^3 = 8+27-125 = -90 = 3\cdot 2\cdot 3\cdot(-5)$.

Definition 6.8.6. Ein multivariates Polynom $p = \sum_{m\in\mathbb{N}^n} \alpha_m t^m \in K[t_1,t_2,\ldots,t_n]$ heißt *symmetrisch*, falls für alle $\sigma \in \mathrm{Sym}(\{1,2,\ldots,n\})$ gilt

$$p \ = \sum_{m\in\mathbb{N}^n} \alpha_m t_1^{m_1}t_2^{m_2}\cdots t_n^{m_n} \ = \sum_{m\in\mathbb{N}^n} \alpha_m t_{\sigma(1)}^{m_1}t_{\sigma(2)}^{m_2}\cdots t_{\sigma(n)}^{m_n}\,.$$

Proposition 6.8.7. *Die symmetrischen Polynome bilden einen K-Unterraum von $K[t_1,t_2,\ldots,t_n]$, und die elementar-symmetrischen Polynome bilden eine Basis.*

KAPITEL 7

Euklidische und unitäre Vektorräume

7.1. Bilinearformen

Sei K ein beliebiger Körper und V ein K-Vektorraum.

DEFINITION 7.1.1. Eine Abbildung $f : V \times V \to K$ heißt *K-Bilinearform*, falls gilt

$$
\begin{aligned}
f(\alpha u + \beta u', v) &= \alpha f(u,v) + \beta f(u',v) \quad \text{und} \\
f(u, \alpha v + \beta v') &= \alpha f(u,v) + \beta f(u,v')
\end{aligned}
\tag{24}
$$

für alle $\alpha, \beta \in K$ und $u, u', v, v' \in V$. Eine K-Bilinearform f heißt *reflexiv*, falls

$$f(u,v) = 0 \quad \Leftrightarrow \quad f(v,u) = 0$$

für alle $u, v \in V$. Eine reflexive Bilinearform heißt *ausgeartet*, falls ein $u \neq 0$ existiert mit der Eigenschaft, dass $f(u,v) = 0 = f(v,u)$ ist für alle $v \in V$.

BEISPIEL 7.1.2 (Standardskalarprodukt). Sei $V = K^n$ der Standard-K-Vektorraum der Dimension n. Für

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \quad \text{und} \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

setze $f(u,v) = u^{\mathrm{tr}}\, v = u_1\, v_1 + \cdots + u_n\, v_n$. Offenbar ist f bilinear und reflexiv. Außerdem ist f nicht ausgeartet: Sei $u \neq 0$. Ohne Einschränkung sei dann $u_1 \neq 0$. Wähle $v_2, \ldots, v_n \in K$ beliebig. Setze $x = u_2 v_2 + \cdots + u_n v_n$ und $v_1 = (1-x) \cdot u_1^{-1}$. Dann ist

$$u_1 v_1 + \cdots + u_n v_n = u_1 v_1 + x = u_1 (1-x) u_1^{-1} + x = 1\,.$$

BEMERKUNG 7.1.3. Das Standardskalarprodukt ist *symmetrisch*, das heißt es gilt

$$f(u,v) = f(v,u) \tag{25}$$

für alle $u, v \in K^n$. Jede symmetrische Bilinearform ist offensichtlich auch reflexiv.

7.2. Das euklidische Skalarprodukt auf $\mathbb{R}^n$

DEFINITION 7.2.1. Im Fall $K = \mathbb{R}$ heißt das Standardskalarprodukt auch *euklidisches Skalarprodukt*. Eine übliche Notation ist

$$\langle u, v \rangle := u^{\mathrm{tr}} v = u_1 v_1 + \cdots + u_n v_n$$

PROPOSITION 7.2.2. *Das euklidische Skalarprodukt ist* $\mathbb{R}$*-bilinear, symmetrisch und* positiv-definit*, das heißt zusätzlich zu* (24) *und* (25) *für alle* $v \in V$ *gilt*

$$
\begin{aligned}
\langle v, v \rangle &\geq 0 \quad \textit{und} \\
\langle v, v \rangle &> 0 \quad \textit{falls } v \neq 0\,.
\end{aligned}
\tag{26}
$$

DEFINITION 7.2.3. Die *euklidische Norm* (oder *Länge*) eines Vektors $v \in \mathbb{R}^n$ wird definiert als

$$\|v\| := \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + \cdots + v_n^2}\,.$$

7.3. Das hermitesche Skalarprodukt auf $\mathbb{C}^n$

Wir suchen ein „passendes“ Skalarprodukt auf $\mathbb{C}^n$.

Beispiel 7.3.1. Sei $f : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ das Standardskalarprodukt auf $\mathbb{C}^2$. Dann gilt $f((1,i),(1,i)) = 1^2 + i^2 = 0$. Daher ist das Standardskalarprodukt auf $\mathbb{C}^2$ keine Erweiterung des euklidischen Skalarprodukts, die irgendwie mit einem Längenbegriff in Zusammenhang zu bringen wäre.

Bekanntermaßen ist die *Konjugation*

$$\bar{\ }: z = x + iy \mapsto \bar{z} := x - iy$$

ein Körperautomorphismus der komplexen Zahlen mit der Eigenschaft $\bar{\bar{z}} = z$ für alle $z \in \mathbb{C}$. Die komplexen Zahlen lassen sich als Paare reeller Zahlen auffassen: $\mathbb{C} =_{\uparrow} \mathbb{R}^2$ und besitzen somit bereits eine euklidische Länge:

$$|z| \ = \ |x+iy| \ = \ \sqrt{x^2+y^2} \ = \ \sqrt{(x+iy)(x-iy)} \ = \ \sqrt{z\bar{z}}\,.$$

Wir setzen die Konjugation fort auf $\mathbb{C}^n$ durch

$$\bar{z} \ = \ \overline{\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}} \ := \ \begin{pmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_n \end{pmatrix}$$

für $z \in \mathbb{C}^n$.

Definition 7.3.2. Das *hermitesche Skalarprodukt* auf $\mathbb{C}^n$ ist definiert durch

$$\langle z, w\rangle \ = \ \langle \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \rangle \ := \ z^{\mathrm{tr}}\bar{w} \ = \ z_1\,\bar{w}_1 + \cdots + z_n\,\bar{w}_n\,.$$

Proposition 7.3.3. *Das hermitesche Skalarprodukt ist $\mathbb{C}$-semi-bilinear, hermitesch und positiv definit, das heißt es gilt für alle $\alpha, \beta \in \mathbb{C}$ und $z, z', w, w' \in \mathbb{C}^n$:*

$$\left\{ \begin{array}{lcl} \langle \alpha\, z + \beta z', w\rangle & = & \alpha\langle z, w\rangle + \beta\langle z', w\rangle \\ \langle z, \alpha w + \beta w'\rangle & = & \bar{\alpha}\langle z, w\rangle + \bar{\beta}\langle z, w'\rangle \end{array} \right. \tag{27}$$

$$\langle z, w\rangle \ = \ \overline{\langle w, z\rangle} \tag{28}$$

$$\left\{ \begin{array}{lcl} \langle z, z\rangle & \geq & 0 \\ \langle z, z\rangle & = & 0 \quad \textit{falls } z = 0 \end{array} \right. \tag{29}$$

Das hermitsche Skalarprodukt ist zwar *nicht* symmetrisch, aber immerhin noch reflexiv.

Bemerkung 7.3.4. Allgemein werden nicht ausgeartete, reflexive Semi-Bilinearformen als *Skalarprodukte* bezeichnet. Semi-Bilinearformen heißen auch *Sesquilinearformen.*

7.4. Euklidische und unitäre Räume

Definition 7.4.1. Sei V ein $\mathbb{R}$-Vektorraum mit einer symmetrischen, positiv-definiten $\mathbb{R}$-Bilinearform $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$. Dann heißt $(V, \langle\cdot,\cdot\rangle)$ *euklidischer Raum.*

Definition 7.4.2. Sei V ein $\mathbb{C}$-Vektorraum mit einer hermiteschen, positiv definiten $\mathbb{C}$-Semi-Bilinearform $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$. Dann heißt $(V, \langle\cdot,\cdot\rangle)$ *unitärer Raum.*

Beispiel 7.4.3. Außer $(\mathbb{R}^n, \langle\cdot,\rangle)$ ist auch jeder Teilraum mit dem eingeschränkten euklidischen Skalarprodukt ein euklidischer Raum.

Analog für Teilräume von $\mathbb{C}^n$.

Sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Raum. Insbesondere ist V ein komplexer Vektorraum mit Skalarmultiplikation

$$\mathbb{C} \times V \to V : (\lambda, v) \mapsto \lambda \cdot v.$$

Diese läßt sich einschränken auf reelle Skalare:

$$\mathbb{R} \times V \to V : (\lambda, v) \mapsto \lambda v$$

und man erhält einen reellen Vektorraum $V_{\mathbb{R}}$ Weiter ist dann

$$(v, w) \;:=\; \mathrm{Re}\langle v, w\rangle \;=\; \frac{1}{2}\left(\langle v, w\rangle + \langle w, v\rangle\right) \;\overset{(28)}{=}\; \frac{1}{2}\left(\langle v, w\rangle + \overline{\langle v, w\rangle}\right)$$

eine $\mathbb{R}$-Bilinearform auf $V_{\mathbb{R}}$, die symmetrisch und wiederum positiv definit ist, das heißt $(V_{\mathbb{R}}, (\cdot,\cdot))$ ist ein euklidischer Raum.

Definition 7.4.4. Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum. Dann definiert

$$\|v\| \;:=\; \sqrt{\langle v, v\rangle}$$

die *Norm* von $v \in V$. Ferner heißen $v, w \in V$ *orthogonal*, falls gilt $\langle v, w\rangle = 0$. In diesem Fall schreiben wir $v \perp w$. Vektoren der Norm 1 heißen *Einheitsvektoren*.

Die Orthogonalitätsrelation ist symmetrisch, weil das euklidische und das hermitesche Skalarprodukt reflexiv sind.

Weil wir im Weiteren den euklidischen und den unitären Fall weitgehend parallel behandeln wollen, ist es bequem stets $\mathbb{K}$ für den Koeffizientenkörper zu schreiben. Das heißt $\mathbb{K} = \mathbb{R}$, falls V euklidisch ist, und $\mathbb{K} = \mathbb{C}$ im unitären Fall. Dazu beachte, dass $\bar{r} = r$ für $r \in \mathbb{R}$. Damit werden die Bilinearität (24) und die Symmetrie (25) zu Spezialfällen der Semi-Bilinearität (27) bzw. der Hermitizität (28).

Bemerkung 7.4.5. Für $\alpha \in \mathbb{K}$ und $v \in V$ gilt

$$\|\alpha v\| \;=\; \sqrt{\langle \alpha v, \alpha v\rangle} \;=\; \sqrt{\alpha\bar{\alpha}} \cdot \sqrt{\langle v, v\rangle} \;=\; |\alpha| \cdot \|v\|\,.$$

Bemerkung 7.4.6. Es gelten die *Polarisierungsidentitäten*

(i) im euklidischen Fall:

$$\langle x, y\rangle \;=\; \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$$

(ii) und im unitären Fall:

$$\langle x, y\rangle \;=\; \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\right).$$

Das heißt sowohl im euklidischen als auch im unitären Raum ist das Skalarprodukt durch die Norm bestimmt.

7.5. Geometrische Eigenschaften euklidischer und unitärer Räume.

Im Folgenden werden die Beweise stets simultan für den euklidischen und den unitären Fall durchgeführt. Sei nun $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum. Der Koeffizientenkörper wird wieder mit $\mathbb{K}$ bezeichnet.

Satz 7.5.1 (Ungleichung von Cauchy-Schwarz). *Für $x, y \in V$ gilt*

$$|\langle x, y\rangle| \;\leq\; \|x\| \cdot \|y\|$$

Beweis. Da die Abbildung $x \mapsto x^2$ monoton auf $\mathbb{R}_{\geq 0}$ ist, genügt es zu zeigen $|\langle x, y\rangle|^2 \leq \|x\|^2\,\|y\|^2$. Damit ist die Behauptung äquivalent zur Ungleichung

$$\langle x, y\rangle \cdot \overline{\langle x, y\rangle} \;=\; \langle x, y\rangle \cdot \langle y, x\rangle \;\leq\; \langle x, x\rangle \cdot \langle y, y\rangle\,.$$

Wir unterscheiden zwei Fälle. Angenommen es gilt $y = 0$. Dann ist $\langle x, y\rangle \cdot \langle y, x\rangle = 0 = \langle x, x\rangle\langle y, y\rangle$. Sei also nun $y \neq 0$. Aus der Positivität folgt dann $\langle y, y\rangle > 0$. Setze $\gamma := \langle x, y\rangle / \langle y, y\rangle \in \mathbb{C}$. Dann gilt

$$\begin{aligned}
0 &\overset{(29)}{\leq} \langle x - \gamma y, x - \gamma y\rangle \\
&\overset{(29)}{=} \langle x, x\rangle - \langle x, \gamma y\rangle - \langle \gamma y, x\rangle + \langle \gamma y, \gamma y\rangle \\
&\overset{(29)}{=} \langle x, x\rangle - \bar{\gamma}\langle x, y\rangle - \gamma\langle y, x\rangle + \gamma\bar{\gamma}\langle y, y\rangle \\
&= \langle x, x\rangle - \bar{\gamma}\gamma\langle y, y\rangle - \gamma\langle y, x\rangle + \gamma\bar{\gamma}\langle y, y\rangle \\
&= \langle x, x\rangle - \gamma\langle y, x\rangle\,.
\end{aligned}$$

Also folgt $\gamma\langle y, x\rangle \leq \langle x, x\rangle$, das heißt

$$\frac{\langle x, y\rangle \cdot \langle y, x\rangle}{\langle y, y\rangle} \;\leq\; \langle x, x\rangle\,.$$

□

Definition 7.5.2. Zu $x, y \in V\backslash\{0\}$ sei der *Winkel* $\gamma \in [0, \pi]$ zwischen x und y definiert durch

$$\cos\gamma \;=\; \frac{\operatorname{Re}\langle x, y\rangle}{\|x\| \cdot \|y\|}.$$

Bemerkung 7.5.3. Aus der Cauchy-Schwarz-Ungleichung 7.5.1 folgt, dass

$$-1 \;\leq\; \frac{\operatorname{Re}\langle x, y\rangle}{\|x\|\,\|y\|} \;\leq\; 1$$

ist. Damit ist $\gamma \in [0, \pi]$ eindeutig festgelegt. Insbesondere gilt $x \perp y$ genau dann, wenn $\cos\gamma = 0$, also $\gamma = \pi/2$ ist.

— Bild —

Satz 7.5.4 (Satz des Pythagoras). *Für $x, y \in V$ mit $x \perp y$ ist $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Beweis. Sei $x \perp y$, das heißt $\langle x, y\rangle = 0 = \langle y, x\rangle$. Weiter gilt wegen (27), dass

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \;=\; \|x\|^2 + \|y\|^2
\end{aligned}$$

ist. □

7.6. Metrische Räume

DEFINITION 7.6.1. Sei M eine beliebige Menge. Eine Abbildung

$$\delta \,:\, M \times M \to \mathbb{R}_{\geq 0}$$

heißt *Metrik*, falls gilt für alle $x, y, z \in M$:

$$\begin{aligned} \delta(x,y) &= \delta(x,y) \\ \delta(x,y) &\geq 0 \quad \text{und} \quad \delta(x,y) = 0 \Leftrightarrow x = y \\ \delta(x,z) &\leq \delta(x,y) + \delta(y,z) \end{aligned}$$

In diesem Fall heißt das Paar (M, δ) *metrischer Raum.*

SATZ 7.6.2. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum. Dann definiert*

$$\delta(x,y) \;:=\; \|x-y\| \;=\; \sqrt{\langle x-y, x-y\rangle}$$

eine Metrik auf V.

BEWEIS. Die Symmetrie von δ ergibt sich aus

$$\|y-x\| \;=\; \|(-1)(x-y)\| \;=\; |1| \cdot \|x-y\| \;=\; \|x-y\|\,,$$

und die positive Definitheit folgt aus (29). Weiter ergibt sich

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &\overset{7.5.1}{\leq} \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

Aus der Monotonie von $x \mapsto x^2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ folgt damit die Dreiecksungleichung. □

PROPOSITION 7.6.3. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum. Dann gilt:*

(i) $\|x+y\| = \|x\| + \|y\| \;\Leftrightarrow\;$ *es existiert $\alpha \in \mathbb{R}_{>0}$ mit $y = \alpha x$ sowie*
(ii) $|\langle x,y\rangle| = \|x\| \cdot \|y\| \;\Leftrightarrow\;$ *x und y sind linear abhängig.*

BEWEIS. Übungsaufgabe. □

LEMMA 7.6.4. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Raum mit zugehöriger Metrik δ. Für alle $x, y \in V$ existiert $m_{x,y} \in V$ eindeutig mit*

$$\delta(x, m_{x,y}) \;=\; \delta(y, m_{x,y}) \;=\; \frac{1}{2}\, d(x,y)\,.$$

Der Punkt $m_{x,y}$ heißt Mittelpunkt *zwischen x und y.*

BEWEIS. Übungsaufgabe. □

SATZ 7.6.5. *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Raum mit zugehöriger Metrik δ. Sei ferner $\varphi : V \to V$ eine Abbildung mit $\varphi(0) = 0$ und*

$$\delta(\varphi(x), \varphi(y)) \;=\; d(x,y)$$

für alle $x, y \in V$, also eine (abstandserhaltende) Isometrie. *Dann ist φ linear.*

BEWEIS. Übungsaufgabe. Man zeige der Reihe nach:

▷ $\varphi(m_{x,y}) \;=\; m_{\varphi(x),\varphi(y)}$ für alle $x, y \in V$,
▷ $\varphi(2x) = 2\varphi(x)$,

▷ $\varphi(x+y) = \varphi(x) + \varphi(y)$ und schließlich
▷ $\varphi(\alpha x + \beta y) = \alpha\varphi(x) + \beta\varphi(y)$ für alle $x, y \in V$ und alle $\alpha, \beta \in \mathbb{R}$.

□

7.7. Weitere Beispiele

7.7.1. Bilinearformen und Matrizen. Sei K eine beliebiger Körper und f eine K-Bilinearform auf dem K-Vektorraum V. Wir setzen hier voraus, dass $\dim_K V = n < \infty$ gilt. So erhalten wir nach Wahl einer Basis $\mathcal{B} = (b_1, b_2, \dots, b_n)$ von V eine Matrix

$$[f]_{\mathcal{B}} \;=\; (f_{ij})_{i,j} \;\in\; K^{n\times n}$$

mit $f_{ij} = f(b_i, b_j)$. Für alle $u, v \in V$ gilt

$$f(u,v) \;=\; [u]_{\mathcal{B}}{}^{\mathrm{tr}} \cdot [f]_{\mathcal{B}} \cdot [v]_{\mathcal{B}}\,. \tag{30}$$

Umgekehrt definiert jede $n \times n$-Matrix via (30) eine K-Bilinearform auf K^n. Eine Matrix $M \in K^{n\times n}$ heißt *symmetrisch*, falls $M^{\mathrm{tr}} = M$ ist. Offenbar ist f genau dann eine symmetrische Bilinearform, wenn die Matrix $[f]_{\mathcal{B}}$ symmetrisch ist.

Beispiel 7.7.1. Die symmetrische Matrix $H = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \in K^{2\times 2}$ definiert die K-Bilinearform h auf K^2 mit

$$h(u,v) \;=\; (u_1, u_2) \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \;=\; (u_2, u_1) \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \;=\; u_2 v_1 + u_1 v_2\,.$$

Betrachtet man speziell $K = \mathbb{R}$, so ist die symmetrische Bilinearform h *nicht* positiv definit, zum Beispiel gilt

$$h(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}) \;=\; 1\cdot 1 + 1 \cdot 1 \;=\; 2\,,$$

$$h(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}) \;=\; 0\cdot 1 + 1 \cdot 0 \;=\; 0$$

und

$$h(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}) \;=\; (-1)\cdot 1 + 1 \cdot (-1) \;=\; -2\,.$$

Beispiel 7.7.2. Für $m, n \in \mathbb{N}$ definiert die relle Diagonalmatrix

$$\operatorname{diag}(\underbrace{1, 1, \dots, 1}_{m}, \underbrace{-1, -1, \dots, -1}_{n})$$

eine $\mathbb{R}$-Bilinearform $f_{m,n}$ auf $\mathbb{R}^{m+n}$ mit

$$f_{m,n}(u,v) \;=\; u_1 v_1 + u_2 v_2 + \cdots + u_m v_m - u_{m+1} v_{m+1} - u_{m+2} v_{m+2} - \cdots - u_{m+n} v_{m+n}\,.$$

Diese symmetrische Bilinearform ist genau dann positiv definit, wenn $n = 0$ ist. Die Bilinearform $f_{m,0}$ ist das euklidische Skalarprodukt auf $\mathbb{R}^m$. Für $m = 0$ ist die Bilinearform *negativ definit*.

Analoges gilt für (komplexe) Semi-Bilinearformen. Eine Matrix $M \in \mathbb{C}^{n\times n}$ heißt *hermitesch*, falls gilt $M^{\mathrm{tr}} = \overline{M}$. Jede hermitesche Matrix definiert eine hermitesche Semi-Bilinearform f auf $\mathbb{C}^n$ via

$$f(u,v) \;=\; u^{\mathrm{tr}} \cdot M \cdot \bar{v}\,,$$

denn es gilt

$$f(v,u) \;=\; v^{\mathrm{tr}} \cdot M \cdot \bar{u} \;=\; \bar{u}^{\mathrm{tr}} \cdot M^{\mathrm{tr}} \cdot \bar{v} \;=\; \bar{u}^{\mathrm{tr}} \cdot \overline{M} \cdot \bar{v} \;=\; \overline{u^{\mathrm{tr}} \cdot M \cdot \bar{v}} \;=\; \overline{f(u,v)}\,.$$

Jede hermitesche Semi-Bilinearform auf $\mathbb{C}^n$ entsteht auf diese Weise.

7.7.2. Bilinearformen auf Funktionenräumen.

Beispiel 7.7.3. Sei $V = \mathcal{C}[0,1]$ die Menge aller stetigen Funktionen $f : [0,1] \to \mathbb{R}$. Via

$$(f+g)(x) \;=\; f(x)+g(x) \qquad \text{für } f,g \in V$$

und

$$(\lambda f)(x) \;=\; \lambda f(x) \qquad \text{für } \lambda \in \mathbb{R}$$

ist V ein $\mathbb{R}$-Vektorraum. Wir werden später zeigen, dass V unendlich-dimensional ist. Setze

$$\langle f,g\rangle \;:=\; \int_0^1 f(x)\,g(x)\,dx\,.$$

Die Bilinearität folgt aus den Rechenregeln für Integrale, die Symmetrie ergibt sich aus der Kommutativität der Multiplikation auf $\mathbb{R}$.

Proposition 7.7.4. *Das Paar* $(\mathcal{C}[0,1], \langle\cdot,\cdot\rangle)$ *ist ein euklidischer Raum.*

Beweis. Es bleibt, die positive Definitheit zu zeigen. Dazu sei $f \in \mathcal{C}[0,1]$ beliebig. Dann ist die Funktion $f^2 : [0,1] \to \mathbb{R} : x \mapsto f(x)^2$ punktweise nicht negativ, weshalb

$$\langle f,f\rangle \;=\; \int_0^1 f^2(x)\,dx \;\geq\; 0$$

folgt. Außerdem ist für $\langle f,f\rangle = \int_0^1 f^2(x)\,dx = 0$ auch $f^2 = 0 = f$, da f stetig ist. □

Bemerkung 7.7.5. Es folgen die Cauchy-Schwarzsche Ungleichung etc. Die zugehörige Norm

$$\|f\| \;=\; \sqrt{\langle f,f\rangle} \;=\; \left(\int_0^1 f^2(x)\,dx\right)^{1/2}$$

ist ein Spezialfall der *L^p-Norm*

$$\|f\|_p \;=\; \left(\int_0^1 |f|^p(x)\,dx\right)^{1/p} \qquad \text{mit } 1 \leq p < \infty$$

auf $\mathcal{C}[0,1]$ für $p=2$.

7.8. Orthonormalbasen

Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum. Wir lassen in diesem Abschnitt ausdrücklich zu, dass V unendliche Dimension über $\mathbb{K}$ hat.

Definition 7.8.1. Eine Familie $(v_1,\dots,v_m)$ in $V\setminus\{0\}$ heißt *Orthogonalsystem*, falls gilt

$$v_i \perp v_j \quad \text{für} \quad i \neq j\,.$$

Gilt zusätzlich $\|v_i\| = 1$, das heißt v_i ist Einheitsvektor, dann heißt $(v_1,\dots,v_m)$ *Orthonormalsystem*. Ein Orthonormalsystem, das eine Basis (von V) ist, heißt *Orthonormalbasis*.

Lemma 7.8.2. *Jedes Orthogonalsystem* $(v_1,\dots,v_m)$ *ist linear unabhängig.*

Beweis. Angenommen, es existieren $\lambda_1,\dots,\lambda_m \in \mathbb{K}$, so dass $\sum_{i=1}^m \lambda_i v_i = 0$. Dann gilt für jedes $k \in \{1,\dots,m\}$, dass

$$0 \;=\; \langle \sum_{i=1}^m \lambda_i v_i, v_k\rangle \;=\; \sum_{i=1}^m \lambda_i \langle v_i, v_k\rangle \;=\; \lambda_k \langle v_k, v_k\rangle$$

ist, woraus wegen $v_k \neq 0$ weiter $\lambda_k = 0$ folgt. □

Beispiel 7.8.3.

(i) Die Standardbasen des $\mathbb{R}^n$ und $\mathbb{C}^n$ sind Orthonormalbasen,
(ii) Für beliebiges $\gamma \in \mathbb{R}$ ist

$$\left(\begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix}, \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}\right)$$

eine Orthonormalbasis von $\mathbb{R}^2$.

Bemerkung 7.8.4 (Koordinaten bezüglich einer Orthonormalbasis). Sei $v_1, \ldots, v_n$ eine Orthonormalbasis von V. Dann lässt sich ein beliebiges $v \in V$ eindeutig darstellen als

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n, \qquad \text{für } \lambda_i \in \mathbb{K}$$

Dies bedeutet

$$\langle v, v_i\rangle = \langle \sum \lambda_j v_j, v_i\rangle = \sum \lambda_j \langle v_j, v_i\rangle = \lambda_j\langle v_j, v_i\rangle = \lambda_i .$$

Anders ausgedrückt ist

$$v = \langle v, v_i\rangle\, v_1 + \cdots + \langle v, v_n\rangle\, v_n .$$

Beispiel 7.8.5. Betrachte die Orthonormalbasis

$$\left(\begin{pmatrix} \cos\pi/4 \\ \sin\pi/4 \end{pmatrix}, \begin{pmatrix} -\sin\pi/4 \\ \cos\pi/4 \end{pmatrix}\right)$$

von $\mathbb{R}^2$. Dabei ist $\cos\pi/4 = \sin\pi/4 = \frac{1}{2}\sqrt{2}$. Für $v = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ gilt:

$$\begin{aligned}
\langle v, \begin{pmatrix} \cos\pi/4 \\ \sin\pi/4 \end{pmatrix}\rangle &= \langle \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{pmatrix}\rangle &&= -\tfrac{1}{2}\sqrt{2} + \tfrac{3}{2}\sqrt{2} &&= \sqrt{2} \\
\langle v, \begin{pmatrix} -\sin\pi/4 \\ \cos\pi/4 \end{pmatrix}\rangle &= \langle \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{pmatrix}\rangle &&= \tfrac{1}{2}\sqrt{2} + \tfrac{3}{2}\sqrt{2} &&= 2\sqrt{2}
\end{aligned}$$

Also ist

$$v = \sqrt{2}\begin{pmatrix} \cos\pi/4 \\ \sin\pi/4 \end{pmatrix} + 2\sqrt{2}\begin{pmatrix} -\sin\pi/4 \\ \cos\pi/4 \end{pmatrix}.$$

7.9. Trigonometrische Polynome

Lemma 7.9.1. *Für $m, n \in \mathbb{N}$ gilt*

$$\int_0^1 \sin(2\pi\, m\, x)\sin(2\pi\, n\, x)\, dx = \begin{cases} 0 & \textit{für}\quad m \neq n \\ 1/2 & \textit{für}\quad m = n > 0 \\ 0 & \textit{für}\quad m = n = 0\,. \end{cases}$$

Beweis. Wir betrachten den Fall $m \neq n$. Partielle Integration liefert

$$\begin{aligned}
&\int_0^1 \sin(2\pi\, m\, x)\ \sin(2\pi\, n\, x)\, dx \\
&\qquad = \frac{1}{4\pi}\left(\frac{\sin(2\pi(m-n)x)}{m-n} - \frac{\sin(2\pi(m+n)x)}{m+n}\right)\Bigg|_0^1 = 0\,.
\end{aligned}$$

Für $m = n = 0$ ergibt sich

$$\int_0^1 \sin^2(0) dx = 0\,,$$

und für $m = n > 0$ gilt

$$\int_0^1 \sin^2(2\pi\, m\, x)\, dx \;=\; \frac{2\pi\, m\, x - \sin(2\pi\, m\, x)\cos(2\pi\, m\, x)}{4\pi\, m}\Bigg|_0^1 \;=\; \frac{2\pi\, m}{4\pi\, m} \;=\; \frac{1}{2}\,.$$

□

Analog beweist man die beiden folgenden Lemmata.

LEMMA 7.9.2. *Für $m, n \in \mathbb{N}$ gilt*

$$\int_0^1 \sin(2\pi\, m\, x)\ \cos(2\pi\, n\, x)\, dx \;=\; 0\,.$$

LEMMA 7.9.3. *Für $m, n \in \mathbb{N}$ gilt*

$$\int_0^1 \cos(2\pi\, m\, x)\ \cos(2\pi\, n\, x)\, dx \;=\; \begin{cases} 0 & \text{für} \quad m \neq n \\ 1/2 & \text{für} \quad n = m > 0 \\ 1 & \text{für} \quad m = n = 0\,. \end{cases}$$

Die vorstehenden Lemmata zusammen genommen ergeben das Folgende.

PROPOSITION 7.9.4. *Die Funktionen $v_0, v_1, \ldots, w_1, w_2, \ldots,$ die für $x \in [0,1]$ definiert sind durch*

$$v_k(x) \;=\; \cos(2\pi\, k\, x) \quad \text{und} \quad w_k(x) \;=\; \sin(2\pi\, k\, x),$$

für $k \in \mathbb{N}$, bilden ein Orthogonalsystem in $\mathcal{C}[0,1]$.

Durch Skalierung (mit $\sqrt{2}$ für $k \geq 1$) erhalten wir ein Orthonormalsystem von Funktionen

$$\tilde{v}_0, \tilde{v}_1, \ldots, \tilde{w}_1, \tilde{w}_2, \ldots$$

Aus Lemma 7.8.2 folgt, dass $\mathcal{C}[0,1]$ unendliche Dimension über $\mathbb{R}$ hat.

DEFINITION 7.9.5. Der lineare Teilraum

$$\mathcal{T}_n \;:=\; \operatorname{lin}_{\mathbb{R}}(v_0, \ldots, v_n, w_1, \ldots, w_n)$$

von $\mathcal{C}[0,1]$ heißt Raum der *trigonometrischen Polynome* vom Grad $\leq n$.

Die Elemente von $\mathcal{T}_n$ bestehen aus Linearkombinationen

$$T \;=\; \alpha_0\, v_0 + \sum_{k=1}^{n} (\alpha_k\, v_k + \beta_k\, w_k)\,.$$

Das heißt, es ist

$$T(x) \;=\; \alpha_0 + \sum_{k=1}^{n} (\alpha_k\, \cos(2\pi\, k\, x) + \beta_k\, \sin(2\pi\, k\, x))\,.$$

7.10. Orthonormalisierungsverfahren nach Gram-Schmidt

Sei $(b_1, \dots, b_m)$ eine linear unabhängige Familie in V. Wir setzen $U := \mathrm{lin}_K(b_1, \dots, b_m)$. Es gilt $\dim_{\mathbb{K}} U = m$. Im Folgenden konstruieren wir eine Orthonormalbasis für U. Dazu setze

$$\begin{array}{lclcl} u_1 & := & b_1 & \text{und} & v_1 := \frac{u_1}{\|u_1\|} \\ u_2 & := & b_2 - \langle b_2, v_2 \rangle v_1 & \text{und} & v_2 := \frac{u_2}{\|u_2\|} \\ u_3 & := & b_3 - \langle b_3, v_1 \rangle v_1 - \langle b_3, v_2 \rangle v_2 & \text{und} & v_3 := \frac{u_3}{\|u_3\|} \\ & \vdots & & & \vdots \\ u_m & := & b_m - \sum_{i=1}^{m-1} \langle b_m, v_i \rangle v_i & \text{und} & v_m := \frac{u_m}{\|u_m\|} \end{array}$$

Satz 7.10.1. *Die Familie* $(v_1, \dots, v_k)$ *ist eine Orthonormalbasis für* $\mathrm{lin}(b_1, \dots, b_k)$ *für alle* $k \in \{1, \dots, m\}$.

Beweis. Per Induktion nach k. Der *Induktionsanfang* für $k = 1$ ist klar.

Induktionsschritt: Wir nehmen an, dass $(v_1, \dots, v_i)$ eine Orthonormalbasis ist für $\mathrm{lin}(b_1, \dots, b_i)$. Dann ist

$$u_{i+1} = b_{i+1} - \sum_{j=1}^{i} \langle b_{i+1}, v_j \rangle v_j$$

und es gilt $b_{i+1} \notin \mathrm{lin}(b_1, \dots, b_i) = \mathrm{lin}(v_1, \dots, v_i)$. Also folgt, dass $u_{i+1} \neq 0$ ist und damit auch $\|u_{i+1}\| \neq 0$; der Vektor $v_{i+1} = \frac{u_{i+1}}{\|u_{i+1}\|}$ ist definiert. Außerdem ist

$$b_{i+1} = u_{i+1} + \sum_{i=1}^{i} \langle b_{i+1}, v_j \rangle v_i \, .$$

Es ergibt sich $b_{i+1} \in \mathrm{lin}(v_1, \dots, v_i, u_{i+1}) = \mathrm{lin}(v_1, \dots, v_{i+1})$, das heißt also $\mathrm{lin}(b_1, \dots, b_{i+1}) = \mathrm{lin}(v_1, \dots, v_{i+1})$. Wir berechnen für $k < i + 1$:

$$\begin{aligned} \langle u_{i+1}, v_k \rangle &= \langle b_{i+1} - \sum_{j=1}^{i} \langle b_{i+1}, v_j \rangle v_j, v_k \rangle \\ &= \langle b_{i+1}, v_k \rangle - \sum_{j=1}^{i} \langle b_{i+1}, v_j \rangle \langle v_j, v_k \rangle \\ &= \langle b_{i+1}, v_k \rangle - \langle b_{i+1}, v_k \rangle = 0 \, . \end{aligned}$$

Dies bedeutet

$$\langle v_{i+1}, v_k \rangle = \langle \frac{u_{i+1}}{\|u_{i+1}\|}, v_k \rangle = \frac{1}{\|u_{i+1}\|} \langle u_{i+1}, v_k \rangle = 0 \, .$$

□

Korollar 7.10.2. *Jeder endlich-dimensionale Teilraum von* V *besitzt eine Orthonormalbasis.*

7.11. Orthogonale Teilräume

Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Vektorraum. Wiederum wird $\dim_{\mathbb{K}} V = \infty$ zugelassen.

Definition 7.11.1. Zu $M \subseteq V$ setze

$$M^\perp \ := \ \{v \in V : \langle v, m\rangle = 0 \text{ für alle } m \in M\} \ .$$

Lemma 7.11.2. *$M^\perp$ ist linearer Teilraum von V.*

Beweis. Übungsaufgabe. □

Lemma 7.11.3. *Seien $A, B \subseteq V$. Dann gilt:*

(i) $A \subseteq B \Rightarrow B^\perp \subseteq A^\perp$
(ii) $A \subseteq B^\perp \Leftrightarrow B \subseteq A^\perp$
(iii) $A \subseteq (A^\perp)^\perp$
(iv) $A^\perp = ((A^\perp)^\perp)^\perp$.

Beweis. Wir zeigen die erste Eigenschaft. Sei also $A \subseteq B$ und $u \in B^\perp$. Dann steht u auf sämtlichen Elementen aus B senkrecht, also auch auf denen aus A.
Für die zweite Eigenschaft sei $A \subseteq B^\perp$ und $v \in B$. Also gilt für alle $u \in A$, dass $u \perp v$ ist. Das heißt $v \in A^\perp$. Die zweite Implikation ist symmetrisch.
Die beiden übrigen Eigenschaften sind Übungsaufgaben. □

Bemerkung 7.11.4. Wir betrachten kurz den endlich-dimensionalen Spezialfall $V = \mathbb{K}^n$. Für $a \in V$ ist

$$a^\perp \ := \ \{a\}^\perp \ = \ \{x \in \mathbb{K}^n : \langle a, x\rangle = 0\} \ .$$

Mit

$$a \ = \ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad x \ = \ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

gilt $\langle a, x\rangle = a_1\,x_1 + \cdots + a_n\,x_n$. Das heißt $a^\perp$ ist genau die Lösungsmenge der linearen Gleichung $a_1\,x_1 + \cdots + a_n\,x_n = 0$. Für $a \neq 0$ ist $a^\perp$ also eine lineare Hyperebene. Weiter gilt für $a, b, c, \ldots \in \mathbb{K}^n$, dass

$$\{a, b, c, \ldots\}^\perp \ = \ a^\perp \cap b^\perp \cap c^\perp \cap \cdots$$

die Lösungsmenge des homogenen linearen Gleichungssystems

$$\langle a, x\rangle = 0, \ \langle b, x\rangle = 0, \ \langle c, x\rangle = 0, \ldots$$

ist.

Sei V jetzt wieder ein beliebiger euklidischer oder unitärer Raum.

Lemma 7.11.5. *Sei $a_1, \ldots, a_k \in V$ und $U := \mathrm{lin}_{\mathbb{K}}(a_1, \ldots, a_k)$. Dann gilt $U^\perp = \{a_1, \ldots, a_k\}^\perp$.*

Beweis. Wegen Lemma 7.11.3(i) ist $U^\perp \subseteq \{a_1, \ldots, a_k\}^\perp$. Sei umgekehrt v enthalten in $\{a_1, \ldots, a_k\}^\perp$. Dann gilt

$$\langle v, a_1\rangle \ = \ \ldots \ = \ \langle v, a_k\rangle \ = \ 0\,.$$

Sei $u \in U$ beliebig. Dann existieren $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$, so dass $u = \lambda_1\,a_1 + \cdots + \lambda_m\,a_m$. Man erhält

$$\langle v, u\rangle \ = \ \sum_{i=1}^{k} \lambda_i \langle v, a_i\rangle \ = \ 0\,,$$

das heißt $v \in u^\perp$. Da dies aber für alle $u \in U$ gilt, folgt die Behauptung. □

Satz 7.11.6. *Sei $U \leq V$ ein endlich-dimensionaler Teilraum mit der Orthonormalbasis $(u_1, u_2 \ldots, u_k)$. Für die Abbildung*

$$\pi \,:\, V \to V \,:\, v \mapsto \sum_{i=1}^{k} \langle v, u_i \rangle \, u_i$$

gelten die folgenden Eigenschaften:

(i) π *ist linear,*
(ii) $\pi(v) \in U$ *für alle* $v \in V$*,*
(iii) $\pi(u) = u$ *für alle* $u \in U$
(iv) $\pi \circ \pi = \pi$*,*
(v) $\operatorname{img} \pi = U$ *und* $\ker \pi = U^\perp$*,*
(vi) $v - \pi(v) \in U^\perp$ *für alle* $v \in V$*,*
(vii) $\|v - \pi(v)\| \leq \|v - u\|$ *für alle* $v \in V$ *und* $u \in U$*; Gleichheit gilt nur für* $u = \pi(u)$*.*

Beweis. Die Linearität von π ist klar (Spezialfall von Beispiel 3.1.3). Das Bild liegt stets in U, weil U ein Teilraum ist und π linear.

Sei $u \in U$. Dann existieren $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ mit $u = \lambda_1 u_1 + \lambda_2 u_2 + \ldots \lambda_k u_k$. Aus Bemerkung 7.8.4 folgt, dass $\langle u, u_i \rangle = \lambda_i$ ist für alle i. Damit ist

$$\pi(u) \;=\; \sum_{i=1}^{k} \langle u, u_i \rangle u_i \;=\; \sum_{i=1}^{k} \lambda_i u_i \;=\; u\,. \tag{31}$$

Da für $v \in V$ gilt, dass $\pi(v) \in U$ ist, folgt $\pi(\pi(v)) = \pi(v)$ aus (31).

Die Surjektivität von π auf den Unterraum U folgt ebenfalls aus (31). Sei $v \in \ker \pi$. Da die Vektoren $u_1, u_2, \ldots, u_k$ linear unabhängig sind, folgt aus

$$0 \;= \pi(v) \;=\; \sum_{i=1}^{k} \langle u, u_i \rangle u_i$$

unmittelbar $\langle u, u_i \rangle = 0$. Also $v \in u_1^\perp \cap u_2^\perp \cap \cdots \cap u_k^\perp = U^\perp$. Diese Argumente lassen sich alle umkehren, so dass auch $U^\perp \subseteq \ker \pi$ folgt.

Sei $v \in V$ beliebig. dann gilt $\pi(v - \pi(v)) = \pi(v) - \pi(\pi(v)) = \pi(v) - \pi(v) = 0$, also $v - \pi(v) \in \ker \pi = U^\perp$.

Die verbleibende Aussage ist eine Übungsaufgabe. □

Die Abbildung π heißt *Orthogonalprojektion* von V auf U.

Bemerkung 7.11.7. Aus Satz 7.11.6(vii) folgt, dass π nicht von der Wahl der Orthonormalbasis $(u_1, \ldots, u_m)$ abhängt.

7.12. Fouriertransformation

Es sei $\pi_n : \mathcal{C}[0,1] \to \mathcal{T}_n$ die orthogonale Projektion auf dem Raum der trigonometrischen Polynome vom Grad $\leq n$. Nach Satz 7.11.6 ist $\pi_n(f)$ die zu $f \in \mathcal{C}[0,1]$ bezüglich der

L^2-Norm beste Approximation durch ein trigonometrisches Polynom vom Grad $\leq n$. Es gilt:

$$\begin{aligned}\pi_n(f) &= \langle f, \tilde{v}_0\rangle\, \tilde{v}_0 + \sum_{k=1}^{n} \left(\langle f, \tilde{v}_k\rangle\, \tilde{v}_k + \langle f, \tilde{w}_k\rangle\, \tilde{w}_k\right) \\ &= \frac{\langle f, v_0\rangle}{\|v_0\|^2}\, v_0 + \dots \\ &= \alpha_0 + \sum_{k=1}^{n} \left(\alpha_k \cos(2\pi\, k\, x) + \beta_k \sin(2\pi\, k\, x)\right)\end{aligned}$$

mit

$$\alpha_0 = \int_0^1 f(x)\, dx\,, \quad \alpha_k = \sqrt{2} \int_0^1 f(x) \cos(2\pi\, k\, x)\, dx \ \text{ für } \ k > 0\,,$$

und

$$\beta_k = \sqrt{2} \int_0^1 f(x) \sin(2\pi\, k\, x)\, dx \ \text{ für } \ k > 0\,.$$

Definition 7.12.1. Die Koeffizienten α_k und β heißen *Fourierkoeffizienten* von f. Die unendliche Reihe

$$\alpha_0 + \sum_{k=1}^{\infty} \alpha_k \cos(2\pi\, k\, x) + \beta_k \sin(2\pi\, k\, x)$$

heißt *Fourierreihe* von f.

Satz 7.12.2. *Falls f zweimal stetig differenzierbar und periodisch ist, dann konvergiert die Fourierreihe gleichmäßig gegen f.*

Für die *Periodizität* wird verlangt, dass $f(0) = f(1)$, $f'(0) = f'(1)$ und $f''(0) = f''(1)$ gilt.

7.13. Summen von Vektorräumen

7.13.1. Innere direkte Summe von Teilräumen. Sei V ein möglicherweise unendlichdimensionaler Vektorraum über einem beliebigen Körper K.

Definition 7.13.1. Für beliebige Teilmengen $A, B \subseteq V$ heißt

$$A + B \;:=\; \{a + b :\ a \in A,\ b \in B\}$$

die *(Minkowski)-Summe* von A und B.

Lemma 7.13.2. *Falls $A, B \leq V$ Teilräume sind, so ist auch $A + B \leq V$ ein Teilraum.*

Beweis. Offenbar ist $0 \in A \cap B$, also $0 = 0 + 0 \in A + B$. Seien $a, a' \in A$, $b, b' \in B$ und $\lambda, \mu \in K$. Dann ist

$$\lambda(a + b) + \mu(a' + b') \;=\; (\lambda a + \mu a') + (\lambda b + \mu b') \;\in\; A + B\,.$$

□

Beispiel 7.13.3. Sei $K = \mathbb{R}$ und $V = \mathbb{R}^2$. Betrachte die Mengen

$$A = [(0,0),\,(2,2)]\,, \quad B = [(0,0),\,(1,0)] \ \text{ und } \ C = [(0,0),\,(0,2)]$$

und deren Minkowski-Summe $A + b + C$. Hier ist

$$[x, y] = \operatorname{conv}(x, y) = \{\lambda x + (1 - \lambda) y :\ 0 \leq \lambda \leq 1\}$$

die *Verbindungsstrecke* von x und y.

— Bild —

LEMMA 7.13.4. *Seien $A, B \leq V$ endlichdimensionale Teilräume. Dann gilt*

$$\dim(A+B) \;=\; \dim A + \dim B - \dim(A \cap B)\,.$$

BEWEIS. Sei $(c_1, c_2, \ldots, c_k)$ eine Basis des endlichdimensionalen Teilraums $A \cap B \leq V$. Nach dem Basisergänzungssatz 2.9.5 existieren $a_1, a_2, \ldots, a_m \in A$ und $b_1, b_2, \ldots, b_n \in B$, so dass $(a_1, \ldots, a_m, c_1, \ldots, c_k)$ Basis von A und $(b_1, \ldots, b_n, c_1, \ldots, c_k)$ Basis von B ist. Damit ist $(a_1, \ldots, a_m, b_1, \ldots, b_n, c_1, \ldots, c_k)$ eine Basis von $A + B$, und es folgt

$$\begin{aligned} \dim(A+B) \;&=\; m+n+k \;=\; (m+k)+(n+k)-k \\ &=\; \dim A + \dim B + \dim(A \cap B)\,. \end{aligned}$$

□

DEFINITION 7.13.5. Seien $A, B \leq V$ Teilräume mit $A \cap B = 0$. Dann heißt $A \oplus B := A + B$ die (innere) *direkte Summe* von A und B.

BEMERKUNG 7.13.6. Falls $\dim_K V < \infty$ und $V = A \oplus B$, so ist $\dim A + \dim B = \dim V$.

7.13.2. Äußere direkte Summe von Vektorräumen. Seien V und W Vektorräume über demselben Körper K. Dann ist das cartesische Produkt $V \times W$ mit komponentenweiser Addition und komponentenweiser Skalarmultiplikation ebenfalls ein K-Vektorraum. Für $V \times W$ ist auch der Begriff *äußere direkte Summe* üblich. Denn die Unterräume

$$V \times \{0\} \quad \text{und} \quad \{0\} \times W$$

sind zu V bzw. W isomorphe Teilräume von $V \times W$, und es gilt

$$(V \times \{0\}) \oplus (\{0\} \times W) \;=\; V \times W\,.$$

Deswegen schreiben wir im Folgenden auch $V \oplus W$ für $V \times W$. Die Unterscheidung, ob es sich um einen innere oder äußere direkte Summe handelt, ergibt sich aus dem Kontext.

Seien nun f und g Bilinearformen auf V bzw. W. Dann ist

$$f \oplus g \;:\; (V \oplus W) \times (V \oplus W) \;:\; (v + w, v' + w') \mapsto f(v, v') + g(w, w')$$

eine Bilinearform.

LEMMA 7.13.7. *Sind f und g beide symmetrisch/reflexiv/nicht ausgeartet, so ist auch $f \oplus g$ symmetrisch/reflexiv/nicht ausgeartet.*

Analoges gilt für (komplexe) Semi-Bilinearformen.

Im endlichdimensionalen Fall gilt Folgendes. Seien $\mathcal{B} = (b_1, b_2, \ldots, b_m)$ und $\mathcal{C} = (c_1, c_2, \ldots, c_n)$ Basen von V bzw. W. Dann ist

$$(\mathcal{B}, \mathcal{C}) \;=\; ((b_1, 0), (b_2, 0), \ldots, (b_m, 0)), (0, c_1), (0, c_2), \ldots, (0, c_n))$$

eine Basis von $V \oplus W$. Insbesondere addieren sich die Dimensionen.

LEMMA 7.13.8. *Für $A = [f]_{\mathcal{B}}$ und $B = [g]_{\mathcal{C}}$ ist*

$$[f \oplus g]_{(\mathcal{B},\mathcal{C})} \;=\; \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \;=:\; A \oplus B\,.$$

Zu $\lambda \in K$ sei (λ) die 1×1-Matrix mit einzigem Koeffizienten λ.

BEISPIEL 7.13.9. Für das Standardskalarprodukt $f : K^n \times K^n \to K$ gilt

$$[f] \;=\; \underbrace{(1) \oplus (1) \oplus \cdots \oplus (1)}_{n \text{ mal}}\,.$$

Beispiel 7.13.10. Für die Bilinearform $f_{m,n}$ auf $\mathbb{R}^{m+n}$ gilt

$$[f_{m,n}] \;=\; \underbrace{(1) \oplus (1) \oplus \cdots \oplus (1)}_{m \text{ mal}} \oplus \underbrace{(-1) \oplus (-1) \oplus \cdots \oplus (-1)}_{n \text{ mal}} \,.$$

7.13.3. Orthogonale Summe in euklidischen und unitären Räumen. Sei nun wieder $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum.

Proposition 7.13.11. *Sei $U \le V$ endlich-dimensionaler Teilraum. Dann gilt $V = U \oplus U^\perp$. Falls* $\dim_K V < \infty$*, so ist* $\dim U^\perp = \dim V - \dim U$.

Beweis. Es gilt $U \cap U^\perp = 0$, weil das Skalarprodukt positiv definit ist, also kein Vektor außer dem Nullvektor auf sich selbst senkrecht steht. Wir zeigen nun $V = U + U^\perp$. Sei π die orthogonale Projektion von V auf U. Dann ist $\pi(v) \in U$ und $v - \pi(v) \in U^\perp$ nach Satz 7.11.6 (vi). □

Korollar 7.13.12. *Sei $U \le V$ endlich-dimensionaler Teilraum. Dann gilt $(U^\perp)^\perp = U$.*

Beweis. Es gilt stets $U \subseteq (U^\perp)^\perp$. Aus Proposition 7.13.11 folgt nun $V = U \oplus U^\perp$, also

$$(U^\perp)^\perp \;=\; (U \oplus U^\perp) \cap (U^\perp)^\perp \;=\; \underbrace{(U \cap (U^\perp)^\perp)}_{U} + \underbrace{(U^\perp \cap (U^\perp)^\perp)}_{0} \;=\; U \,.$$

□

7.14. Orthogonale und unitäre Abbildungen

Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Raum.

Definition 7.14.1. Eine lineare Abbildung $g : V \to V$ heißt (im euklidischen Fall) *orthogonal* oder (im unitären Fall) *unitär*, falls gilt

$$\langle g(x), g(y)\rangle \;=\; \langle x, y\rangle \quad \text{für alle } x, y \in V \,.$$

Orthogonale Abbildungen erhalten Längen, Winkel und die Orthogonalitätsrelation, da sich diese mit Hilfe des Skalarprodukts ausdrücken lassen.

Lemma 7.14.2. *Jede orthogonale oder unitäre lineare Abbildung ist injektiv.*

Beweis. Angenommen $g : V \to V$ ist linear, aber nicht injektiv. Dann existiert $v \in \ker g \setminus \{0\}$. Es gilt also $v \ne 0$ und $g(v) = 0$. Damit folgt aus

$$0 \;=\; \langle g(v), g(v)\rangle \quad \text{und} \quad \langle v, v\rangle \;>\; 0$$

wegen der positiven Definitheit des Skalarprodukts, dass g nicht orthogonal (bzw. unitär) ist. □

Definition 7.14.3. Eine bijektive orthogonale lineare Abbildung heißt *orthogonale Transformation.* Analog werden unitäre Transformationen definiert.

Definition 7.14.4. Wir bezeichnen (im euklidischen Fall) mit

$$\mathrm{O}(V) \;:=\; \{g \in \mathrm{GL}(V) :\; g \text{ orthogonal}\}$$

die *orthogonale Gruppe* von $(V, \langle\cdot,\cdot\rangle)$ und (im unitären Fall) mit

$$\mathrm{U}(V) \;:=\; \{g \in \mathrm{GL}(V) :\; g \text{ unitär}\}$$

die *unitäre Gruppe* von $(V, \langle\cdot,\cdot\rangle)$.

Aus Lemma 7.14.2 folgt, dass jede orthogonale lineare Abbildung auf einem endlich-dimensionalen Vektorraum eine orthogonale Transformation ist.

BEMERKUNG 7.14.5. Es ist zu verifizieren, dass $\mathrm{O}(V)$ und $\mathrm{U}(V)$ tatsächlich Untergruppen von $\mathrm{GL}(V)$ sind.

BEISPIEL 7.14.6 (Spiegelung an einer linearen Hyperebene). Sei $(V, \langle\cdot,\cdot\rangle)$ euklidisch. Dann ist das Skalarprodukt symmetrisch und bilinear. Für $a \in V \setminus \{0\}$ ist

$$\langle a, \cdot\rangle \,:\, V \to \mathbb{R} \,:\, x \mapsto \langle a, x\rangle$$

eine *Linearform*, das heißt, ein Element vom $\mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$. Die Abbildung

$$s_a \,:\, V \to V \,:\, x \mapsto x - 2\frac{\langle a, x\rangle}{\langle a, a\rangle} a$$

heißt *Spiegelung* an der Hyperebene $\ker\langle a, \cdot\rangle$. Offensichtlich is s_a linear. Es gilt

$$\begin{aligned}
\langle s_a(x), s_a(y)\rangle &= \langle x - 2\frac{\langle a, x\rangle}{\langle a, a\rangle} a, y - 2\frac{\langle a, y\rangle}{\langle a, a\rangle} a\rangle \\
&= \langle x, y\rangle - 2\frac{\langle a, y\rangle}{\langle a, a\rangle}\langle x, a\rangle - 2\frac{\langle a, x\rangle}{\langle a, a\rangle}\langle y, a\rangle + 4\frac{\langle a, x\rangle\langle a, y\rangle}{\langle a, a\rangle^2}\langle a, a\rangle \\
&= \langle x, y\rangle \,.
\end{aligned}$$

Es gilt genau dann $s_a(x) = x$, wenn x im Kern der Linearform $\langle a, \cdot\rangle$ liegt. Zusätzlich ist

$$\begin{aligned}
s_a(s_a(x)) &= s_a(x - 2\frac{\langle a, x\rangle}{\langle a, a\rangle} a) \;=\; x - 2\frac{\langle a, x\rangle}{\langle a, a\rangle} a - 2\frac{\langle a, x - 2\frac{\langle a,x\rangle}{\langle a,a\rangle} a\rangle}{\langle a, a\rangle} a \\
&= x - 2\frac{\langle a, x\rangle}{\langle a, a\rangle} a - 2\frac{\langle a, x\rangle}{\langle a, a\rangle} a + 4\frac{\langle a, x\rangle\langle a, a\rangle}{\langle a, a\rangle^2} a \;=\; x \,.
\end{aligned}$$

Die Spiegelung s_a ist ein Beispiel für eine *Involution*.

Im Folgenden sei V endlich-dimensional.

BEMERKUNG 7.14.7. Aus Satz 7.6.5 und den Polarisierungsidentitäten 7.4.6 folgt, dass jede Isometrie von $\mathbb{R}^n$ eine orthogonale lineare Transformation ist.

PROPOSITION 7.14.8. *Sei $(v_1, v_2, \ldots, v_n)$ eine Orthonormalbasis von V. Eine lineare Abbildung $g : V \to V$ ist genau dann orthogonal (bzw. unitär), wenn $(g(v_1), g(v_2), \ldots, g(v_n))$ ebenfalls eine Orthonormalbasis von V ist.*

BEWEIS. Nach Voraussetzung gilt $\langle v_i, v_j\rangle = \delta_{ij}$ für alle i, j. Ist also g orthogonal (bzw. unitär), so ist $\langle g(v_i), g(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$, und $(g(v_1), g(v_2), \ldots, g(v_n))$ eine Orthonormalbasis.
Für die Umkehrung seien $x = \sum_i \lambda_i v_i$ und $y = \sum_i \mu_i v_i$ beliebig. Dann ist

$$\begin{aligned}
\langle g(x), g(y)\rangle &= \langle g(\sum_i \lambda_i v_i), g(\sum_i \mu_i v_i)\rangle \;=\; \sum_{i,j} \lambda_i \bar{\mu}_j \langle g(v_i), g(v_j)\rangle \\
&= \sum_{i,j} \lambda_i \bar{\mu}_j \langle v_i, v_j\rangle \;=\; \langle \sum_i \lambda_i v_i, \sum_i \mu_i v_i\rangle \;=\; \langle x, y\rangle \,.
\end{aligned}$$

□

Ist V endlich-dimensional so definiert die Determinante durch

$$\mathrm{SL}(V) \;:=\; \{g \in \mathrm{GL}(V) :\; \det(g) = 1\}$$

eine Untergruppe, die *spezielle lineare Gruppe*.

Definition 7.14.9. Wir bezeichnen (im euklidischen Fall) mit

$$\mathrm{SO}(V) \; := \; \mathrm{O}(V) \cap \mathrm{SL}(V)$$

die *spezielle orthogonale Gruppe* von $(V, \langle\cdot,\cdot\rangle)$ und (im unitären Fall) mit

$$\mathrm{SU}(V) \; := \; \mathrm{U}(V) \cap \mathrm{SL}(V)$$

die *spezielle unitäre Gruppe* von $(V, \langle\cdot,\cdot\rangle)$.

Definition 7.14.10. Eine Matrix $Q \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, falls

$$Q \cdot Q^{\mathrm{tr}} \; = \; E_n$$

ist. Eine Matrix $Q \in \mathbb{C}^{n\times n}$ heißt *unitär*, falls

$$Q \cdot Q^* \; = \; E_n$$

ist. Die Menge der orthogonalen (bzw. unitären) Matrizen wird mit $\mathrm{O}_n\,\mathbb{R}$ bzw. $\mathrm{U}_n\,\mathbb{C}$ bezeichnet. Zusätzlich definieren wir $\mathrm{SO}_n\,\mathbb{R} := \mathrm{O}_n\,\mathbb{R} \cap \mathrm{SL}_n\,\mathbb{R}$ und $\mathrm{SU}_n\,\mathbb{C} := \mathrm{U}_n\,\mathbb{C} \cap \mathrm{SL}_n\,\mathbb{C}$.

Hier bezeichnet Q^* die Matrix $\bar{Q}^{\mathrm{tr}} = \overline{Q^{\mathrm{tr}}}$. Für Q reell gilt $Q^* = Q^{\mathrm{tr}}$. Die Gruppe $\mathrm{SL}_n\,\mathbb{K}$ ist die Gruppe der (invertierbaren) Matrizen mit Determinate 1.

Lemma 7.14.11. *Für Q orthogonal gilt* $\det Q \in \{\pm 1\}$. *Für Q unitär gilt* $|\det Q| = 1$.

Aufgabe 7.14.12. Die Mengen $\mathrm{O}_n\,\mathbb{R}$, $\mathrm{U}_n\,\mathbb{C}$, $\mathrm{SO}_n\,\mathbb{R}$ und $\mathrm{SU}_n\,\mathbb{C}$ bilden Untergruppen von $\mathrm{GL}_n\,\mathbb{R}$ bzw. $\mathrm{GL}_n\,\mathbb{C}$. Es gilt $\mathrm{O}_n\,\mathbb{R} \cong \mathrm{O}(\mathbb{R}^n)$ und $\mathrm{U}_n\,\mathbb{C} \cong \mathrm{U}(\mathbb{C}^n)$ sowie $\mathrm{SO}_n\,\mathbb{R} \cong \mathrm{SO}(\mathbb{R}^n)$ und $\mathrm{SU}_n\,\mathbb{C} \cong \mathrm{SU}(\mathbb{C}^n)$.

Satz 7.14.13. *Für Matrizen $Q \in \mathbb{K}^{n\times n}$ sind die folgenden Eigenschaften äquivalent.*

(i) Die lineare Abbildung $\varphi_Q : \mathbb{K}^n \to \mathbb{K}^n$ ist orthogonal (bzw. unitär).
(ii) Die Spalten von Q bilden ein Orthonormalsystem.
(iii) Die Zeilen von Q bilden ein Orthonormalsystem.
(iv) $Q \cdot Q^* = E_n$.
(v) $Q^* \cdot Q = E_n$.
(vi) $Q \in \mathrm{GL}_n\,\mathbb{K}$ und $Q^{-1} = Q^*$.

Beweis. Die Standardbasis von $\mathbb{K}^n$ eine Orthonormalbasis. Daher ist nach Proposition 7.14.8 die Abbildung φ_Q genau dann orthogonal bzw. unitär, wenn das Bild der Standardbasis unter φ_Q eine Orthonormalbasis ist. Die Bilder der Standardbasisvektoren sind genau die Spalten von Q.

Betrachte die Matrix $Q \cdot Q^*$. Der Koeffizient r_{ij} an der Stelle (i, j) ist genau das Standardskalarprodukt von $q_{i,\cdot}$ und $\overline{q_{\cdot,j}}$. Also gilt genau dann $r_{ij} = \delta_{ij}$, wenn die Zeilen von Q ein Orthonormalsystem bilden. □

Satz 7.14.14 (QR-Zerlegung). *Jede Matrix $A \in \mathbb{K}^{m\times n}$ mit vollem Spaltenrang n lässt sich schreiben als*

$$A \; = \; Q \cdot R\,,$$

wobei die Spalten von $Q \in \mathbb{K}^{m\times n}$ ein Orthonormalsystem bilden und die Matrix $R \in \mathbb{K}^{n\times n}$ eine obere Dreiecksmatrix mit vollem Rang ist.

Beweis. Seien $s_1, s_2, \dots, s_n$ die Spalten von A. Wegen $\operatorname{rank} A = n$, bilden die Spalten eine Basis des Spaltenraums $U = \operatorname{lin}_{\mathbb{K}}\{s_1, s_2, \dots, s_n\} \le \mathbb{K}^m$. Wenden wir das Gram-Schmitt-Verfahren auf $(s_1, s_2, \dots, s_n)$ an, so erhalten wir eine Orthonormalbasis $(q_1, q_2, \dots, q_n)$

von U. Nach Konstruktion gilt dann

$$\begin{aligned} s_1 &= \|q_1'\|q_1 \\ s_2 &= \langle s_2, q_1\rangle q_1 + \|q_2'\|q_2 \\ &\vdots \\ s_n &= \langle s_n, q_1\rangle q_1 + \langle s_n, q_2\rangle q_2 + \cdots + \langle s_n, q_{n-1}\rangle q_{n-1} + \|q_n'\|q_n\,. \end{aligned}$$

Hier ist

$$q_k' = s_k - \sum_{i=1}^{k-1} \langle s_k, q_i\rangle q_i \quad \text{und} q_k = \frac{1}{\|q_k'\|} q_k'\,.$$

Wir setzen nun

$$R := \begin{pmatrix} \|q_1'\| & \langle s_2, q_1\rangle & \cdots & \langle s_n, q_1\rangle \\ 0 & \|q_2'\| & \cdots & \langle s_n, q_2\rangle \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \|q_n'\| \end{pmatrix}$$

und erhalten $A = QR$. □

Aufgabe 7.14.15. Ist die Zerlegung aus Satz 7.14.14 stets eindeutig?

Bemerkung 7.14.16. Aus Proposition 7.14.8 folgt insbesondere für $m = n$ in Satz 7.14.14, dass jede lineare Transformation in $\mathrm{GL}_n\,\mathbb{K}$ sich als Produkt einer orthogonalen (bzw. unitären) Transformation und einer oberen Dreiecksmatrix schreiben lässt.

Beispiel 7.14.17.

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

7.15. Exkurs: Spiegelungsgruppen

Für $\gamma \in \mathbb{R}$ ist $a = (\cos\gamma, \sin\gamma) \in \mathbb{R}^2 \setminus \{0\}$. Es gilt $\|a\| = 1$. Bezüglich der Standardbasis hat die Spiegelung s_a aus Beispiel 7.14.6 die Matrix

$$\begin{aligned} [s_a] &= \begin{pmatrix} 1 - 2\cos^2\gamma & -2\cos\gamma\sin\gamma \\ -2\cos\gamma\sin\gamma & 2\cos^2\gamma - 1 \end{pmatrix} = \begin{pmatrix} -\cos(2\gamma) & -\sin(2\gamma) \\ -\sin(2\gamma) & \cos(2\gamma) \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\gamma+\pi) & \sin(2\gamma+\pi) \\ \sin(2\gamma+\pi) & -\cos(2\gamma+\pi) \end{pmatrix}\,. \end{aligned}$$

Betrachtet man zwei Winkel $\alpha, \beta \in \mathbb{R}$ und führt die beiden Spiegelungen hintereinander aus, so erhält man

$$\begin{aligned} &\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{pmatrix} \\ &\qquad = \begin{pmatrix} \cos\alpha\cos\beta + \sin\alpha\sin\beta & \sin\alpha\cos\beta - \cos\alpha\sin\beta \\ -\sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta + \sin\alpha\sin\beta \end{pmatrix} \\ &\qquad = \begin{pmatrix} \cos(\alpha-\beta) & \sin(\alpha-\beta) \\ -\sin(\alpha-\beta) & \cos(\alpha-\beta) \end{pmatrix}\,, \end{aligned}$$

die Drehung um den Winkel $\alpha - \beta$.

Bemerkung 7.15.1. In die vorige Rechnung fließen diverse trigonometrische Identitäten ein, die sich alle leicht aus Eulers Gleichung

$$e^{i\gamma} \;=\; \cos\gamma + i\sin\gamma$$

gewinnen lassen. Beispielsweise gilt das folgende Additionstheorem:

$$\begin{aligned}
\sin(\alpha-\beta) \;&=\; \frac{1}{2i}(e^{i(\alpha-\beta)} - e^{-i(\alpha-\beta)}) \;=\; \frac{1}{2i}(e^{i\alpha}e^{-i\beta} - e^{-i\alpha}e^{i\beta}) \\
&=\; \frac{1}{2i}\big((\cos\alpha + i\sin\alpha)(\cos\beta - i\sin\beta) - (\cos\alpha - i\sin\alpha)(\cos\beta + i\sin\beta)\big) \\
&=\; \frac{1}{2i}\big(\cos\alpha\cos\beta - \sin\alpha\sin\beta + i(\sin\alpha\cos\beta - \sin\beta\cos\alpha) \\
&\qquad - \cos\alpha\cos\beta + \sin\alpha\sin\beta + i(\sin\alpha\cos\beta - \sin\beta\cos\alpha)\big) \\
&=\; \sin\alpha\cos\beta - \cos\alpha\sin\beta\,.
\end{aligned}$$

Beispiel 7.15.2. Betrachte die Ecken

$$\begin{pmatrix}1\\0\end{pmatrix},\ \frac{1}{2}\begin{pmatrix}1\\\sqrt{3}\end{pmatrix},\ \frac{1}{2}\begin{pmatrix}-1\\\sqrt{3}\end{pmatrix},\ \begin{pmatrix}-1\\0\end{pmatrix},\ -\frac{1}{2}\begin{pmatrix}1\\\sqrt{3}\end{pmatrix},\ \frac{1}{2}\begin{pmatrix}1\\-\sqrt{3}\end{pmatrix}$$

des regelmäßigen Sechsecks auf dem Einheitskreis. Beachte, dass $\cos(\pi/3) = 1/2$ und $\sin(\pi/3) = 1/\sqrt{3}$ gilt. Das Produkt der beiden Spiegelungen

$$\sigma_1 \;=\; s_{(1,0)} \quad \text{und} \quad \sigma_2 \;=\; s_{\frac{1}{2}(-1,\sqrt{3})}$$

ist die Drehung um $2\pi/3$. Insgesamt erzeugen σ_1 und σ_2 eine Gruppe der Ordnung 6, die *Diedergruppe* D_3, die isomorph ist zur symmetrischen Gruppe vom Grad 3. Dies ist die Gruppe aller Drehungen und Spiegelungen des gleichseitigen Dreiecks

$$\operatorname{conv}\Big\{\begin{pmatrix}1\\0\end{pmatrix},\ \frac{1}{2}\begin{pmatrix}-1\\\sqrt{3}\end{pmatrix},\ -\frac{1}{2}\begin{pmatrix}1\\\sqrt{3}\end{pmatrix}\Big\}\,.$$

In dieser Gruppe gelten die Relationen

$$\sigma_1{}^2 \;=\; \sigma_2{}^2 \;=\; (\sigma_1\sigma_2)^3 \;=\; \mathrm{id}\,. \tag{32}$$

Die Diedergruppe ist durch das Erzeugendensystem $\{\sigma_1, \sigma\}$ und die Relationen (32) charakterisiert. Ähnlich lassen sich alle Diedergruppen D_n, also die Gruppen aller Drehungen und Spiegelungen eines regelmäßigen n-Ecks aus zwei Spieglungen erzeugen.

KAPITEL 8

Eigenwerte und Eigenvektoren

8.1. Definitionen und Beispiele

Sei V ein endlich-dimensionaler K-Vektorraum und $\varphi : V \to V$ ein K-linearer Endomorphismus.

DEFINITION 8.1.1. Ein Skalar $\lambda \in K$ heißt *Eigenwert* von φ, falls exisitiert $v \in V \setminus \{0\}$, so dass $\varphi(v) = \lambda v$. Jeder von 0 verschiedene Vektor w für den gilt $\varphi(w) = \lambda w$ heißt *Eigenvektor* zum Eigenwert λ. Der Unterraum

$$V_\lambda \ := \ V_\lambda(\varphi) \ := \ \ker(\varphi - \lambda \,\mathrm{id}_v) \ = \ \{v \in V : \ \varphi(v) = \lambda v\}$$

heißt *Eigenraum* zum Eigenwert λ bzgl. φ. Die Dimension $d_\lambda := \dim V_\lambda \geq 1$ heißt *geometrische Vielfachheit* von λ.

BEMERKUNG 8.1.2. Eigenwerte linearer Abbildungen treten in der Physik auf als Eigenfrequenzen schwingfähiger Systeme.

DEFINITION 8.1.3. Eine lineare Abbildung φ heißt *diagonalisierbar*, falls es eine Basis von V gibt, die aus Eigenvektoren (zu verschiedenen Eigenwerten) besteht. Das heißt die zugehörige Matrix bzgl. dieser Basis aus Eigenvektoren ist eine Diagonalmatrix.

BEISPIEL 8.1.4. Sei $V = \mathbb{R}^2$ und G eine Gerade durch 0. Wähle Orthogonalbasis (v, w) von V, so dass $\mathrm{lin}(v) = G$. Sei $\sigma = s_w$ die orthogonale Spiegelung an G; d.h. $\sigma(v) = v$ und $\sigma(w) = -w$. Dann gilt

$$\begin{cases} v & \text{Eigenvektor zum Eigenwert } 1 \\ w & \text{Eigenvektor zum Eigenwert } -1\,, \end{cases}$$

und

$$[\sigma]_{(v,w)} \ = \ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

ist eine Diagonalmatrix.

BEISPIEL 8.1.5. Sei wiederum G eine Ursprungsgerade in $\mathbb{R}^2$. Sei π die orthogonale Projektion auf G; d.h. $\pi(v) = v$ und $\pi(w) = 0$. Wiederum ist die Matrix

$$[\pi]_{(v,w)} \ = \ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

eine Diagonalmatrix.

BEISPIEL 8.1.6. Sei $\gamma \in (0, \pi)$ und ρ die Drehung in $\mathbb{R}^2$ um den Winkel γ, d.h.

$$[\rho] \ = \ \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} .$$

Dann besitzt ρ keinen Eigenwert (in $\mathbb{R}$).
Angenommen ρ hat Eigenwert λ. Dann ist

$$\begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \ = \ \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} ,$$

also
$$\begin{pmatrix} \cos\gamma - \lambda & -\sin\gamma \\ \sin\gamma & \cos\gamma - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; 0\,.$$
Für $(x_1, x_2) \neq 0$ folgt, dass
$$\det \begin{pmatrix} \cos\gamma - \lambda & -\sin\gamma \\ \sin\gamma & \cos\gamma - \lambda \end{pmatrix} \;=\; (\cos\gamma - \lambda)^2 + \sin^2\gamma \;=\; 0$$
ist. Dies ist aber äquivalent zu
$$\lambda^2 - 2\lambda\cos\gamma + 1 \;=\; 0\,.$$
Daher folgt
$$\lambda \;=\; \cos\gamma \pm \sqrt{\cos^2\gamma - 1}\,.$$
Weil aber λ reell sein soll, muss $\cos\gamma = \pm 1$ gelten, im Widerspruch zur Annahme $\pi \in (0, \pi)$.

Bemerkung 8.1.7. Es gilt:
$$\begin{aligned} \lambda \text{ Eigenwert von } \varphi \;&\Leftrightarrow\; \exists v \neq 0 : (\varphi - \lambda\,\mathrm{id})(v) = 0 \\ &\Leftrightarrow\; \ker(\varphi - \lambda\,\mathrm{id}) \neq \{0\} \end{aligned}$$
sowie
$$\begin{aligned} v \text{ Eigenvektor von } \varphi \text{ bzgl. } \lambda \;&\Leftrightarrow\; v \neq 0 \text{ und } (\varphi - \lambda\,\mathrm{id})(v) = 0 \\ &\Leftrightarrow\; v \in \ker(\varphi - \lambda\,\mathrm{id}) \setminus \{0\}\,. \end{aligned}$$

Satz 8.1.8 (Charakteristische Gleichung). *Sei* $\varphi \in \mathrm{End}(V)$ *und* $\dim_K V < \infty$. *Das Skalar* $\lambda \in K$ *ist genau dann Eigenwert von* φ, *wenn* $\det(\varphi - \lambda\,\mathrm{id}) = 0$ *ist.*

Beweis. Es gilt: λ ist kein Eigenwert von $\varphi \Leftrightarrow \ker(\varphi - \lambda\,\mathrm{id}) = \{0\} \Leftrightarrow \varphi - \lambda\,\mathrm{id}$ injektiv $\Leftrightarrow \varphi - \lambda\,\mathrm{id}$ bijektiv $\Leftrightarrow \det(\varphi - \lambda\,\mathrm{id}) \neq 0$. □

Definition 8.1.9. Sei $M \in K^{n\times n}$. Dann heißt $\lambda \in K$ *Eigenwert* von M, falls λ Eigenwert von
$$\varphi_M \;:\; K^n \to K^n, v \mapsto \varphi_M(v) = Mv$$
ist. Analog für die Eigenvektoren.

Beispiel 8.1.10. Sei $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$. Dann ist
$$\det(A - \lambda E_n) \;=\; \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} \;=\; \lambda^2 + 1 \;\neq\; 0$$
für alle $\lambda \in \mathbb{R}$; das heißt A hat keine Eigenwerte in $\mathbb{R}$.
Es gilt $A = [\rho]$ für $\gamma = \pi/2$ in Beispiel 8.1.6.

Beispiel 8.1.11. Sei $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in \mathbb{C}^{2\times 2}$. Dann ist
$$\det(A - \lambda E_n) \;=\; \lambda^2 + 1 = 0 \;\Leftrightarrow\; \lambda = i \text{ oder } \lambda = -i\,.$$
Lösen der linearen Gleichungssysteme
$$(A - iE_n) = 0 \quad \text{bzw.} \quad (A + iE_n) = 0$$
führt zu den Eigenvektoren
$$v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ (zum Eigenwert } i\text{)}$$
und
$$v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ (zum Eigenwert } -i\text{)}$$

Die Vektoren (v_1, v_2) bilden eine Basis von $\mathbb{C}^2$, und A ist diagonalisierbar über $\mathbb{C}$, das heißt in $\mathbb{C}^{2\times 2}$ ist A ähnlich zu

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} .$$

Bemerkung 8.1.12. Sei $A \in \mathbb{C}^{n\times n}$ eine Matrix mit reellen Einträgen. Falls v Eigenvektor zum Eigenwert λ, dann ist auch $\overline{v}$ Eigenvektor zum Eigenwert $\overline{\lambda}$.

Beispiel 8.1.13. Sei $K = \mathbb{F}_3$ der Körper mit drei Elementen und

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

[charakteristisches Polynom: $-t^3 - t^2 + 1 = (-t^2 - t + 1)(t-1)$]

8.2. Diagonalisierbarkeit von Matrizen

Sei V ein endlich-dimensionaler K-Vektorraum und $\varphi \in \operatorname{End}(V)$ ein Endomorphismus.

Satz 8.2.1. *Seien $\lambda_1, \dots, \lambda_r$ paarweise verschiedene Eigenwerte von φ und $v_1, \dots, v_r$ zugehörige Eigenvektoren. Dann ist $(v_1, \dots, v_r)$ linear unabhängig.*

Beweis. Per Induktion nach r. Für $r = 1$ ist $v_1 \neq 0$, also (v_1) linear unabhängig. Für den Induktionsschritt können wir also annehmen, daß $(v_1, \dots, v_{r-1})$ linear unabhängig ist. Seien $\mu_1, \dots, \mu_r \in K$ mit $\mu_1 v_1 + \cdots + \mu_r v_r = 0$. Dann folgt

$$\begin{aligned} 0 = \varphi(0) &= \varphi(\mu_1 v_1) + \cdots + \varphi(\mu_r v_r) \\ &= \mu_1 \varphi(v_1) + \cdots + \mu_r \varphi_r(v_r) \\ &= \mu_1 \lambda_1 v_1 + \cdots + \mu_r \lambda_r v_r \, . \end{aligned}$$

Dies impliziert

$$\mu_1(\lambda_1 - \lambda_r)v_1 + \cdots + \mu_{r-1}(\lambda_{r-1} - \lambda_r)v_{r-1} = 0 \, ,$$

also gilt nach Induktionsannahme

$$\mu_1(\lambda_1 - \lambda_r) = \dots = \mu_{r-1}(\lambda_{r-1} - \lambda_r) = 0 \, .$$

Weiter folgt wegen $\lambda_i \neq \lambda_j$ für $i \neq j$ dann

$$\mu_1 = \dots = \mu_{r-1} = 0 \quad \text{und damit auch} \quad \mu_r = 0 \, .$$

Damit sind $v_1, \dots, v_r$ linear unabhängig. □

Korollar 8.2.2. *Eigenräume $V_{\lambda_1}, V_{\lambda_2}$ zu verschiedenen Eigenwerten $\lambda_1 \neq \lambda_2$ schneiden sich nur in der* 0.

Korollar 8.2.3. *Sei V ein K-Vektorraum mit $\dim_K V = n$ und $\lambda_1, \dots, \lambda_r$ paarweise verschiedene Eigenwerte von $\varphi \in \operatorname{End}(V)$ mit $d_{\lambda_1} + \cdots + d_{\lambda_r} = n$. Dann ist φ diagonalisierbar.*

Beweis. Seien $\lambda_1, \dots, \lambda_r$ paarweise verschiedene Eigenwerte von φ mit den geometrischen Vielfachheiten $d_1, \dots, d_r$. Sei $(b_{i,1}, \dots, b_{i,d_i})$ eine Basis des Eigenraums V_{λ_i}. Dann ist

$B = (b_{1,1}, \ldots, b_{1,d_1}, \ldots, b_{r,1}, \ldots, b_{r,d_r})$ eine Basis von $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$. Gilt nun $d_1 + \cdots + d_r = n$, dann ist $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r} = V$, und φ ist diagonalisierbar.

$$[\varphi]_B = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix} \lambda_r & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{matrix}} \end{pmatrix}$$

□

8.3. Das charakteristische Polynom

8.3.1. Wiederholung: Polynome. Vergleiche Kapitel 6. Sei K ein beliebiger Körper.

- ▷ Ein Polynom mit Koeffizienten in K ist eine Abbildung
$$p : \mathbb{N} \to K : i \mapsto p_i$$
mit $\#\{i \in \mathbb{N} : p_i \neq 0\} < \infty$.
- ▷ $(K[t], +, \cdot)$ ist die K-Algebra der Polynome in der Unbestimmten t. Die Monome $1, t, t^2, t^3, \ldots$ bilden eine K-Basis. Insbesondere gilt $\dim_K K[t] = \infty$.
- ▷ Zu einem Polynom $p \in K[t]$ gibt es eine Einsetzungsabbildung
$$p^* : K \to K : x \mapsto p_0 + p_1 x + \cdots + p_d x^d ,$$
wobei $d = \deg p$ der Grad von p ist.
- ▷ Der Ring $K[t]$ ist nullteilerfrei, weil $\deg(p \cdot q) = \deg p + \deg q$ gilt.
- ▷ Für eine Nullstelle $\lambda \in K$, die also $p^*(\lambda) = 0$ erfüllt, gilt, dass $(t - \lambda)$ das Polynom p im Ring $K[t]$ teilt.
- ▷ Jedes Polynom vom Grad d besitzt höchstens d Nullstellen in K.
- ▷ Der Fundamentalsatz der Algebra besagt, dass jedes Polynom in $\mathbb{C}[t]$ vom Grad ≥ 1 eine komplexe Nullstelle besitzt. Anders ausgedrückt: Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen.

Definition 8.3.1. Seien $a \in K[t]$ und $\lambda \in K$, so dass $a = (t - \lambda)^s b$ für $s \geq 1$ und $b \in K[t]$ mit $b^*(\lambda) \neq 0$. Dann heißt s die *Vielfachheit* der Nullstelle λ von a, und λ heißt *s-fache Nullstelle* von a.

Bemerkung 8.3.2. Seien $\lambda_1, \ldots, \lambda_r$ die verschiedenen Nullstellen des Polynoms $a \in K[t]$ mit Vielfachheiten $s_1, \ldots, s_r$. Dann existiert eindeutig ein Polynom $b \in K[t]$ ohne Nullstellen, so dass
$$a = (t - \lambda_1)^{s_1} \ldots (t - \lambda_r)^{s_r} b.$$

8.3.2. Definition und Beispiele. Sei $A \in K^{n \times n}$ eine quadratische Matrix.

Definition 8.3.3. Der Ausdruck
$$\chi_A(t) = \det(A - tE_n)$$
heißt *charakteristisches Polynom* von A.

Beispiel 8.3.4. Vergleiche Beispiel 8.1.6. Sei $K = \mathbb{R}$ und

$$A \;=\; \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \quad \text{für } \gamma \in (0, \pi).$$

Dann ist

$$\begin{aligned} \chi_A \;&=\; \det(A - tE_n) \;=\; \det \begin{pmatrix} \cos\gamma - t & -\sin\gamma \\ \sin\gamma & \cos\gamma - t \end{pmatrix} \\ &=\; \cos^2\gamma - 2\cos\gamma t + t^2 + \sin^2\gamma \\ &=\; 1 - 2\cos\gamma t + t^2 \end{aligned}$$

Beispiel 8.3.5. Sei nun K beliebig und

$$A \;=\; \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \;\in\; \mathbb{R}^{2\times 2}$$

Dann gilt

$$\begin{aligned} \chi_A(t) \;&=\; \det \begin{pmatrix} a_{11} - t & a_{12} \\ a_{21} & a_{22} - t \end{pmatrix} \\ &=\; t^2 - (a_{11} + a_{22})t + a_{11}a_{22} - a_{12}a_{21} \\ &= t^2 - (\operatorname{trace} A)t + \det A\,. \end{aligned}$$

Definition 8.3.6. Für eine quadratische Matrix $A \in K^{n\times n}$ heißt die Summe der Diagonalelemente auch *Spur* von A und wird mit $\operatorname{trace} A$ bezeichnet.

Satz 8.3.7. *Das charakteristische Polynom $\chi_A(t) \in K[A]$ ist ein Polynom vom Grad n. Es gilt*

$$\chi_A(t) \;=\; (-1)^n t^n + (-1)^{n-1}(\operatorname{trace} A)t^{n-1} + \cdots + \det A\,.$$

Beweis. Aus der Leibnizformel 5.4.4 folgt

$$\begin{aligned} \chi_A(t) \;&=\; (a_{11} - t)(a_{22} - t)\cdots(a_{nn} - t) \\ &\qquad + \text{Produkte, in denen höchstens } n-2 \text{ Faktoren } (a_{ii} - t) \text{ auftreten} \\ &=\; (-1)^n t^n + (-1)^{n-1}(\operatorname{trace} A)t^{n-1} + \text{Terme niedrigeren Grades}\,. \end{aligned}$$

□

Bemerkung 8.3.8. Wir hatten die Leibniz-Formel (und andere Eigenschaften der Determinante) nur für Matrizen mit Körperkoeffizienten bewiesen. Der Polynomring $K[t]$ ist nullteilerfrei, und daher in seinen Quotientenkörper $K(t)$ einbettbar; vergleiche Abschnitt 6.7. Daher gelten alle Aussagen über Determinanten von Matrizen mit Koeffizienten in $K[t]$ entsprechend.

Bemerkung 8.3.9. Die Eigenwerte von A sind genau die Nullstellen des charakteristischen Polynoms χ_A. Da $\deg_{\chi_A} = n$, folgt (erneut), dass A höchstens n Eigenwerte hat.

Definition 8.3.10. Sei λ Eigenwert von A. Dann heißt die Vielfachheit der Nullstelle λ in $\chi_A(t) = \det(A - tE_n)$ die *algebraische Vielfachheit* des Eigenwertes λ.

Proposition 8.3.11. *Für jeden Eigenwert λ von A ist die geometrische Vielfachheit d_λ stets kleiner oder gleich der algebraischen Vielfachheit.*

Beweis. Übungsaufgabe. □

Beispiel 8.3.12. Sei $K = \mathbb{R}$ und $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$. Dann ist

$$\chi_A(t) = \begin{vmatrix} 1-t & 2 \\ 0 & 1-t \end{vmatrix} = (1-t)^2 .$$

Also ist 1 einziger Eigenwert von A (mit der algebraischen Vielfachheit 2).

Um die Eigenvektoren von A zum Eigenwert 1 zu bestimmen, betrachten wir das lineare Gleichungssystem

$$(A - 1E_n)v = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} v = 0 .$$

Es folgt, dass $v \in \mathbb{R}\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ ist, also die geometrische Vielfachheit $d_1 = 1$ ist. Das heißt, es kommt tatsächlich vor, dass die geometrische Vielfachheit echt kleiner ist als die algebraische.

Satz 8.3.13. *Die Matrix $A \in K^{n \times n}$ ist genau dann über K diagonalisierbar, wenn $\chi_A \in K[t]$ in Linearfaktoren zerfällt, d.h.*

$$\chi_A(t) = \pm(t - \lambda_1)^{e_1} \dots (t - \lambda_r)^{e_r} ,$$

und falls für jeden Eigenwert λ_i die algebraische Vielfachheit mit der geometrischen übereinstimmt.

Beweis. Seien $\lambda_1, \lambda_2 \dots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A. Ist A diagonalisierbar und sind $d_1, d_2, \dots, d_r$ die geometrischen Vielfachheiten der Eigenwerte, so existiert eine invertierbare Matrix $S \in \mathrm{GL}_n K$, so dass

$$A' := S^{-1}AS = \begin{pmatrix} \left.\boxed{\begin{matrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}}\right\} d_1 & & 0 \\ & \ddots & \\ 0 & & \left.\boxed{\begin{matrix} \lambda_r & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{matrix}}\right\} d_r \end{pmatrix}$$

ist. Hierbei sind die Spalten von S Eigenvektoren zu den jeweils entsprechenden Eigenwerten. Es folgt weiter

$$\begin{aligned} \chi_A(t) &= \det(A - tE_n) = \det(SA'S^{-1} - tE_n) = \det(SA'S^{-1} - tSE_nS^{-1}) \\ &= \det(S(A' - tE)S^{-1}) = \det(A' - tE_n) \\ &= \chi_{A'}(t) = \prod_{i=1}^{r} (\lambda_i - t)^{d_i} . \end{aligned} \tag{33}$$

Es gilt $d_i = e_i$ für alle i.

Nehmen wir umgekehrt an, dass $\chi_A(t) = (\lambda_1 - t)^{e_1} \dots (\lambda_r - t)^{e_r}$ in Linearfaktoren und zerfällt und zusaätzlich $e_i = d_i$ gilt für alle i. Dann ist

$$n = \deg \chi_A = e_1 + e_2 + \dots + e_r = d_1 + d_2 + \dots + d_r ,$$

und A ist diagonalisierbar. $\square$

Bemerkung 8.3.14. Aus der Rechnung (33) im Beweis zu Satz 8.3.13 folgt, dass ähnliche Matrizen dasselbe charakteristische Polynom besitzen.

KOROLLAR 8.3.15. *Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann diagonalisierbar, wenn für jeden Eigenwert $\lambda_i \in \mathbb{C}$ die geometrische und die algebraische Vielfachheit übereinstimmen.*

BEWEIS. Der Fundamentalsatzes besagt, dass χ_A in Linearfaktoren zerfällt. □

KOROLLAR 8.3.16. *Jede komplexe Matrix hat einen Eigenwert.*

DEFINITION 8.3.17. Die *Begleitmatrix* des normierten Polynoms

$$p(t) \;=\; t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \;\in\; K[t]$$

vom Grad n ist die Matrix

$$C_p \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix} \;\in\; K^{n\times n}\,.$$

SATZ 8.3.18. *Das charakteristische Polynom der Begleitmatrix des normierten Polynoms*

$$p(t) \;=\; t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in K[t]$$

vom Grad $n \geq 1$ ist

$$\det(C_p - tE_n) \;=\; (-1)^n p(t)\,.$$

BEWEIS. Der Beweis erfolgt durch vollständige Induktion. Für $n = 1$ ist die Aussage klar, und für $n > 1$ erhält man durch Streichen der ersten Zeile und der ersten Spalte von C_p die Begleitmatrix des Polynoms $q(t) = t^{n-1} + a_{n-1}t^{n-2} + \cdots + a_2 t + a_1$. Damit ergibt sich durch Entwicklung nach der ersten Spalte

$$\det(C_p - tE_n) \;=\; (-t)(-1)^{n-1}q(t) + (-1)^{n+1}(-a_0)\,.$$

Wir erhalten

$$\det(C_p - tE_n) \;=\; (-1)^n p(t)\,.$$

□

8.4. Exkurs: Das PageRank-Verfahren von Google

Das Page-Rank-Verfahren ergibt das Hauptkriterium, nach dem die Suchmaschine Google die Ergebnisse einer Suchanfrage ordnet. Es ist also ein Maß für die „Wichtigkeit“ einer Webseite. Da allerdings das Verfahren in seiner reinen Form relativ leicht manipuliert werden kann (Stichwort „Link-Farming“), kommen in der Realität weitere Kritierien hinzu. Hier wird nur die ursprüngliche Idee skizziert.

Das hier vorgestellte Verfahren wird in der folgenden Arbeit beschrieben:

▷ Sergey Brin & Lawrence Page: The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 33:107–117, 1998.

Die Methode beruht auf dem Konzept der *Markovketten* aus der Stochastik. Da wahrscheinlichkeitstheoretische Grundlagen nicht vorausgesetzt werden, ist die Diskussion hier in einzelnen Punkten unvollständig. Wir analysieren die Situation vor allem anhand eines konkreten (ausgedachten) Beispiels.

Das Modell geht aus von einem Internet-Benutzer, der zufällig auf einer Webseite beginnt und anschließend zufällig auf einen der Links auf dieser Seite klickt und sich so durch das Internet bewegt. Kommt der Surfer dabei auf eine Seite ohne ausgehende Links, fängt er wieder mit einer zufälligen Seite an. Das Maß für die Wichtigkeit einer Webseite in diesem Modell ist der Zeitanteil, den der Surfer erwartungsgemäß auf einer Seite verbringt.

Betrachte ein Webstruktur aus vier Seiten A, B, C, D, die wie folgt verlinkt sind:

$$\begin{tikzcd} A \arrow[r] \arrow[d, leftrightarrow] \arrow[dr] & B \arrow[dl] \\ C \arrow[r, leftrightarrow] & D \end{tikzcd}$$

Wir bezeichnen die Wahrscheinlichkeit, dass sich der Surfer zu einem diskreten Zeitpunkt $t \in \mathbb{N}$ auf der Webseite X aufhält mit $\mathrm{Pr}_t(X)$. Die Wahrscheinlichkeit auf einer der vier Seiten zu beginnen, beträgt jeweils 1/4:

$$\mathrm{Pr}_0(A) \;=\; \mathrm{Pr}_0(B) \;=\; \mathrm{Pr}_0(C) \;=\; \mathrm{Pr}_0(D) \;=\; 1/4\,.$$

Im nächsten Schritt kann etwa die Seite C von A oder von B aus erreicht werden, wobei A zwei Links hat und B nur eins. Also gilt

$$\mathrm{Pr}_{t+1}(C) \;=\; \frac{\mathrm{Pr}_t(A)}{2} + \frac{\mathrm{Pr}_t(B)}{1} + \frac{\mathrm{Pr}_t(D)}{1}\,,$$

d.h.

$$\mathrm{Pr}_1(C) \;=\; 1/8 + 1/4 + 1/4 \;=\; 1/2$$

und entsprechend für die übrigen Seiten.

Bemerkung 8.4.1. Empirisch wurde ein *Dämpfungsfaktor* von $d = 0{,}85$ eingeführt; dies ist die Wahrscheinlichkeit, dass der imaginäre Websurfer weiter klickt. Sonst fängt er wieder neu bei einer zufälligen Seite an.

Für Webseiten $p_1, p_2, \ldots, p_n$ entsteht dann folgendes Rechenschema:

$$\mathrm{Pr}_{t+1}(p_k) \;=\; \frac{1-d}{n} + d \sum_{p_i \in M(p_k)} \frac{\mathrm{Pr}_t(p_i)}{\ell(p_i)}$$

wobei $M(p_k)$ die Menge der Seiten von $p_1, \ldots, p_n$ ist, die einen Link auf p_k gesetzt haben, und

$$\ell(p_i) \;=\; \#\text{Links von } p_i \text{ auf andere Seiten}\,.$$

Diese Formel berücksichtigt nicht direkt, dass eine Seite ohne ausgehende Links wiederum mit einer Zufallsseite startet. Das lässt sich aber dadurch beheben, dass man eine künstliche zusätzliche Seite einführt, auf die alle Seiten ohne echte Links verweisen, und die selbst wieder Links auf alle Seiten besitzt. Man kann dann zeigen, dass

$$\lim_{t\to\infty} \mathrm{Pr}_t(p_k) \;=:\; \mathrm{Pr}(p_k)$$

existiert.

In unserem Beispielweb ist

$$M(A) = \{C\}\,, \quad M(B) = \{A\}\,, \quad M(C) = \{A, B, D\}\,, \quad M(D) = \{C\}$$

und

$$\ell(A) = 2\,, \quad \ell(B) = 1\,, \quad \ell(C) = 2\,, \quad \ell(D) = 1\,.$$

Was aber hat das alles mit linearer Algebra zu tun? Setze

$$R \;:=\; \begin{pmatrix} \mathrm{Pr}(p_1) \\ \vdots \\ \mathrm{Pr}(p_n) \end{pmatrix} \;\in\; \mathbb{R}^n\,.$$

Dann gilt

$$(34) \qquad R = \begin{pmatrix} (1-d)/n \\ \vdots \\ (1-d)/n \end{pmatrix} + d \begin{pmatrix} s_{11} & \dots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \dots & s_{nn} \end{pmatrix} \cdot R \,,$$

wobei

$$s_{ki} = \begin{cases} 1/\ell(p_i) & \text{falls } p_i \in M(p_k) \\ 0 & \text{sonst}\,. \end{cases}$$

Es gilt also für alle $i \in \{1, 2, \dots, n\}$, dass

$$\sum_{k=1}^{n} s_{ki} = 1\,.$$

Anders ausgedrückt, jede Spaltensumme der Matrix $S = (s_{ki})$ beträgt 1 (oder 0 für Seiten ohne Links). In unserem Beispiel ist

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

Ignoriert man die Dämpfung, setzt also $d = 1$, dann besagt die Formel (34), dass R ein Eigenvektor der Matrix S zum Eigenwert 1 ist. In unserem Fall stellt sich heraus, dass S die Eigenwerte 1, 0 und $-1/2$ besitzt. Der Eigenraum zum Eigenwert 1 ist eindimensional, und ein Eigenvektor ist

$$\begin{pmatrix} 2 \\ 1 \\ 4 \\ 2 \end{pmatrix}.$$

Um dies als Vektor von Wahrscheinlichkeiten interpretieren zu können, muss man mit 1/9 skalieren.

Will man die Dämpfung einbeziehen, so stellt man fest, dass (34) äquivalent ist zu

$$\begin{pmatrix} 1 \\ R \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ (1-d)/n & ds_{11} & \dots & ds_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ (1-d)/n & ds_{n1} & \dots & ds_{nn} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ R \end{pmatrix}$$

d.h. $\left(\begin{smallmatrix} 1 \\ R \end{smallmatrix}\right)$ ist Eigenvektor der $(n+1) \times (n+1)$-Matrix

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ (1-d)/n & ds_{11} & \dots & ds_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ (1-d)/n & ds_{n1} & \dots & ds_{nn} \end{pmatrix}$$

zum Eigenwert 1.

Weitere Referenzen sind

- ▷ Amy N. Langville & Carl D. Meyer: Google's PageRank and Beyond, Princeton 2006.
- ▷ Ehrhard Behrends: Introduction to Markov Chains, Vieweg 2000.

8.5. Der Satz von Cayley-Hamilton

Zusätzlich zur Auswertungsabbildung in K hatten wir in Abschnitt 6.4 einem Polynom $a = a_0 + a_1 t + \cdots + a_N t^N \in K[t]$ auch eine zweite Auswertungsabbildung

$$a^* \,:\, K^{n\times n} \to K^{n\times n} \,:\, M \mapsto a_0 E_n + a_1 M + \cdots + a_N M^N$$

zugeordnet.

Bemerkung 8.5.1. Die Abbildung $\varphi_M : K[t] \to K^{n\times n} : a \mapsto a^*(M)$ ist ein K-Algebra-Homomorphismus. Man kann ebenso eine Auswertungsabbildung in $\operatorname{End}_K V$ betrachten für beliebige K-Vektorräume V.

Beispiel 8.5.2. Wir betrachten das Polynom

$$a \;=\; t^2 - 2t + 1 \;=\; (t-1)^2$$

in $\mathbb{R}[t]$ sowie die reelle 2×2-Matrix $M = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Dann ist

$$a^*(M) \;=\; (M - E_2)^2 \;=\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 \;=\; \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Der Ring $K[t]$ ist nullteilfrei, aber der Ring $K^{n\times n}$ ist dies nicht (für $n \geq 2$). Es gilt für alle λ in K:

$$a^*(\lambda E_n) \;=\; a^*(\lambda) E_n\,.$$

Das heißt insbesondere, falls λ Nullstelle von a ist, gilt $a(\lambda E_n) = 0$.

Angesichts der Überlegungen in Abschnitt 6.4 ist es klar, dass es für jede Matrix $A \in K^{n\times n}$ ein Polynom (vom Grad höchstens n^2) geben muss, dessen Einsetzungsabbildung auf A angewendet 0 ergibt. Weitreichender gilt aber sogar das Folgende.

Satz 8.5.3 (Cayley-Hamilton). *Sei $A \in K^{n\times n}$. Dann gilt $\chi_A^*(A) = 0$.*

Wir wiederholen zunächst einiges zur Adjunkten einer Matrix. Sei $C \in K^{n\times n}$ eine quadratische Matrix. Aus (16) wissen wir, dass

$$\det(C) = \sum_{j=1}^{n} (-1)^{i+j} \gamma_{ij} \det(C_{ij}),$$

wobei $C = (\gamma_{ij})$ und

$$C_{ij} \;=\; \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{ij} & \cdots & \gamma_{1n} \\ \vdots & & \vdots & & \vdots \\ \gamma_{1i} & \cdots & \gamma_{ij} & \cdots & \gamma_{in} \\ \vdots & & \vdots & & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nj} & \cdots & \gamma_{nn} \end{pmatrix}_{i,j} \;=\; \left(\begin{array}{cc|cc} \gamma_{11} & \cdots & \cdots & \gamma_{1n} \\ \vdots & & & \vdots \\ \hline \vdots & & & \vdots \\ \gamma_{n1} & \cdots & \cdots & \gamma_{nn} \end{array}\right) \;\in\; K^{(n-1)\times(n-1)}$$

gilt. Setze $\gamma'_{ij} := (-1)^{i+j} \det(C_{ij})$. Die Matrix

$$\operatorname{adj}(C) \;:=\; C^{\#} \;=\; (\gamma'_{ij})_{i,j} \in K^{n\times n}$$

heißt *Adjunkte* von C. Es gilt nach Proposition 5.7.1

$$\operatorname{adj}(C) \cdot C \;=\; (\det C) E_n\,.$$

Beweis des Satzes von Cayley-Hamilton. Es gilt

$$\operatorname{adj}(A - tE_n) \cdot (A - tE_n) \;=\; \det(A - tE_n) \cdot E_n\,. \tag{35}$$

Die Koeffizienten der Matrix $\operatorname{adj}(A - tE_n)$ sind Polynome in t vom Grad $\leq n-1$ mit Koeffizienten in K. Also existieren Matrizen $B_0, \dots, B_{n-1} \in K^{n\times n}$, so dass

$$\operatorname{adj}(A - tE_n) \;=\; B_0 + B_1 t + \cdots + B_{n-1}t^{n-1}\,.$$

Dieses Polynom mit Koeffizienten in $K^{n\times n}$ setzen wir in (35) ein; per Koeffizientenvergleich erhalten wir

$$\left\{\begin{array}{rcrcr} B_0A & & & = & a_0E_n \\ B_1A & - & B_0 & = & a_1E_n \\ B_2A & - & B_1 & = & a_2E_n \\ \vdots & & & & \vdots \\ B_{n-1}A & - & B_{n-1} & = & a_{n-1}E_n \\ & - & B_{n-1} & = & (-1)^nE_n \end{array}\right.$$

wobei $\chi_A = \det(A - tE_n) = a_0 + a_1t + \cdots + (-1)^nt^n$ ist. Multiplizieren der ersten Gleichung in (35) mit E_n, der zweiten mit A, …, der $(n+1)$-ten mit A^n und anschließendes Aufsummieren ergibt

$$0 \;=\; \chi_a^*(A)\,.$$

□

Satz 8.5.4 (Satz vom Minimalpolynom). *Für jede Matrix $A \in K^{n\times n}$ existiert genau ein Polynom μ_A minimalen Grades mit Leitkoeffizient 1, für das gilt*

$$\mu_A^*(A) \;=\; 0\,.$$

Das Polynom $\mu_A \in K[t]$ heißt *Minimalpolynom* von A.

Beweis. Nach Satz 8.5.3 gilt $\chi_A^*(A) = 0$, und χ_a hat den Leitkoeffizienten $(-1)^n$. Also existiert mit $(-1)^n\chi_A$ ein Polynom mit Leitkoeffizient 1, das A annulliert. Hieraus folgt dann die Existenz eines Polynoms mit denselben Eigenschaften und minimalen Grades. Dieser Grad sei mit d bezeichnet.

Um die Eindeutigkeit zu zeigen, nehmen wir an es existieren zwei Polynome μ und ν in $K[t]$ mit $\mu = \mu_0 + \mu_1t + \cdots + t^d$ und $\nu = \nu_0 + \nu_1t + \cdots + t^d$ mit

$$\mu^*(A) \;=\; 0 \;=\; \nu^*(A)\,.$$

Dann ist $(\mu - \nu)^*(A) = 0$ und $\deg(\mu - \nu) < d$ im Widerspruch zur Minimalität von d. □

Bemerkung 8.5.5. Die Existenz annullierender Polynome zu linearen Abbildungen ist nur in endlich-dimensionalen Vektorräumen gesichert.

Wie in der Bemerkung 8.3.14 erläutert haben ähnliche Matrizen dasselbe charakteristische Polynom. Dasselbe gilt auch für das Minimalpolynom.

Satz 8.5.6. *Ähnliche Matrizen haben dasselbe Minimalpolynom.*

Beweis. Seien $A \in K^{n\times n}$, $S \in \mathrm{GL}_n K$ und $B = S^{-1}AS$ ähnlich zu A. Für $\mu_A = \mu_0 + \mu_1t + \cdots + \mu_{d-1}t^{d-1} + t^d$ gilt dann

$$\begin{aligned} \mu_A^*(B) \;&=\; \mu_0E_n + \mu_1B + \cdots + \mu_{d-1}B^{d-1} + B^d \\ &=\; \mu_0E_n + \mu_1S^{-1}AS + \cdots + \mu_{d-1}(S^{-1}AS)^{d-1} + (S^{-1}AS)^d \\ &=\; \mu_0S^{-1}E_nS + \mu_1S^{-1}AS + \cdots + \mu_{d-1}S^{-1}A^{d-1}S + S^{-1}A^dS \\ &=\; S^{-1}\mu_A^*(A)S \;=\; S^{-1}0S \;=\; 0\,. \end{aligned} \tag{36}$$

Aus Satz 8.5.4 folgt hieraus, dass $\deg \mu_A \geq \deg \mu_B$ ist. Analog folgt $\mu_B^*(A) = 0$ und $\deg \mu_A \leq \deg \mu_B$. Wiederum wegen (36) folgt aus der Eindeutigkeit des Minimalpolynoms, dass $\mu_A = \mu_B$ ist. □

Satz 8.5.7. *Sei $A \in K^{n\times n}$. Dann haben χ_A und μ_A dieselben Nullstellen.*

Beweis. Übungsaufgabe. □

Beispiel 8.5.8. Wir betrachten die Matrix

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

Für das charakteristische Polynom gilt dann

$$\chi_A = \begin{vmatrix} 1-t & 1 & 1 \\ -1 & 3-t & 1 \\ 0 & 0 & 2-t \end{vmatrix} = (2-t)((3-t)(1-t)+1) = (2-t)^3.$$

Also ist der Grad des Minimalpolynoms μ_A höchstens drei, und es gilt

$$\mu_A \in \{t-2, (t-2)^2, (t-2)^3\}.$$

Wir berechnen

$$\begin{aligned} (t-2)^*(A) &= A - 2E_3 = \begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \neq 0 \quad \text{und} \\ [(t-2)^2]^*(A) &= (A-2E_3)^2 = 0 \end{aligned}$$

Damit folgt $\mu_A = (t-2)^2$.

8.6. Diagonalisierung normaler Matrizen

Im Folgenden ist stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $(\mathbb{K}^n, \langle\cdot,\cdot\rangle)$ der Standard-euklidische bzw. -unitäre Raum.
Für $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ ist

$$\overline{A} = (\overline{a}_{ij})$$

die *komplex-konjugierte Matrix* und

$$A^* = \left(\overline{A}\right)^{\mathrm{tr}} = \overline{A^{\mathrm{tr}}} = (\overline{a}_{ji})$$

die *Adjungierte* von A. Es gelten die folgenden Rechenregeln

$$\begin{aligned} (A+B)^* &= A^* + B^* \\ (\lambda A)^* &= \overline{\lambda} A^* \quad \text{für } \lambda \in \mathbb{C} \\ (AB)^* &= B^* A^* \\ A^{**} &= A \end{aligned}$$

Falls A reell ist, so gilt $A^* = A^{\mathrm{tr}}$. Wiederholung: Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt

(i) *symmetrisch*, falls $A = A^{\mathrm{tr}}$;
(ii) *schiefsymmetrisch*, falls $A = -A^{\mathrm{tr}}$;
(iii) *hermitesch* (oder *selbstadjungiert*), falls $A = A^*$;
(iv) *schiefhermitesch*, falls $A = -A^*$.

Beispiel 8.6.1. Die Matrix

$$\begin{pmatrix} 1 & -i & 1-i \\ i & 2 & 2+i \\ 1+i & 2-i & -2 \end{pmatrix}$$

aus $\mathbb{C}^{3\times 3}$ ist hermitesch.

Lemma 8.6.2. *Für $A \in \mathbb{C}^{n\times n}$ gilt:*

$$\begin{aligned} A &= \frac{1}{2}\underbrace{(A+\overline{A})}_{\text{reell}} + \frac{1}{2}\underbrace{(A-\overline{A})}_{\text{imaginär}} \\ &= \frac{1}{2}\underbrace{(A+A^{\mathrm{tr}})}_{\text{symmetrisch}} + \frac{1}{2}\underbrace{(A-A^{\mathrm{tr}})}_{\text{schiefsymmetrisch}} \\ &= \frac{1}{2}\underbrace{(A+A^{*})}_{\text{hermitesch}} + \frac{1}{2}\underbrace{(A-A^{*})}_{\text{schiefhermitesch}} . \end{aligned}$$

Bemerkung 8.6.3. Für komplexe 1×1-Matrizen ist die erste Zerlegung genau die übliche Zerlegung in Real- und Imaginärteil.

Für $A \in \mathbb{C}^{n\times n}$ und $v, w \in \mathbb{C}^n$ gilt

$$\langle Av, w\rangle = (Av)^{\mathrm{tr}}\overline{w} = v^{\mathrm{tr}}A^{\mathrm{tr}}\overline{w} = v^{\mathrm{tr}}\overline{A^*}\overline{w} = v^{\mathrm{tr}}\overline{A^* w} = \langle v, A^* w\rangle .$$

Hieraus folgt weiter

$$\langle v, Aw\rangle = \langle v, A^{**}w\rangle = \langle A^* v, w\rangle .$$

Definition 8.6.4. Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt *normal*, falls $AA^* = A^*A$.

Beispiel 8.6.5.

(i) Unitäre Matrizen sind normal.
(ii) Reelle orthogonale Matrizen sind normal.
(iii) Diagonalmatrizen sind normal.
(iv) $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$ ist nicht normal, denn

$$A^*A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = AA^*$$

(v) Hermitesche und schiefhermitesche Matrizen sind normal.
(vi) Reelle symmetrische und schiefsymmetrische Matrizen sind normal.

Lemma 8.6.6. *Sei $A \in \mathbb{C}^{n\times n}$ normal. Dann gilt*

(i) $A - \lambda E_n$ ist normal für alle $\lambda \in \mathbb{C}$,
(ii) Q^*AQ ist normal für alle $Q \in \mathrm{U}_n\,\mathbb{C}$.

Beweis. Sei $\lambda \in \mathbb{C}$. Dann gilt

$$\begin{aligned} (A-\lambda E_n)(A-\lambda E_n)^* &= (A-\lambda E_n)(A* - \overline{\lambda}E_n) \\ &= AA^* - \lambda A^* - \overline{\lambda}A + \lambda\overline{\lambda} = A^*A - \overline{\lambda}A - \lambda A^* + \overline{\lambda}\lambda \\ &= (A^* - \overline{\lambda}E_n)(A-\lambda E_n) = (A-\lambda E_n)^*(A-\lambda E_n) . \end{aligned}$$

Die zweite Aussage beweist man ähnlich. □

Das (euklidische bzw. unitäre) Skalarprodukt auf $\mathbb{K}^n$ und die zugehörige Norm sind positiv definit. Dies bedeutet für $A \in \mathbb{K}^{n\times n}$, $v \in \mathbb{K}^n \setminus \{0\}$ und $\lambda \in \mathbb{K}$:

(37) v Eigenvektor zum Eigenwert λ von $A \iff Av - \lambda v = 0 \iff \|Av - \lambda v\| = 0\,.$

Lemma 8.6.7. *Sei $A \in \mathbb{C}^{n\times n}$ normal, $\lambda \in \mathbb{C}$ Eigenwert von A und $v \in \mathbb{C}^n$ ein zugehöriger Eigenvektor. Dann ist $\overline{\lambda}$ Eigenwert von A^* mit demselben Eigenvektor v.*

Beweis. Für beliebige $\mu \in \mathbb{C}$ und $w \in \mathbb{C}^n \setminus \{0\}$ gilt:

$$\begin{aligned}\|Aw - \mu w\|^2 &= \langle (A - \mu E_n)w, (A - \mu E_n)w\rangle = \langle (A - \mu E_n)^*(A - \mu E_n)w, w\rangle \\ &= \langle (A - \mu E_n)(A - \mu E_n)^*w, w\rangle = \langle (A - \mu E_n)^*w, (A - \mu E_n)w\rangle \\ &= \langle (A^* - \overline{\mu}E_n)w, (A^{**} - \overline{\overline{\mu}}E_n)w\rangle = \|A^*w - \overline{\mu}w\|^2\,.\end{aligned}$$

Die Behauptung folgt dann aus (37). □

Proposition 8.6.8. *Sei $A \in \mathbb{C}^{n\times n}$ hermitesch. Dann sind alle Eigenwerte von A reell. Insbesondere sind die komplexen Eigenwerte reeller symmetrischer Matrizen stets reell.*

Beweis. Sei $A^* = A$ und $v \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert λ von A. Dann gilt $Av = \lambda v$, also $A^*v = \overline{\lambda}v$ wegen Lemma 8.6.7. Damit ist also

$$\lambda v = Av = A^*v = \overline{\lambda}v\,.$$

Wegen $v \neq 0$ folgt, dass $\lambda = \overline{\lambda}$ reell ist. □

Bemerkung 8.6.9. Komplexe symmetrische Matrizen können nicht reelle Eigenwerte haben, z.B.

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Lemma 8.6.10. *Sei $A \in \mathbb{C}^{n\times n}$ normal. Dann sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal.*

Beweis. Seien v und w Eigenvektoren zu den Eigenwerten λ bzw. μ. Dann gilt

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Av, w\rangle = \langle v, A^*w\rangle = \langle v, \overline{\mu}w\rangle = \mu\langle v, w\rangle\,.$$

Aus $\lambda \neq \mu$ folgt nun $\langle v, w\rangle = 0$. □

Lemma 8.6.11. *Sei $A \in \mathbb{C}^{n\times n}$ normal und v ein Eigenvektor von A (zu einem beliebigen Eigenwert). Dann gilt $A(v^\perp) \subseteq v^\perp$.*

Das heißt, der lineare Unterraum $v^\perp$ ist *invariant* bezüglich A. Wegen $v \neq 0$ ist $\dim_{\mathbb{C}} v^\perp = n - 1$.

Beweis. Es sei v Eigenvektor von A zum Eigenwert λ, und es sei $w \in v^\perp$ beliebig. Es ist $\langle v, w\rangle = 0$, und wir müssen zeigen, dass auch $\langle v, Aw\rangle = 0$ gilt.

$$\langle v, Aw\rangle = \langle A^*v, w\rangle = \langle \overline{\lambda}v, w\rangle = \overline{\lambda}\langle v, w\rangle = 0\,.$$

□

Satz 8.6.12 (Hauptsatz über normale Matrizen). *Sei $A \in \mathbb{C}^{n\times n}$ eine Matrix. Dann sind die folgenden Aussagen äquivalent:*

(i) Die Matrix A ist normal.
(ii) Der $\mathbb{C}$-Vektorraum $\mathbb{C}^n$ besitzt eine Orthonormalbasis aus Eigenvektoren zu A.
(iii) Es existiert eine unitäre Matrix $Q \in \mathrm{U}_n\,\mathbb{C}$, so dass Q^*AQ eine Diagonalmatrix ist.

Beweis. Sei A normal. Dann konstruieren wir per Induktion nach n eine Orthonormalbasis von $\mathbb{C}^n$. Für $n = 1$ ist das charakteristische Polynom linear, und der einzige Koeeffizient der 1×1-Matrix A ist auch Eigenwert von A. Der Vektor 1 ist Eigenvektor und bildet eine Orthonormalbasis von $\mathbb{C}^1$. Sei nun $n > 1$. Da $\mathbb{C}$ algebraisch abgeschlossen ist, besitzt A eine Eigenwert λ_1 und einen zugehörigen Eigenvektor $v_1 \in \mathbb{C}^n$. Ohne Einschränkung ist $\|v_1\| = 1$. Nun gilt wegen Lemma 8.6.11, dass $A(v_1^\perp) \subseteq v_1^\perp$ ist. Das heißt, A induziert eine lineare Abbildung auf dem $(n-1)$-dimensionalen Unterraum $v_1^\perp$. Induktiv können wir annehmen, dass es eine (Orthonormal-)basis $(v_2, v_3, \ldots, v_n)$ von $v_1^\perp$ aus Eigenvektoren von A gibt zu Eigenwerten $\lambda_2, \lambda_3, \ldots, \lambda_n$. Für jeden der paarweise verschiedenen Eigenwerte liefert die Orthonormalisierung nach Gram-Schmidt 7.10.1 eine Orthonormalbasis des zugehörigen Eigenraums. Die Eigenräume zu verschiedenen Eigenwerten stehen paarweise aufeinander senkrecht nach Lemma 8.6.10. Dies liefert eine Orthonormalbasis von $\mathbb{C}^n$ aus Eigenvektoren von A.

Angenommen $\mathbb{C}^n$ besitzt eine Orthonormalbasis $(v_1, v_2, \ldots, v_n)$ aus Eigenvektoren von A. dann gilt für die Matrix Q mit den Spalten $v_1, v_2, \ldots, v_n$, dass $Q^{-1}AQ = Q^*AQ$ eine Diagonalmatrix ist. Nach Satz 7.14.13 ist Q unitär.

Es sei $Q \in \mathrm{U}_n\,\mathbb{C}$, so dass $D = Q^{-1}AQ$ eine Diagonalmatrix ist. Also ist $A = QDQ^*$ normal wegen Lemma 8.6.6, da D als Diagonalmatrix normal ist. □

Satz 8.6.13 (Hauptsatz über hermitesche Matrizen). *Für $A \in \mathbb{C}^{n\times n}$ sind äquivalent:*

(i) Die Matrix A ist hermitesch.
(ii) Die Matrix A ist normal, und alle Eigenwerte sind reell.
(iii) Es existiert eine unitäre Matrix $Q \in U_n\mathbb{C}$, so dass Q^*AQ eine reelle Diagonalmatrix ist.

Beweis. Sei A hermitesch. Dann ist A normal, und alle Eigenwerte sind reell wegen Proposition 8.6.8.

Sei A ist normal mit der Eigenschaft, dass alle Eigenwerte reell sind. Dann folgt aus Satz 8.6.12, dass es eine unitäre Matrix Q gibt, so dass Q^*AQ eine Diagonalmatrix aus den Eigenwerten von A ist.

Angenommen, es existiert eine unitäre Matrix $Q \in U_n\mathbb{C}$, so dass $D = Q^*AQ$ eine reelle Diagonalmatrix ist. Dann ist D hermitesch und damit auch $A = QDQ^*$. □

Beispiel 8.6.14. Die Matrix $A = \left(\begin{smallmatrix} 1 & i \\ -i & 1 \end{smallmatrix}\right) \in \mathbb{C}^{2\times 2}$ is hermitesch, und ihr charakteristisches Polynom ist

$$\chi_A(t) \;=\; t^2 - 2t = \; t(t-2)\,.$$

Es sind also 0 und 2 die beiden Eigenwerte von A. Die jeweils eindimensionalen Eigenräume werden von Zugehörige Eigenvektoren sind

$$\begin{pmatrix} -i \\ 1 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} i \\ 1 \end{pmatrix}$$

aufgespannt. Es gilt $\|\left(\begin{smallmatrix} -i \\ 1 \end{smallmatrix}\right)\| = \|\left(\begin{smallmatrix} i \\ 1 \end{smallmatrix}\right)\| = \sqrt{2}$. Für die unitäre Matrix

$$Q \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$$

gilt $Q^*AQ = \operatorname{diag}(0,2)$.

Lemma 8.6.15. *Sei $A \in \mathbb{R}^{n\times n}$ eine Matrix, deren (komplexe) Eigenwerte sämtlich reell sind. Dann ist A genau dann über $\mathbb{C}$ diagonalisierbar, wenn A über $\mathbb{R}$ diagonalisierbar ist.*

Beweis. Sei A eine reelle Matrix, deren Eigenwerte sämtlich reell sind, und die über $\mathbb{C}$ diagonalisierbar ist. Es sei λ ein Eigenwert von A. Da λ reell ist, ist auch die Matrix $A - \lambda E_n$ reell. Der Rang von $A - \lambda E_n$ über $\mathbb{R}$ ist derselbe, wie der Rang über $\mathbb{C}$. Daher ist die geometrische Vielfachheit von λ über $\mathbb{C}$ dieselbe wie die geometrische Vielfachheit von λ über $\mathbb{R}$. Die algebraische Vielfachheit von λ über $\mathbb{C}$ ist dieselbe wie die algebraische Vielfachheit von λ über $\mathbb{R}$. Die Behauptung folgt nun aus Satz 8.3.13. □

Satz 8.6.16 (Hauptsatz über reelle symmetrische Matrizen). *Für $A \in \mathbb{R}^{n\times n}$ sind äquivalent:*

(i) Die Matrix A ist symmetrisch.
(ii) Der $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ hat eine Orthonormalbasis aus Eigenvektoren von A.
(iii) Es existiert eine orthogonale Matrix $Q \in \mathrm{O}_n\,\mathbb{R}$, so dass $Q^{\mathrm{tr}}AQ$ eine (reelle) Diagonalmatrix ist.

Beweis. Sei A reell symmetrisch. Dann ist A normal und daher wegen Satz 8.6.12 über $\mathbb{C}$ diagonalisierbar. Mit Proposition 8.6.8 sind sämtliche Eigenwerte von A reell. Also lässt sich Lemma 8.6.15 anwenden, und A ist auch reell diagonalisierbar. Wie im Beweis zu Satz 8.6.12 zeigt man, dass $\mathbb{R}^n$ eine Orthonormalbasis aus Eigenvektoren von A besitzt.

Die übrigen Beweisschritte lassen sich ebenfalls unter Spezialisierung auf den reellen Fall aus dem Beweis zu Satz 8.6.12 übernehmen. □

Beispiel 8.6.17. Die Matrix $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ ist symmetrisch. Das charakteristische Polynom ist

$$\chi_A(t) \;=\; \begin{vmatrix} -t & 1 \\ 1 & -t \end{vmatrix} \;=\; t^2 - 1 \;=\; (t+1)(t-1)\,.$$

Die beiden Eigenwerte sind 1 und -1. Zugehörige Eigenvektoren sind

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Beide Vektoren haben euklidische Norm $\sqrt{2}$. Die Matrix

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

ist orthogonal, und es gilt $Q^{\mathrm{tr}}AQ = \mathrm{diag}(1, -1)$.

Lemma 8.6.18. *Sei λ ein Eigenwert der invertierbaren Matrix $A \in \mathrm{GL}_n\,K$ über einem beliebigen Körper K, und es sei v ein zugehöriger Eigenvektor. Dann ist $\lambda \neq 0$ und λ^{-1} ist Eigenwert von A, und v ein zugehöriger Eigenvektor.*

Beweis. Es gilt $Av = \lambda v$, also

$$A^{-1}Av \;=\; v \;=\; \lambda A^{-1}v\,.$$

Daher ist $\lambda \neq 0$ und $A^{-1}v = \lambda^{-1}v$. □

Satz 8.6.19 (Hauptsatz über unitäre Matrizen). *Für $A \in \mathbb{C}^{n\times n}$ sind äquivalent:*

(i) Die Matrix A ist unitär.
(ii) Die Matrix A ist normal, und alle Eigenwerte haben den Betrag 1.
(iii) Es existiert eine unitäre Matrix $Q \in U_n\mathbb{C}$, so dass Q^*AQ eine Diagonalmatrix ist, deren Einträge alle Betrag 1 haben.

Beweis. Sei A unitär. Dann ist A normal. Sei λ Eigenwert von A mit zugehörigem Eigenvektor v. Nach Lemma 8.6.18 ist λ^{-1} Eigenwert von A^{-1} mit Eigenvektor v. Nach Lemma 8.6.7 ist $\overline{\lambda}$ Eigenwert von A^*, ebenfalls mit dem Eigenvektor v. Da aber $A^{-1} = A^*$ gilt, folgt

$$\lambda^{-1}v \;=\; A^{-1}v \;=\; A^*v \;=\; \overline{\lambda}v\,.$$

Wegen $v \neq 0$ ist $\lambda^{-1} = \overline{\lambda}$, also $|\lambda| = \lambda\overline{\lambda} = 1$.

Die beiden anderen Implikationen folgen wie oben. □

Beispiel 8.6.20. Die Matrix $A = 1/\sqrt{2}\left(\begin{smallmatrix} i & -i \\ 1 & 1 \end{smallmatrix}\right) \in \mathbb{C}^{2\times 2}$ ist unitär. Für

$$B \;:=\; 2\sqrt{2}A \;=\; \begin{pmatrix} 2i & -2i \\ 2 & 2 \end{pmatrix}.$$

ist

$$\chi_B(t) \;=\; t^2 - (2+2i)t + 8i \;=\; (t - (1+\sqrt{3} + (1-\sqrt{3})i))(t - (1-\sqrt{3} + (1+\sqrt{3})i)\,.$$

Das heißt die beiden Eigenwerte sind $(1 \pm \sqrt{3}) + (1 \mp \sqrt{3})i$. Es gilt

$$|1 \pm \sqrt{3} + (1 \mp \sqrt{3})i| \;=\; 8 \;=\; (2\sqrt{2})^2\,.$$

Bemerkung 8.6.21. Die Eigenwerte unitärer Matrizen sind komplexe Zahlen vom Betrag 1. Solche komplexe Zahlen lassen sich eindeutig schreiben als

$$e^{i\varphi} \;=\; \cos\varphi + i\sin\varphi$$

für $\varphi \in [0, 2\pi)$. Für jede unitäre Matrix $A \in \mathrm{U}_n\,\mathbb{C}$ gibt es also eine zweite unitäre Matrix Q, so dass $Q^{-1}AQ = \operatorname{diag}(e^{i\varphi_1}, e^{i\varphi_2}, \ldots, e^{i\varphi_n})$ ist.

Bemerkung 8.6.22. Sei $A \in \mathbb{R}^{n\times n}$ reell und λ Eigenwert von A mit Eigenvektor v. Dann folgt aus $Av = \lambda v$ durch Konjugieren

$$A\overline{v} \;=\; \overline{A}\overline{v} \;=\; \overline{Av} \;=\; \overline{\lambda v} \;=\; \overline{\lambda}\overline{v}\,.$$

Das heißt, dass $\overline{\lambda}$ ebenfalls Eigenwert von A ist und $\overline{v}$ ein zugehöriger Eigenvektor. Ist v selbst nicht reell, so sind v und $\overline{v}$ linear unabhängig.

Satz 8.6.23 (Normalform orthogonaler Matrizen). *Es sei $A \in \mathrm{O}_n\,\mathbb{R}$ reell orthogonal. Dann gibt es eine zweite orthogonale Matrix $Q \in \mathrm{O}_n\,\mathbb{R}$, so dass*

$$Q^{-1}AQ \;=\; \begin{pmatrix} P & & \\ & N & \\ & & D \end{pmatrix}$$

gilt, wobei $P = \operatorname{diag}(1, 1, \ldots, 1) \in \mathbb{R}^{p\times p}$, $N = \operatorname{diag}(-1, -1, \ldots, -1) \in \mathbb{R}^{q\times q}$ und

$$D \;=\; \begin{pmatrix} \boxed{\begin{matrix} \cos\gamma_1 & -\sin\gamma_1 \\ \sin\gamma_1 & \cos\gamma_1 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \cos\gamma_r & -\sin\gamma_r \\ \sin\gamma_r & \cos\gamma_r \end{matrix}} \end{pmatrix} \in \mathbb{R}^{2r\times 2r}$$

mit $\gamma_1, \gamma_2, \ldots, \gamma_r \in (0, \pi)$ und $p + q + 2r = n$.

Hier ist p die (algebraische und geometrische) Vielfachheit des Eigenwerts 1 und q die (algebraische und geometrische) Vielfachheit des Eigenwerts -1.

Beweis. Als reelle orthogonale Matrix ist A auch (komplex) unitär. Damit gelten die Konsequenzen aus Satz 8.6.19. Insbesondere haben alle (komplexen) Eigenwerte von A den Betrag 1. Die nicht-rellen Eigenwerte treten in komplex-kunjugierten Paaren auf. Insgesamt können wir die Eigenwerte anordnen wie folgt:

$$\underbrace{1, 1, \ldots, 1}_{p}, \quad \underbrace{-1, -1, \ldots, -1}_{q}, \quad \underbrace{e^{i\gamma_1}, e^{-i\gamma_1}, e^{i\gamma_2}, e^{-i\gamma_2}, \ldots, e^{i\gamma_r}, e^{-i\gamma_r}}_{2r}$$

für $\gamma_k \in (0, \pi)$. Hierbei kann natürlich $p = 0$, $q = 0$ oder $r = 0$ auftreten.

Da A unitär ist, besitzt $\mathbb{C}^n$ eine Orthonormalbasis von Eigenvektoren zu A. Nach Lemma 8.6.15 existiert insbesondere eine Orthonormalbasis $(u_1, u_2, \ldots, u_p)$ aus reellen Vektoren des Eigenraums zum Eigenwert 1. Ebenso existiert eine reelle Orthonormalbasis $(v_1, v_2, \ldots, v_q)$ des Eigenraums zum Eigenwert -1.

Angesichts der Bemerkung 8.6.22 wissen wir, dass es konjugiert-komplexe Eigenvektoren w_k und $\overline{w_k}$ zu den Eigenwerten $e^{i\gamma_k}$ und $e^{-i\gamma_k}$ gibt, so dass insgesamt

$$u_1, u_2, \ldots, u_p, \quad v_1, v_2, \ldots, v_q, \quad w_1, \overline{w_1}, w_2, \overline{w_2}, \ldots, w_r, \overline{w_r}$$

eine Basis des komplexen Vektorraums $\mathbb{C}^n$ bildet. Ohne Einschränkung können wir annehmen, dass $\|w_k\| = \|\overline{w_k}\| = 1$ gilt.

Wir konstruieren die reellen Vektoren

$$x_k := \frac{1}{\sqrt{2}}(w_k + \overline{w_k}) \quad \text{und} \quad y_k := \frac{i}{\sqrt{2}}(w_k - \overline{w_k}),$$

und man stellt fest, dass $\langle x_k, y_k \rangle = 0$ ist sowie $\|x_k\| = \|y_k\| = 1$. Das heißt aber wegen Lemma 8.6.10, dass

$$u_1, u_2, \ldots, u_p, \quad v_1, v_2, \ldots, v_q, \quad x_1, y_1, x_2, y_2, \ldots, x_r, y_r \tag{38}$$

eine Orthonormalbasis von $\mathbb{R}^n$ ist. Wie in Beispiel 8.3.4 ist die orthogonale Matrix, die durch Einschränkung auf den zweidimensionalen invarianten Unterraum $\operatorname{lin}_{\mathbb{R}}(x_k, y_k)$ entsteht, eine Drehung mit der Matrix

$$\begin{pmatrix} \cos\gamma_k & -\sin\gamma_k \\ \sin\gamma_k & \cos\gamma_k \end{pmatrix}$$

bezüglich der Basis (x_k, y_k). Die gesuchte Matrix Q entsteht spaltenweise aus den Vektoren (38). Dies beendet den Beweis. □

KAPITEL 9

Quadratische Formen

9.1. Definitionen und Beispiele

Sei K ein beliebiger Körper, in dem $2 := 1 + 1 \neq 0$ gilt. Man sagt, die *Charakteristik* des Körpers K ist ungleich zwei. Sei V ein K-Vektorraum.

Definition 9.1.1. Eine Abbildung $Q : V \to K$ heißt *quadratische Form*, wenn gilt:

(i) Für alle $\lambda \in K$ und $v \in V$ ist $Q(\lambda v) = \lambda^2 Q(v)$.

(ii) Die Abbildung

$$\beta_Q \;:\; V \times V \to K \;:\; (v, w) \mapsto \frac{1}{2}(Q(v + w) - Q(v) - Q(w))$$

ist eine symmetrische Bilinearform, genannt die *Polarform* zu Q.

Lemma 9.1.2. *Ist $f : V \times V \to K$ eine symmetrische Bilinearform, so ist*

$$Q_f \;:\; V \to K \;:\; v \mapsto f(v, v)$$

eine quadratische Form, und es gilt

$$\beta_{Q_f} \;=\; f\,.$$

Beweis. Seien $v, w \in V$. Dann ist

$$\begin{aligned}
\frac{1}{2}(Q_f(v + w) - Q_f(v) - Q_f(w)) \;&=\; \frac{1}{2}(f(v + w, v + w) - f(v, v) - f(w, w)) \\
&=\; \frac{1}{2}(f(v, v) + f(v, w) + f(w, v) + f(w, w) - f(v, v) - f(w, w)) \\
&=\; \frac{1}{2}(f(v, w) + f(w, v)) \;=\; f(v, w)\,.
\end{aligned}$$

□

Beispiel 9.1.3. Sei $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ das euklidische Skalarprodukt auf $\mathbb{R}^n$. Dann ist

$$Q_{\langle\cdot,\cdot\rangle} \;:\; \mathbb{R}^n \to \mathbb{R} \;:\; v \mapsto \langle v, v\rangle = \|v\|^2$$

die zugehörige quadratische Form.

Bemerkung 9.1.4. Vergleiche Lemma 9.1.2 mit den Polarsierungsidentitäten 7.4.6.

Bemerkung 9.1.5. Sei f eine symmetrische Bilinearform auf K^n. Dann bezeichnet $[f]$ die Matrix von f bezüglich der Standardbasis von K^n. Für $v \in K^n$ gilt

$$Q_f(v) \;=\; f(v, v) \;=\; v^{\mathrm{tr}} \cdot [f] \cdot v\,.$$

Beispiel 9.1.6. Sei $q = \sum_{1 \leq i \leq j \leq n} \alpha_{ij} t_i t_j \in K[t_1, t_2, \ldots, t_n]$ ein homogenes Polynom vom Grad 2. Dann ist

$$Q \;:\; K^n \to K \;:\; v \mapsto q^*(v) = \sum_{1 \leq i \leq j \leq n} \alpha_{ij} v_i v_j$$

eine quadratische Form.

Aufgabe 9.1.7. Jede quadratische Form auf K^n entsteht wie in Beispiel 9.1.6.

9.2. Hauptachsentransformation

Es sei Q eine quadratische Form auf $\mathbb{R}^n$.

SATZ 9.2.1. *Es existiert eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^n$ (bezüglich des euklidischen Skalarprodukts), so dass die Matrix der Polarform β_Q bezüglich $\mathcal{B}$ Diagonalgestalt* $\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *hat.*

BEWEIS. Es sei $A = [\beta_Q]$ die Matrix von β_Q bezüglich der Standardbasis von $\mathbb{R}^n$. Weil β_Q eine symmetrische Bilinearform ist, ist die Matrix A symmetrisch. Nach Satz 8.6.16 existiert eine orthogonale Transformation $S \in O_n\mathbb{R}$, so dass $S^{-1}AS$ Diagonalgestalt hat. Die Spalten von S bilden die gesuchte Orthonormalbasis $\mathcal{B}$. $\square$

DEFINITION 9.2.2. Die Menge $\{v \in V : Q(v) = 1\}$ heißt *Quadrik* zur quadratischen Form Q.

Wir benutzen weiter die Notation von Satz 9.2.1. Wegen Aufgabe 9.1.7 ist eine Quadrik eine (affine) algebraische Hyperfläche im Sinne von Definition 6.8.2. Ist $A = [\beta_Q] = (a_{ij})$ die (symmetrische) Matrix zur Polarform von Q, so erfüllt ein Punkt $x \in \mathbb{R}^n$ auf der Quadrik die Gleichung

$$\sum_{i=1}^{n} a_{ii}x_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij}x_ix_j \; = \; 1\,. \tag{39}$$

Es sei $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ die Orthogonalbasis von $\mathbb{R}^n$ aus Satz 9.2.1. Schreiben wir $y = [x]_{\mathcal{B}} = S^{-1}x$, so gilt

$$\sum_{i=1}^{n} \lambda_i y_i^2 \; = \; 1\,. \tag{40}$$

Die Umwandlung der Gleichung (39) in die Form (40) heißt *Hauptachsentransformation*. Im einzelnen sind hierzu die folgenden Schritte durchzuführen:

(i) Bestimme die Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_n$ der Matrix A mit ihren algebraischen Vielfachheiten.
(ii) Für jeden der paarweise verschiedenen Eigenwerte λ bestimme eine Basis des Eigenraums $\ker(A - \lambda E_n)$ durch Gauß-Jordan-Elimination.
(iii) Konstruiere aus jeder dieser Basen eine Orthonormalbasis mit dem Gram-Schmidt-Verfahren. Die Vereinigung dieser Orthonormalbasen aller Eigenräume bilden eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^n$. Die Transformationsmatrix S wird spaltenweise aus $\mathcal{B}$ gebildet.

Danach gilt $S^{-1}AS = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) =: D$ und

$$Q(x) \; = \; x^{\mathrm{tr}}Ax \; = \; (Sy)^{\mathrm{tr}}ASy \; = \; y^{\mathrm{tr}}(S^{-1}AS)y \; = \; y^{\mathrm{tr}}Dy\,.$$

Die Vektoren der Basis $\mathcal{B}$ heißen *Hauptachsen* von Q. Die transformierte Quadrik ist kongruent zur ursprünglichen Quadrik, weil S eine orthogonale Transformation ist.

BEISPIEL 9.2.3. Betrachte die quadratische Form

$$x_1^2 + 6x_1x_2 + x_2^2$$

auf $\mathbb{R}^2$. Die Matrix der Polarform ist $A = \left(\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right)$. Das charakteristische Polynom ist

$$\chi_A(t) \; = \; \begin{vmatrix} 1-t & 3 \\ 3 & 1-t \end{vmatrix} \; = \; (1-t)^2 - 9\,.$$

Die Eigenwerte sind $\lambda_1 = 4$ und $\lambda_2 = -2$. Die normierten Eigenvektoren sind

$$b_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad \text{und} \quad b_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}.$$

Die Transformationsmatrix ist

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix} = \begin{pmatrix}\cos\frac{\pi}{4} & -\sin\frac{\pi}{4}\\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4}\end{pmatrix}.$$

Die Gleichung (40) hat also die Form

$$4y_1^2 - 2y_2^2 = 1\,. \tag{41}$$

Satz 9.2.4 (Trägheitssatz von Sylvester). *Es existiert eine Basis $\mathcal{C}$ von $\mathbb{R}^n$, so dass die Matrix der Polarform β_Q bezüglich $\mathcal{C}$ Diagonalgestalt* $\operatorname{diag}(1,1,\dots,1,-1,-1,\dots,-1,0,0,\dots,0)$ *hat.*

Beweis. Nach Satz 9.2.1 existiert eine (Orthonormal-)basis $\mathcal{B} = (b_1, b_2, \dots, b_n)$ bezüglich, der β_Q die Gestalt $\operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ hat. Setze

$$c_i := \begin{cases} \frac{1}{\sqrt{\lambda_i}} b_i & \text{falls } \lambda_i > 0\,,\\ \frac{1}{\sqrt{-\lambda_i}} b_i & \text{falls } \lambda_i < 0 \text{ und}\\ b_i & \text{falls } \lambda_i = 0\,. \end{cases}$$

Bis auf Umordnung ist $\mathcal{C} = (c_1, c_2, \dots, c_n)$ die gesuchte Basis. □

Beispiel 9.2.5. Führt man das Beispiel 9.2.3 fort, so ergibt sich aus dem Trägheitssatz von Sylvester aus der Gleichung (41) nun die Quadrik mit der Gleichung

$$z_1^2 - z_2^2 = 1\,. \tag{42}$$

Die Transformation, die dem Trägheitssatz von Sylvester zu Grunde liegt ist nicht mehr notwendig kongruent. Die transformierte Quadrik ist aber immer noch ein affines Bild der ursprünglichen Quadrik.

Bemerkung 9.2.6. Eine Anwendung der Hauptachsentransformation in der Bildverarbeitung wird unter

`http://de.wikipedia.org/wiki/Hauptkomponentenanalyse`

diskutiert.

9.3. Klassifikation der Quadriken im ebenen Fall $n = 2$

Mit der Notation wie oben sei $D = \operatorname{diag}(\lambda_1, \lambda_2)$. Die Gleichung (40) hat nach der Hauptachsentransformation die Form

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1\,. \tag{43}$$

Es treten die folgenden Fälle auf:

(i) $\underline{\lambda_1 > 0 \text{ und } \lambda_2 > 0}$: Für $a := \sqrt{1/\lambda_1}$ und $b := \sqrt{1/\lambda_2}$ wird die Gleichung (43) zu

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1\,,$$

und die Quadrik ist die achsenparallele Ellipse mit den *Halbachsen* a und b.

(ii) $\underline{\lambda_1 > 0 \text{ und } \lambda_2 < 0}$: Für $a := \sqrt{1/\lambda_1}$ und $b := \sqrt{-1/\lambda_2}$ wird die Gleichung (43) zu

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} \;=\; 1\,,$$

und die Quadrik ist eine Hyperbel mit den *Asymptoten* $x_2 = \pm\frac{b}{a}x_1$.

(iii) $\underline{\lambda_1 \neq 0 \text{ und } \lambda_2 = 0}$: Die Gleichung (43) vereinfacht sich zu

$$\lambda_1 x_1^2 \;=\; 1\,.$$

Für $\lambda_1 < 0$ existiert keine reelle Lösung. Für $\lambda_1 > 0$ gilt dann $x_1 = \pm\sqrt{1/\lambda_1}$. Die Quadrik besteht aus zwei affinen Geraden parallel zur x_2-Achse durch $(\pm\sqrt{1/\lambda_1}, 0)$.

(iv) $\underline{\lambda_1 \leq 0 \text{ und } \lambda_2 \leq 0}$: Die zugehörige Quadrik ist leer.

Die Quadrik in Beispiel 9.2.3 ist also eine Hyperbel.

9.4. Die hamiltonschen Quaternionen

Die *(hamiltonschen) Quaternionen* bilden die folgende Menge

$$\mathbb{H} \;:=\; \left\{ \begin{pmatrix} w & z \\ -\overline{z} & \overline{w} \end{pmatrix} \;:\; w, z \in \mathbb{C} \right\}$$

von 2×2-Matrizen mit komplexen Koeffizienten. Für $a, b, c, d \in \mathbb{C}$ gilt

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} + \begin{pmatrix} c & d \\ -\overline{d} & \overline{c} \end{pmatrix} \;=\; \begin{pmatrix} a+c & b+d \\ -\overline{b+d} & \overline{a+c} \end{pmatrix}$$

und

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -\overline{d} & \overline{c} \end{pmatrix} \;=\; \begin{pmatrix} ac - b\overline{d} & ad + b\overline{c} \\ -(\overline{ad + b\overline{c}}) & \overline{ac - b\overline{d}} \end{pmatrix}.$$

Das heißt $\mathbb{H}$ ist abgeschlossen unter Addition und Multiplikation, also ein Teilring von $\mathbb{C}^{2\times 2}$. Wir setzen

$$1_{\mathbb{H}} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i_{\mathbb{H}} := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j_{\mathbb{H}} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k_{\mathbb{H}} := \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

Dann gilt

$${i_{\mathbb{H}}}^2 \;=\; {j_{\mathbb{H}}}^2 \;=\; {k_{\mathbb{H}}}^2 \;=\; -1_{\mathbb{H}}$$

sowie

$$\begin{array}{lll} i_{\mathbb{H}} \cdot j_{\mathbb{H}} \;=\; k_{\mathbb{H}}\,, & j_{\mathbb{H}} \cdot k_{\mathbb{H}} \;=\; i_{\mathbb{H}}\,, & k_{\mathbb{H}} \cdot i_{\mathbb{H}} \;=\; j_{\mathbb{H}}\,, \\ j_{\mathbb{H}} \cdot i_{\mathbb{H}} \;=\; -k_{\mathbb{H}}\,, & k_{\mathbb{H}} \cdot j_{\mathbb{H}} \;=\; -i_{\mathbb{H}} \quad \text{und} & i_{\mathbb{H}} \cdot k_{\mathbb{H}} \;=\; -j_{\mathbb{H}}\,. \end{array}$$

Die Matrixalgebra $\mathbb{C}^{2\times 2}$ ist ein $\mathbb{C}$-Vektorraum (der Dimension 4), also auch ein $\mathbb{R}$-Vektorraum (der Dimension 8). Insbesondere ist eine reelle Skalarmultiplikation auf $\mathbb{H}$ definiert.

Proposition 9.4.1. *Die Menge $\mathbb{H}$, versehen mit der Matrixaddition und -multiplikation sowie der reellen Skalarmultiplikation ist eine 4-dimensionale $\mathbb{R}$-Algebra mit Basis $(1_{\mathbb{H}}, i_{\mathbb{H}}, j_{\mathbb{H}}, k_{\mathbb{H}})$.*

Beweis. Übungsaufgabe. □

Jede komplexe Zahl lässt sich eindeutig schreiben als $u + iv$ für $u, v \in \mathbb{R}$. Via

$$u + iv \;\mapsto\; u1_{\mathbb{H}} + vi_{\mathbb{H}}$$

ist $\mathbb{C}$ eingebettet in $\mathbb{H}$. Bezüglich dieser Einbettung gilt dann $1_{\mathbb{H}} = 1$ und $i_{\mathbb{H}} = i$. Im Folgenden lassen wir daher die Indizes weg. Wir haben die Inklusionen

$$\mathbb{R} \;\subset\; \mathbb{C} \;\subset\; \mathbb{H}\,.$$

LEMMA 9.4.2. *Die von* 0 *verschiedenen Quaternionen sind invertierbar. Das heißt es gilt*

$$\mathbb{H} \setminus \{0\} \;\subset\; \mathrm{GL}_2\,\mathbb{C} \;\subset\; \mathrm{GL}_4\,\mathbb{R}\,.$$

BEWEIS. Übungsaufgabe. □

Insgesamt folgt, dass $\mathbb{H}$ alle Eigenschaften eines Körpers besitzt mit Ausnahme der Kommutativität der Multiplikation. Die Quaternionen sind ein Beispiel für einen *Schiefkörper*. Schiefkörper werden auch *Divisionsalgebren* genannt, wobei man hier Acht geben muss: Gelegentlich wird für Divisionsalgebren auch noch auf die Assoziativität der Multiplikation verzichtet.

BEMERKUNG 9.4.3. Man kann zeigen, dass $\mathbb{H}$ bis auf Isomorphie der einzige echte (das heißt: nicht nicht-kommutative) Schiefkörper ist, der $\mathbb{R}$ als zentralen Teilkörper enthält und der endliche Dimension als $\mathbb{R}$-Vektorraum hat..

Jede Quaternion u besitzt eine eindeutige Darstellung als reelle Linearkombination $u_1 + u_2i + u_3j + u_4k$ der Basiselemente $(1, i, j, k)$. Die Abbildung

$$\bar{\cdot} : \mathbb{H} \to \mathbb{H} : u \mapsto \overline{u} = u_1 - u_2i - u_3j - u_4k$$

ist $\mathbb{R}$-linear.

LEMMA 9.4.4. *Für alle* $v, w \in \mathbb{H}$ *gilt* $\overline{vw} = \overline{w}\,\overline{v}$.

BEWEIS. Übungsaufgabe. □

LEMMA 9.4.5. *Die* Normabbildung $N : \mathbb{H} \to \mathbb{R} : u \mapsto u\overline{u}$ *ist eine quadratische Form auf* $\mathbb{R}^4$*, und es gilt*

$$[\beta_N]_{(1,i,j,k)} \;=\; \mathrm{diag}(1, 1, 1, 1)\,.$$

BEWEIS. Sei $u \in \mathbb{H}$ und $\lambda \in \mathbb{R}$. Dann ist

$$N(\lambda u) \;=\; \lambda u \cdot \overline{\lambda u} \;=\; \lambda^2 u\overline{u} \;=\; \lambda^2 N(u)\,.$$

Ferner ist

$$N(u + v) - N(u) - N(v) \;=\; (u + v)(\overline{u + v}) - u\overline{u} - v\overline{v} \;=\; u\overline{v} + v\overline{u}$$

sowie

$$\begin{aligned} \lambda(u + u')\overline{v} + v\overline{\lambda(u + u')} \;&=\; \lambda u\overline{v} + \lambda u'\overline{v} + v\lambda\overline{u} + v\lambda\overline{u'} \\ &=\; \lambda(u\overline{v} + v\overline{u}) + \lambda(u'\overline{v} + v\overline{u'})\,. \end{aligned}$$

Hieraus folgt, dass die Abbildung

$$\beta_N : \mathbb{H} \to \mathbb{R} : (u, v) \mapsto \frac{1}{2}N(u + v) - N(u) - N(v) = \frac{1}{2}(u\overline{v} + v\overline{u})$$

linear im ersten Argument ist. Analog folgt die Linearität im zweiten Argument. Insgesamt ist β_N eine symmetrische Bilinearform, also ist n eine quadratische Form.

Zur Berechnung der darstellenden Matrix von β_N bezüglich $(1, i, j, k)$ rechnet man beispielsweise

$$\frac{1}{2}(i\overline{i} + i\overline{i}) = 1 \quad \text{und} \quad \frac{1}{2}(i\overline{j} + j\overline{i}) = -\frac{1}{2}(ij + ji) = 0\,.$$

□

Die symmetrische Bilinearform β_N auf $\mathbb{R}^4$ ist genau das euklidische Skalarprodukt. Die Abbildung

$$\kappa_u \,:\, \mathbb{H} \to \mathbb{H} \,:\, v \mapsto u^{-1}vu$$

heißt *Konjugation* mit $u \in \mathbb{H} \setminus \{0\}$. Für $u, v, w \in \mathbb{H}$ und $\lambda, \mu \in \mathbb{R}$ gilt

$$\kappa_u(\lambda v + \mu w) \;=\; u^{-1}(\lambda v + \mu w)u \;=\; \lambda u^{-1}vu + \mu u^{-1}wu \;=\; \lambda\kappa_u(v) + \mu\kappa_u(w)\,,$$

das heißt, die Abbildung κ_u ist $\mathbb{R}$-linear. Man rechnet nach, dass κ_u bijektiv ist also, $\kappa_u \in \mathrm{GL}_4\,\mathbb{R}$. Außerdem gilt

$$\kappa_u(1) \;=\; 1$$

für alle $u \in \mathbb{H} \setminus \{0\}$.

Lemma 9.4.6. *Es gilt*

$$\beta_N(\kappa_u(v), \kappa_u(w)) \;=\; \beta_N(v, w)$$

für alle $u \in \mathbb{H} \setminus \{0\}$ und $v, w \in \mathbb{H}$.

Anders ausgedrückt ist κ_u eine orthogonale Abbildung, also $\kappa_u \in \mathrm{O}_4\,\mathbb{R}$.

Beweis. Es gilt

$$\begin{aligned}
2\beta_N(\kappa_u(v), \kappa_u(w)) &= \kappa_u(v)\overline{\kappa_u(w)} + \kappa_u(w)\overline{\kappa_u(v)} \;=\; u^{-1}vu \cdot \overline{u^{-1}wu} + u^{-1}wu \cdot \overline{u^{-1}vu} \\
&= u^{-1} \cdot v \cdot u\overline{u} \cdot \overline{w} \cdot \overline{u}^{-1} + u^{-1} \cdot w \cdot u\overline{u} \cdot \overline{v} \cdot \overline{u}^{-1} \\
&\overset{N(u)\in\mathbb{R}}{=} u^{-1} \cdot v \cdot \overline{w} \cdot u\overline{u} \cdot \overline{u}^{-1} + u^{-1} \cdot w \cdot \overline{v} \cdot u\overline{u} \cdot \overline{u}^{-1} \\
&= u^{-1} \cdot v \cdot \overline{w} \cdot u + u^{-1} \cdot w \cdot \overline{v} \cdot u \\
&= \kappa_u(v\overline{w} + w\overline{v}) \;=\; \kappa_u(2\beta_N(v, w)) \;\overset{\beta_N(v,w)\in\mathbb{R}}{=}\; 2\beta_N(v, w)\,.
\end{aligned}$$

□

Eine Quaternion u heißt *rein*, falls sie Linearkombination von i, j, k ist. Das heißt $u = u_1 + u_2 i + u_3 j + u_4 k$ ist genau dann rein, falls $u_1 = 0$ gilt. Die reinen Quaternionen bilden einen 3-dimensionalen Teilraum des $\mathbb{R}$-Vektorraums $\mathbb{H}$.

Lemma 9.4.7. *Die reinen Quaternionen bleiben unter Konjugation mit $u \in \mathbb{H} \setminus \{0\}$ invariant.*

Beweis. Die Konjugation κ_u ist linear und respektiert das Skalarprodukt β_N. Gleichzeitig gilt $\kappa_u(1) = 1$. Hieraus folgt, dass κ_u das orthogonale Komplement $1^\perp = \mathrm{lin}_{\mathbb{R}}(i, j, k)$, also die reinen Quaternionen in sich abbildet. □

Satz 9.4.8. *Die Abbildung*

$$\kappa \,:\, \{u \in \mathbb{H} :\, N(u) = 1\} \to \mathrm{SO}(\mathrm{lin}_{\mathbb{R}}(i, j, k)) \,:\, u \mapsto \kappa_u$$

ist ein surjektiver Gruppenhomomorphismus mit der Eigenschaft, dass die Faser $\kappa^{-1}(Q)$ einer Drehmatrix $Q \in \mathrm{SO}(\mathrm{lin}_{\mathbb{R}}(i, j, k))$ aus genau einer Quaternion der Norm 1 und ihrer Negativen besteht.

Beweis. Die Quaternionen der Norm 1 bilden eine Untergruppe der multiplikativen Gruppe von $\mathbb{H}$. Dies folgt aus

$$N(uv) \;=\; (uv)(\overline{uv}) \;=\; u \cdot v \cdot \overline{v} \cdot \overline{u} \;=\; N(u)N(v)\,.$$

Die Einschränkung von β_N auf $\mathrm{lin}_{\mathbb{R}}(i, j, k) \cong \mathbb{R}^3$ ist das euklidische Skalarprodukt. Da κ_u bijektiv ist, folgt aus Lemma 9.4.7, dass κ_u in der orthogonalen Gruppe $\mathrm{O}_3\,\mathbb{R}$ liegt. Man rechnet nach, dass für $N(u) = 1$ gilt $\det \kappa_u = 1$. In diesem Fall ist also $\kappa_u \in \mathrm{SO}_3\,\mathbb{R}$.

Für zwei von 0 verschiedene Quaternionen u und v gilt

$$\kappa_{uv}(w) \;=\; (uv)^{-1}w(uv) \;=\; v^{-1}u^{-1}wuv \;=\; \kappa_v(\kappa_u(w)) \;=\; (\kappa_u \circ \kappa_v)(w)\,.$$

Es bleibt (als Übungsaufgabe) zu zeigen, dass die Abbildung $u \mapsto \kappa_u$ eine surjektiv ist, und der Kern zweielementig ist. □

KAPITEL 10

Jordansche Normalform

Gegenstand des Kapitels ist die Frage, wie man entscheidet, ob zwei Matrizen zueinander ähnlich sind. Nach einer Vorüberlegung zu dieser Frage über einem beliebigen Körper diskutieren wir vor allem den komplexen Fall.

10.1. Schursche Normalform

Sei K ein beliebiger Körper.

SATZ 10.1.1 (Schursche Normalform). *Eine Matrix $A \in K^{n\times n}$ ist genau dann ähnlich zu einer oberen Dreiecksmatrix, falls $\chi_A \in K[t]$ in Linearfaktoren zerfällt.*

BEWEIS. Sei A eine obere Dreiecksmatrix mit Diagonalelementen $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Dann gilt

$$\chi_A(t) \;=\; \det(A - tE_n) \;=\; (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)\,.$$

das charakteristische Polynom zerfällt also in Linearfaktoren.

Nehmen wir umgekehrt an, dass χ_A is Linearfaktoren zerfällt. Wir wollen per Induktion nach n beweisen, dass sich A durch Basiswechsel auf obere Dreiecksgestalt bringen lässt. Die Behauptung ist klar für $n = 1$. Sei nun $n \geq 2$. Weil χ_A in Linearfaktoren zerfällt, besitzt A einen Eigenwert $\lambda \in K$ sowie einen zugehörigen Eigenvektor $b \in K^n$. Wähle eine beliebige geordnete Basis $\mathcal{B}$ von K^n, die mit b beginnt. Dann gilt

$$[\varphi_A]_{\mathcal{B}} \;=\; \begin{pmatrix} \lambda & * \\ 0 & A' \end{pmatrix}$$

für eine Matrix $A' \in K^{(n-1)\times(n-1)}$. Es gilt $\chi_A(t) = (t-\lambda)\chi_{A'}(t)$. Insbesondere zerfällt das charakteristische Polynom von A' in Linearfaktoren. Per Induktion lässt sich A' damit auf obere Dreiecksgestalt bringen, und dies beweist die Behauptung. □

BEMERKUNG 10.1.2. Ob man mit oberen oder unteren Dreiecksmatrizen arbeiten möchte, ist reine Geschmachssache.

10.2. Jordanblöcke

DEFINITION 10.2.1. Ein *Jordanblock* oder *elementare Jordanmatrix* ist eine Matrix

$$J_{m,\lambda} \;=\; \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in K^{m\times m}.$$

Beispiel 10.2.2.

$$J_{2,1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad J_{1,1} = (1)$$
$$J_{3,0} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_{2,\lambda} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

Lemma 10.2.3. *Ein Jordanblock $J_{m,\lambda}$ mit $m > 1$ ist nicht diagonalisierbar. Der Eigenraum zum Eigenwert λ ist 1-dimensional.*

Beweis. Es gilt $\chi_{J_{m,\lambda}}(t) = (\lambda - t)^m$, und damit ist λ der einzige Eigenwert von $J_{m,\lambda}$ (mit algebraischer Vielfachheit m). Betrachte das lineare Gleichungssystem

$$(J_{m,\lambda} - \lambda E_m)x = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} x = 0\,. \tag{44}$$

Für $x = (x_1, \ldots, x_m)^{\mathrm{tr}}$ ist

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} x = \begin{pmatrix} x_2 \\ \vdots \\ x_m \\ 0 \end{pmatrix},$$

und wir erhalten $x = (1, 0, \ldots, 0)^{\mathrm{tr}}$ als einzige Lösung von (44) (bis auf Vielfache). □

Lemma 10.2.4. *Für das Minimalpolynom eines Jordanblocks gilt*

$$\mu_{J_{m,\lambda}}(t) = \chi_{J_{m,\lambda}}(t) = (\lambda - t)^m\,.$$

Beweis. Übungsaufgabe. □

Definition 10.2.5. Eine Matrix $A \in K^{n\times n}$ besitzt *Jordansche Normalform*, falls sie eine Blockdiagonalmatrix aus Jordanblöcken ist:

$$A = \begin{pmatrix} J_{m_1,\lambda_1} & & & \\ & J_{m_2,\lambda_2} & & \\ & & \ddots & \\ & & & J_{m_k,\lambda_k} \end{pmatrix}.$$

10.3. Die Jordansche Normalform einer komplexen Matrix

Wir betrachten von nun an den Fall $K = \mathbb{C}$. Ziel dieses Abschnitts ist der Beweis des folgenden Satzes.

Satz 10.3.1. *Sei $V = \mathbb{C}^n$ und $\varphi \in \mathrm{End}_{\mathbb{C}} V$ ein Endomorphismus. Dann existiert eine Basis $\mathcal{B}$ von V, so dass $[\varphi]_{\mathcal{B}}$ Jordansche Normalform besitzt.*

Eine Basis $\mathcal{B}$ wie im Satz 10.3.1 heißt *Jordanbasis* von $\mathbb{C}^n$ für φ. Bevor wir uns dem Beweis zuwenden, einige Folgerungen.

KOROLLAR 10.3.2. *Sei $A \in \mathbb{C}^{n\times n}$. Dann ist A ähnlich zu einer Matrix in Jordanscher Normalform, das heißt, es existiert eine invertierbare Matrix $S \in \mathrm{GL}_n\,\mathbb{C}$, so dass $S^{-1}AS$ in Jordanscher Normalform ist.*

KOROLLAR 10.3.3. *Zwei komplexe Matrizen sind genau dann ähnlich, wenn sie (bis auf Umordnung der Jordanblöcke) dieselbe Jordansche Normalform besitzen.*

Das heißt, ein Algorithmus zur Berechnung der Jordanschen Normalform dient dazu, zu entscheiden, ob zwei Matrizen zueinander ähnlich sind.

10.3.1. Jordanketten und verallgemeinerte Eigenräume. Sei $V = \mathbb{C}^n$, $\varphi \in \mathrm{End}_{\mathbb{C}}\,V$ ein Endomorphismus und λ ein Eigenwert von φ. Weiter sei $\mathcal{B} = (v_1, \dots, v_m, v_{m+1}, \dots, v_n)$ eine Basis von V, bezüglich der gilt

$$[\varphi]_{\mathcal{B}} = \begin{pmatrix} J_{m,\lambda} & 0 \\ 0 & * \end{pmatrix}.$$

Dann gilt

$$\left\{\begin{array}{rcl} \varphi(v_1) &=& \lambda v_1 \\ \varphi(v_2) &=& v_1 + \lambda v_2 \\ \vdots && \\ \varphi(v_m) &=& v_{m-1} + \lambda v_m \end{array}\right. \iff \left\{\begin{array}{rcl} (\varphi - \lambda\,\mathrm{id})v_1 &=& 0 \\ (\varphi - \lambda\,\mathrm{id})v_2 &=& v_1 \\ \vdots && \\ (\varphi - \lambda\,\mathrm{id})v_m &=& v_{m-1} \end{array}\right. . \tag{45}$$

DEFINITION 10.3.4. Eine Familie $(v_1, v_2, \dots, v_m)$ in V heißt *Jordankette* zum Eigenwert λ von φ, falls $v_1 \neq 0$ und (45) erfüllt ist.

LEMMA 10.3.5. *Eine Jordankette zum Eigenwert λ von φ ist linear unabhängig.*

BEWEIS. Beweis per Induktion nach m. Es sei $(v_1, v_2, \dots, v_m)$ eine Jordankette. Für $m = 1$ ist $v_1 \neq 0$, also linear unabhängig. Sei $m \geq 2$. Nach Induktionsvoraussetzung ist $(v_1, v_2, \dots, v_{m-1})$ linear unabhängig. Angenommen, es existieren $\mu_1, \mu_2, \dots, \mu_{m-1}$, so dass $v_m = \mu_1 v_1 + \dots + \mu_{m-1} v_{m-1}$. Dann folgt

$$\begin{aligned} v_{m-1} &= (\varphi - \lambda\,\mathrm{id})v_m = (\varphi - \lambda\,\mathrm{id})(\mu_1 v_1 + \dots + \mu_{m-1} v_{m-1}) \\ &= \mu_1(\varphi - \lambda\,\mathrm{id})(v_1) + \mu_2(\varphi - \lambda\,\mathrm{id})(v_2) + \dots + \mu_{m-1}(\varphi - \lambda\,\mathrm{id})(v_{m-1}) \\ &= \mu_2 v_1 + \mu_3 v_2 + \dots + \mu_{m-1} v_{m-2}\,. \end{aligned}$$

Aber dies steht im Widerspruch dazu, dass $(v_1, v_2, \dots, v_{m-1})$ linear unabhängig ist. □

DEFINITION 10.3.6. Ein Vektor $v \in V \setminus \{0\}$ heißt *verallgemeinerter Eigenvektor* (oder *Hauptvektor*) von φ, falls eine natürliche Zahl $m > 0$ existiert, so dass

$$(\varphi - \lambda\,\mathrm{id})^m v = 0\,. \tag{46}$$

Die kleinste Zahl m, für die (46) erfüllt ist, heißt *Stufe* von v.

Die Hauptvektoren der Stufe 1 sind genau die Eigenvektoren.

Sei v ein verallgemeinerter Eigenvektor von φ der Stufe m. Wir setzen

$$\begin{array}{rcr} v_1 &:=& (\varphi - \lambda\,\mathrm{id})^{m-1}v \\ v_2 &:=& (\varphi - \lambda\,\mathrm{id})^{m-2}v \\ &\vdots& \\ v_{m-1} &:=& (\varphi - \lambda\,\mathrm{id})v \\ v_m &:=& v \end{array} .$$

Dann gilt offenbar (45), und $(v_1, v_2, \dots, v_m)$ ist eine Jordankette, also linear unabhängig nach Lemma 10.3.5. Insbesondere gilt $m \leq n$.

Definition 10.3.7. Die Menge

$$V^\lambda(\varphi) \ := \ \bigcup_{k\in\mathbb{N}} \ker(\varphi - \lambda\,\mathrm{id})^k \ = \ \mathrm{lin}_{\mathbb{C}}(v_1,\dots,v_m)$$

heißt *verallgemeinerter Eigenraum* von φ bezüglich λ.

Man beachte, dass $\ker(\varphi-\lambda\,\mathrm{id})^k \subseteq \ker(\varphi-\lambda\,\mathrm{id})^{k+1}$ für alle $k\in\mathbb{N}$ ist, weil eine lineare Abbildung 0 auf 0 abbildet.

Proposition 10.3.8. *Es seien φ und ψ Endomorphismen, die mit einander vertauschen, das heißt, $\varphi\circ\psi = \psi\circ\varphi$. Dann ist für jeden Eigenwert λ von φ der verallgemeinerte Eigenraum $V^\lambda(\varphi)$ invariant unter ψ. Insbesondere ist $V^\lambda(\varphi)$ invariant unter φ.*

Beweis. Sei $v\in V^\lambda(\varphi)$. Dann existiert $m\in\mathbb{N}$ mit $v\in\ker(\varphi-\lambda\,\mathrm{id})^m$. Es folgt

$$(\varphi-\lambda\,\mathrm{id})^m(\psi(v)) \ = \ \psi\big((\varphi-\lambda\,\mathrm{id})^m(v)\big) \ = \ 0 \ .$$

Also ist $\psi(v)\in V^\lambda(\varphi)$. □

10.3.2. Kurze Vorüberlegungen zum Minimalpolynom. Sei K in diesem Unterabschnitt ein beliebiger Körper. In der Übung wurde der folgende Satz gezeigt:

Satz 10.3.9. *Sei $V=K^n$ und $A\in K^{n\times n}$ eine quadratische Matrix. Seien $p_1,p_2,\dots,p_r\in K[t]$ teilerfremde Polynome mit $p_1\cdot p_2\cdot\dots\cdot p_r=\mu_A$. Dann gilt*

$$V \ = \ \bigoplus_{i=1}^r \ker p_i^*(A)\,.$$

Beispiel 10.3.10. Betrachte die Matrix

$$A \ = \ \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 1 & 3\end{pmatrix} \ \in \ \mathbb{C}^{3\times 3}\,.$$

Das charakteristische Polynom

$$\chi_A(t) \ = \ t^3-5t^2+8t-4 \ = \ (t-1)(t-2)^2$$

stimmt mit dem Minimalpolynom überein. Es gilt

$$\ker(A-E_3) \ = \ \ker\begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 2\end{pmatrix} \ = \ \mathbb{C}\begin{pmatrix}1\\0\\0\end{pmatrix}$$

und

$$\ker((A-2E_3)^2) \ = \ \ker\begin{pmatrix} -1 & 0 & 0\\ 0 & -1 & -1\\ 0 & 1 & 1\end{pmatrix}^2 \ = \ \ker\begin{pmatrix} 1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \ = \ \mathbb{C}\begin{pmatrix}0\\1\\0\end{pmatrix}+\mathbb{C}\begin{pmatrix}0\\0\\1\end{pmatrix}\,.$$

Lemma 10.3.11. *Für*

$$A \ = \ \begin{pmatrix} J_{m,\lambda} & *\\ 0 & *\end{pmatrix} \ \in \ K^{n\times n}$$

teilt das Minimalpolynom $(t-\lambda)^m$ des Jordanblocks $J_{m,\lambda}$ das Minimalpolynom von A.

Beweis. Übungsaufgabe. □

10.3.3. Zerlegung von $\mathbb{C}^n$ entlang der verallgemeinerten Eigenräume. Ab jetzt wieder $K = \mathbb{C}$.

SATZ 10.3.12. *Sei $V = \mathbb{C}^n$ und $\varphi \in \operatorname{End}_{\mathbb{C}} V$ ein Endomorphismus mit den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$. Dann gilt*

$$V \;=\; V^{\lambda_1}(\varphi) \oplus V^{\lambda_2}(\varphi) \oplus \cdots \oplus V^{\lambda_r}(\varphi)\,.$$

BEWEIS. Sei $A = [\varphi]$. Weil $\mathbb{C}$ algebraisch abgeschlossen ist, zerfällt das Minimalpolynom $\mu_A(t) = (t-\lambda_1)^{k_1}(t-\lambda_2)^{k_2}\cdots(t-\lambda_r)^{k_r}$ in Linearfaktoren. Wir setzen $p_i := (t-\lambda_i)^{k_i}$, und wir wollen zeigen, dass

$$\ker p_i^*(A) \;=\; \ker(A - \lambda_i\varphi)^{k_i} \;=\; V^{\lambda_i}(\varphi)$$

gilt. Die Inklusion "$\subseteq$" ist klar nach Definition von $V^{\lambda_i}(\varphi)$. Nehmen wir es gäbe einen Hauptvektor von φ zum Eigenwert λ_i der Stufe $m > k_i$. Sei $(v_1, v_2, \dots, v_m)$ eine zugehörige Jordankette. Diese Familie von Vektoren ist linear unabhängig, lässt sich also zu einer Basis $\mathcal{B}$ von V ergänzen. Dann gilt

$$[\varphi]_{\mathcal{B}} \;=\; \begin{pmatrix} J_{m,\lambda_i} & * \\ 0 & * \end{pmatrix}.$$

Nach Lemma 10.3.11 teilt das Minimalpolynom des Jordanblocks J_{m,λ_i} das Minimalpolynom von A. Hieraus entsteht der gewünschte Widerspruch, da $\mu_{J_{m,\lambda}} = (t-\lambda_i)^m$ das Minimalpolynom $\mu_A(t)$ eben nicht teilt. Aus Satz 10.3.9 folgt nun die Behauptung. □

BEMERKUNG 10.3.13. $\dim V^{\lambda_i}(\varphi) =$ algebraische Vielfachheit von λ_i.

LEMMA 10.3.14. *Sei $\mathcal{C} = (v_1^1, \dots, v_{k_1}^1, v_1^2, \dots, v_{k_2}^2, \dots, v_1^s, \dots, v_{k_s}^s)$ eine Familie von s Jordanketten zum Eigenwert λ von φ, das heißt,*

$$(\varphi - \lambda\,\mathrm{id})(v_{j+1}^i) = v_j^i \text{ für } 1 \le j < k_i \quad \text{und} \quad (\varphi - \lambda\,\mathrm{id})(v_1^i) = 0\,.$$

Sind die Eigenvektoren $v_1^1, v_1^2, \dots, v_1^s$ linear unabhängig, so sind alle Vektoren in $\mathcal{C}$ linear unabhängig.

BEWEIS. Sei $\sum_{i,j} \alpha_{ij} v_j^i = 0$ für $\alpha_{ij} \in \mathbb{C}$. Angenommen, es existieren i, j, mit $\alpha_{ij} \neq 0$. Sei k das Maximum aller Stufen, für die ein Index i existiert mit $\alpha_{ik} \neq 0$. Dann folgt

$$\begin{aligned} 0 \;&=\; (\varphi - \lambda\,\mathrm{id})^{k-1}\Big(\sum_{i,j} \alpha_{ij} v_j^i\Big) \;=\; \sum_{i,j} \alpha_{ij} (\varphi - \lambda\,\mathrm{id})^{k-1}(v_j^i) \\ &=\; \sum_{i=1}^s \alpha_{ik} (\varphi - \lambda\,\mathrm{id})^{k-1}(v_k^i) \;=\; \sum_{i=1}^s \alpha_{ik} v_1^i\,. \end{aligned}$$

Dies ist ein Widerspruch, da die Vektoren $v_1^1, v_1^2, \dots, v_1^s$ linear unabhängig sind. □

LEMMA 10.3.15. *Sei $\mathcal{C}$ wie oben, aber linear abhängig. Dann existiert eine Familie $\mathcal{C}'$ von Jordanketten mit $\operatorname{lin}(\mathcal{C}) = \operatorname{lin}(\mathcal{C}')$, die einen Vektor weniger enthält.*

BEWEIS. Aus Lemma 10.3.14 folgt, dass die Eigenvektoren $v_1^1, v_1^2, \dots, v_1^s$ linear abhängig sind. Das heißt, es existiert eine nicht-triviale Linearkombination $\sum_{i=1}^s \alpha_i v_1^i = 0$ mit $\alpha_i \neq 0$. Der Index i sei so gewählt, dass die zugehörige Jordankette $v_1^i, \dots, v_{k_i}^i$ minimale Länge hat unter all denen mit $\alpha_i \neq 0$. Ohne Einschränklung sei $i = 1$.

Fall 1: $k_1 = 1$. Streiche v_1^1 aus $\mathcal{C}$, um $\mathcal{C}'$ zu erhalten.

Fall 2: $k_1 \ge 2$. Setze

$$\tilde{v}_j^1 \;:=\; v_{j+1}^1 + \sum_{\alpha_i \neq 0} \frac{\alpha_i}{\alpha_1} v_{j+1}^i$$

für $1 \le j < k_1$. Wegen der Minimalität von k_1 gilt $k_1 \le k_i$. Damit ist die Existenz von v^i_{j+1} gesichert. Setze $\mathcal{C}' = (\tilde{v}^1_1, \dots, \tilde{v}^1_{k_1-1}, v^2_1, \dots, v^2_{k_2}, \dots, v^s_1, \dots, v^s_{k_s}$.
Es ist zu zeigen, dass $(\tilde{v}^1_1, \dots, \tilde{v}^1_{k_1-1})$ eine Jordankette ist, und dass $\mathrm{lin}(\mathcal{C}) = \mathrm{lin}(\mathcal{C}')$ ist. □

Satz 10.3.16. *Für jeden Eigenwert λ von φ besitzt der verallgemeinerte Eigenraum $V^\lambda(\varphi)$ eine* Jordanbasis, *also eine Familie von Jordanketten, die eine Basis von $V^\lambda(\varphi)$ bilden.*

Beweis. Sei $(u_1, u_2, \dots, u_l)$ eine Basis von $V^\lambda(\varphi)$. Jeder Vektor u_i ist ein verallgemeinerter Eigenvektor, zum Beispiel der Stufe k_i. Dies liefert eine Jordankette der Länge k_i. Die Vereinigung dieser l Jordanketten ergibt eine Familie von Jordanketten, die $V^\lambda(\varphi)$ erzeugen. Durch (wiederholtes Anwenden) von Lemma 10.3.15 erhalten wir eine Basis aus Jordanketten. □

10.3.4. Verfahren zur Bestimmung der Jordanschen Normalform. Sei $V = \mathbb{C}^n$ und $\varphi \in \mathrm{End}_{\mathbb{C}} V$. Wir wollen eine Basis J von V bestimmen, so dass $[\varphi]_J$ Jordansche Normalform besitzt.

(i) Bestimme die Eigenwerte $\lambda_1, \dots, \lambda_l$ von φ mit ihren algebraischen Vielfachheiten $k_1, v \dots, k_k$. Es gilt $k_1 + \cdots + k_l = n$.
(ii) Für jeden Eigenwert λ_i: Bestimme eine Basis des verallgemeinerten Eigenraums $V^{\lambda_i}(\varphi)$. Dazu löst man schrittweise die linearen Gleichungssysteme
$$(\varphi - \lambda_i \,\mathrm{id})^j v \;=\; 0$$
für $j = 1, 2, \dots$ bis man k_i linear unabhängige Lösungen v gefunden hat.
(iii) Bilde Jordanketten und verkürze sie schrittweise durch Anwendung von Lemma 10.3.15 bis man eine Basis erhält.
(iv) Die Matrix des Basiswechsels besitzt als Spalten die verallgemeinerte Eigenvektoren (kettenweise aufsteigend!).

Beispiel 10.3.17. Betrachte
$$A \;=\; \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \end{pmatrix} \;\in\; \mathbb{C}^{5\times 5} \,.$$

(i) $\chi_A(t) = (1-t)^5$, das heißt, 1 ist der einzige Eigenwert, und er hat die algebraische Vielfachheit 5. Es folgt, dass $\mathbb{C}^5 = V^1(A)$ ist.
(ii) Wähle als Basis von $\mathbb{C}^5$:
$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\; v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\; v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},\; v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix},\; v_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$
Es gilt $(A - E_5)v_1 = w$ und $(A - E_5)v_2 = -w$ für
$$w = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \end{pmatrix},$$

und w ist ein Eigenvektor. Also sind v_1 und v_2 Hauptvektoren der Stufe 2. Außerdem gilt $(A - E_5)v_3 = (A - E_5)v_4 = v_5$ und $(A - E_5)v_5 = 0$. Das heißt, v_5 ist Eigenvektor, und v_3 und v_4 sind Hauptvektoren der Stufe 2.

(iii) Wir analysieren die Jordanketten im einzelnen.

▷ Die von v_1 bzw. v_2 beginnenden Jordanketten sind linear abhängig. Setze:

$$\tilde{v}_1 \;=\; v_1 + \frac{1}{1}v_2 \;=\; \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Dann ist $(A - E_5)\tilde{v}_1 = 0$, das heißt, $\tilde{v}_1$ ist ein Eigenvektor.

▷ Die von v_3 und v_4 beginnenden Jordanketten sind ebenfalls linear abhängig. Setze:

$$\tilde{v}_3 \;=\; v_3 - \frac{1}{1}v_4 \;=\; \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}.$$

Wiederum ist $\tilde{v}_3$ ein Eigenvektor.

Wir haben jetzt folgendes System von Jordanketten:

$$\tilde{v}_1,\; v_2 \longrightarrow -w,\; \tilde{v}_3,\; v_4 \longrightarrow v_5\,, \tag{47}$$

und es gilt

$$-w - v_5 - \frac{1}{2}\tilde{v}_3 + \frac{1}{2}\tilde{v}_1 \;=\; 0\,.$$

Weiter stellt sich heraus, dass $(-w, v_2, v_5, v_4, \tilde{v}_3)$ linear unabhängig ist. Durch Streichen des Vektors $\tilde{v}_1$ in (47) erhalten wir also die Jordanbasis

$$v_2 \longrightarrow -w,\; \tilde{v}_3,\; v_4 \longrightarrow v_5\,.$$

(iv) Insgesamt erhalten wir die Transformationsmatrix

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

und $S^{-1}AS = J_{2,1} \oplus J_{2,1} \oplus J_{1,1}$.

KAPITEL 11

Extras

In diesem Kapitel stellen wir noch kurz einige weitere Konzepte aus der Linearen Algebra vor.

11.1. Dualräume

Sei V eine Vektorraum über einem beliebigen Körper K.

DEFINITION 11.1.1. Der Vektorraum

$$V^* \;:=\; \operatorname{Hom}_K(V, K)$$

der Linearformen heißt *Dualraum* von V.

PROPOSITION 11.1.2. *Falls* $\dim_K V < \infty$*, so ist* V^* *isomorph zu* V *als* K*-Vektorraum.*

Sei V endlichdimensional, und es sei $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ eine Basis von $V \cong K^n$. Wir definieren Linearformen $b_i^* \in V^*$ durch

$$b_i^*(b_j) \;=\; \delta_{ij}\,.$$

PROPOSITION 11.1.3. *Die Familie* $\mathcal{B}^* := (b_1^*, b_2^*, \ldots, b_n^*)$ *von Linearformen ist eine Basis von* V^**, die zu* $\mathcal{B}$ duale Basis*.*

11.2. Tensorprodukte

Seien V und W Vektorräume über K. Betrachte den Vektorraum

$$F(V \times W) \;:=\; \left\{ \sum_{i=1}^{n} \alpha_i e_{(v_i \times w_i)} \,:\, n \in \mathbb{N}, \alpha_i \in K, (v_i \times w_i) \in V \times W \right\}$$

der formalen K-Linearkombinationen von Elementen in $V \times W$. Auf dieser Menge führen wir eine Relation ein durch

$$\begin{aligned} e_{(v_1+v_2)\times w} &\sim e_{v_1\times w} + e_{v_2\times w}\,, \\ e_{v\times(w_1+w_2)} &\sim e_{v\times w_1} + e_{v\times w_2} \quad \text{und} \\ e_{(\alpha v)\times w} &\sim \alpha e_{v\times w} \;\sim\; e_{v\times(\alpha w)}\,, \end{aligned} \tag{48}$$

für alle $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ und $\alpha \in K$. Diese Relation ist offensichtlich reflexiv. Ihr symmetrischer und transitiver Abschluss ist eine Äquivalenzrelation, die wir wiederum mit $\sim$ bezeichnen.

DEFINITION 11.2.1. Der Quotient

$$V \otimes W \;:=\; F(V \times W)/\sim$$

heißt *Tensorprodukt* von V und W.

Die Elemente des Tensorprodukts schreiben wir als

$$v \otimes w \ := \ [e_{(v,w)}]_\sim \,.$$

Das Tensorprodukt hat die natürliche Struktur eines K-Vektorraums, und die Relationen (48) werden zu den Gleichungen

$$\begin{aligned} (v_1 + v_2) \otimes w \ &= \ v_1 \otimes w + v_2 \otimes w \\ v \otimes (w_1 + w_2) \ &= \ v \otimes w_1 + v \otimes w_2 \\ (\alpha v) \otimes w \ &= \ \alpha(v \otimes w) \ = \ v \otimes (\alpha w) \,. \end{aligned}$$

Proposition 11.2.2. *Seien V und W endlich-dimensional mit $\dim_K V = n$ und $\dim_K W = m$. Dann ist $\dim_K(V \otimes W) = mn$.*

Beweis. Seien $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ und $\mathcal{C} = (c_1, c_2, \ldots, c_m)$ Basen von V bzw. W. Dann ist

$$\{b_i \otimes c_j \,:\, 1 \le i \le n,\, 1 \le j \le m\}$$

eine Basis von $V \otimes W$. □

Bemerkung 11.2.3. Es gilt

$$K^n \otimes K^m \ = \ \mathrm{Hom}_K(K^n, K^m) \ = \ K^{m \times n} \ = \ K^{mn} \,.$$

Betrachte die K-bilineare Abbildung

$$\tau \,:\, V \times W \to V \otimes W \,:\, (v, w) \mapsto v \otimes w \,.$$

Satz 11.2.4 (Universelle Eigenschaft). *Für jedes Paar (X, φ), wobei X ein K-Vektorraum ist und $f : V \times W \to X$ eine K-bilineare Abbildung, existiert eine eindeutige lineare Abbildung $\varphi : V \otimes W \to X$ mit der Eigenschaft*

$$\varphi \circ \tau \ = \ f \,.$$

Beweis. Setze $\varphi(v \otimes w) := f(v, w)$. □

Die Eindeutigkeit der Abbildung φ erzwingt, dass $V \otimes W$ der bis auf Isomorphie einzige K-Vektorraum mit dieser Eigenschaft ist.

Bemerkung 11.2.5. Kategorientheoretisch ausgedrückt ist die Abbildung τ das terminale Objekt in der Kategorie der bilinearen Abbildungen von $V \times W$.