Inhaltsverzeichnis

Vorlesungsmanuskript
Lineare Algebra 5.01.051 im Modul mat050
Wintersemester 2018/2019

Universitäts-Prof. Dr. Andreas Stein
Raum W1 2-216, Institut für Mathematik
Carl-von-Ossietzky Universität Oldenburg, D-26111 Oldenburg
Telefon: (0441) 798-3232, Email: `andreas.stein@uni-oldenburg.de`

Empfohlene Literatur (Deutsch)

- G. Fischer: „Lineare Algebra: Eine Einführung für Studienanfänger (Grundkurs Mathematik)“, Springer Spektrum, 18. Auflage 2014.
- G. Fischer: „Lineare Algebra: Lernbuch Lineare Algebra und Analytische Geometrie: Das Wichtigste ausführlich für das Lehramts- und Bachelorstudium“, Springer Spektrum, 2012.
- S. Bosch: Lineare Algebra, Springer Spektrum, 5. Auflage 2014.
- B. Huppert und W. Willems, Lineare Algebra: Mit zahlreichen Anwendungen in Kryptographie, Codierungstheorie, Mathematischer Physik und Stochastischen Prozessen, Vieweg + Teubner Verlag, 2. Auflage 2010.
- A. Beutelspacher: Lineare Algebra: Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen, Springer Spektrum; 8. Auflage 2013.
- M. Koecher: Lineare Algebra und analytische Geometrie, Springer, 4. Auflage 2013.
- H.-J. Kowalsky, G. Michler: Lineare Algebra, de Gruyter, 2003.
- F. Lorenz: Lineare Algebra I+II, Spektrum-Verlag, 2008.

Empfohlene Literatur (Englisch)

- H. Anton, Chris Rorres: Elementary Linear Algebra with Applications, Wiley, 2010.
- D. Lay: Linear Algebra and Its Applications, International Edition, Pearson 2012.
- L. Spence, A. Insel, S. Friedberg: Elementary Linear Algebra, Pearson Prentice Hall, International Edition 2013.
- S. Friedberg, A. Insel, L. Spence: Linear Algebra, Pearson Prentice Hall, Int. Ed. 2013.
- S. J. Leon: Linear Algebra with Applications, Global Edition, Pearson 2015.
- S. Lang: Linear Algebra, Springer, 3. Auflage 2013.

Themenübersicht: Einige der folgenden Themen sind optional.

- Grundlegende Techniken und Strukturen, Gruppen, Ringe, Körper.
- Lineare Gleichungssysteme, Matrizenrechnung.
- Vektorräume, Dimension, Basis.
- Lineare Abbildungen, Basiswechsel.
- Determinanten.
- Eigenwerte und Eigenvektoren, Diagonalisierung.
- Vektorräume mit Skalarprodukt, spezielle Normalformen.

Beschreibungen:

- **Modulbeschreibung**: Es wird eine Einführung in die wesentlichen Ideen und Methoden der linearen Algebra gegeben. Die Studierenden lernen, mathematische Denk- und Arbeitsweisen anzuwenden und grundlegende mathematische Beweisprinzipien zu beherrschen.

- **Meine Meinung**: Neben der Analysis I zählt die Lineare Algebra zu einem der wichtigsten Grundlagenkurse der Reinen Mathematik. Die Anwendungen sind so zahlreich, dass es schwer ist, sie alle aufzuzählen. Einige Anwendungen werden in der Vorlesung zumindest erwähnt, wenn nicht sogar genauer besprochen. Zudem wird ein gewisser Abstraktionsgrad angesetzt, der für das weitere STudium fundamental ist.

- **Meine Philosophie**: In einem solch typischen Kurs der Algebra können Studierende frühzeitig ihre Tendenzen und Begabungen feststellen. Ich werde deshalb den Erstsemesterkurs in Linearer Algebra so interessant wie möglich machen. Unterrichten ist für mich in jedem Semester immer wieder eine neue Herausforderung. Mit mir zusammen wird hoffentlich allen neuen Studenten gelingen, den Sprung vom Gymnasium zur Universität zu bewältigen.

0 Einführung und Motivation

Zuerst taucht die Frage auf, was der Term "Lineare Algebra" eigentlich bedeutet. Welche mathematische Themen besprochen werden, haben wir bereits aufgelistet. Allgemein versteht man unter "Linearer Algebra" die Theorie der Vektorräume und linearen Abbildungen (Homomorphismen). Wichtig ist dabei das Lösen linearer Gleichungssysteme und die Einführung der Determinantenfunktion.

Die Lineare Algebra hat für Studienanfänger auch andere Aufgaben. Es werden die Grundlagen für viele weiterführende Vorlesungen geschaffen. Die wichtigen Themen werden nämlich in fast allen Bereichen der Naturwissenschaften und Wirtschaftswissenschaften benötigt.

0.1 Mengen

Definition Eine **Menge** A ist die Gesamtheit wohlbestimmter und unterschiedlicher Objekte, und die Objekte heißen **Elemente**. Ist a ein Element aus A, so schreiben wir $a \in A$, andernfalls $a \notin A$.

Wir verlangen, dass eine Menge mit einer präzisen Beschreibung gegeben ist, die uns entscheiden lässt, ob ein gegebenes Objekt darin enthalten ist oder nicht. Mengen können durch Aufzählung der Elemente, eine Bildungsvorschrift oder durch Angabe von Eigenschaften definiert werden.

Wichtige Beispiele und Definitionen

- $\varnothing = \{\}$, die **leere Menge**, d.h. die eindeutige Menge ohne Elemente.
- $\mathbb{N} := \{1, 2, 3, \ldots\}$, die **Menge der natürlichen Zahlen**.
- $\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\} = \mathbb{N} \cup \{0\}$.
- $\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$, die **Menge der ganzen Zahlen**.
- $\mathbb{Q} := \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}$, die **Menge der rationalen Zahlen**.
- Die **Menge der reellen Zahlen** (exakte Definition in der Analysis 1) ist
$$\mathbb{R} := \left\{ r = \sum_{j=-\infty}^{m} a_j 10^j \mid m \in \mathbb{Z}, a_j \in \mathbb{N}_0, 0 \leq a_j \leq 9 \right\} .$$
- $\mathbb{C} := \{ z = x + yi \mid x, y \in \mathbb{R} \}$ die **Menge der komplexen Zahlen**. Dabei haben wir $i^2 = -1$. Man identifiziert $\mathbb{C}$ mit der Menge $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ unter vernünftigen Operationen.

Definition Falls eine Menge A leer ist oder aus nur endlich vielen Elementen besteht, so nennen wir A eine **endliche Menge**. Ist A eine endliche Menge, so nennen wir die Anzahl ihrer Elemente auch die **Mächtigkeit** oder die **Kardinalität** von A, geschrieben $|A|$. Ist A nicht endlich, so ist A eine **unendliche Menge**.

Eine **endliche Menge** A ist entweder gleich der leeren Menge, $A = \varnothing$, oder es existiert ein $n \in \mathbb{N}$, so dass $A = \{a_1, a_2, \ldots, a_n\}$.

Konstruktion von neuen Mengen aus einer gegebenen Menge

Definition Seien A und B zwei Mengen.

(a) B heißt eine **Teilmenge** von A, geschrieben $B \subseteq A$, wenn jedes Element von B auch zu A gehört.

(b) B heißt eine **echte Teilmenge** von A, geschrieben $B \subset A$, wenn $B \subseteq A$ und $B \neq A$ gilt. Wir schreiben auch: $B \subsetneq A$.

(c) Die **Potenzmenge** von A ist $P(A) := \{B \mid B \subseteq A\}$, also die Menge aller Teilmengen von A.

Definition Seien A und B zwei Mengen. Wir nennen A und B **gleich**, geschrieben $A = B$, wenn A und B dieselben Elemente besitzen. Ist dies nicht der Fall, so schreiben wir $A \neq B$.

Operationen mit Mengen

Seien zwei Mengen A und B gegeben. Die folgenden Verknüpfungen kann man wie in der Schule mittels Venndiagrammen darstellen.

- $A \cup B := \{ x \mid x \in A \text{ oder } x \in B \}$ = **Vereinigung** von A und B, also die Menge aller Elemente, die in A oder in B oder in beiden enthalten sind.
- $A \cap B := \{ x \mid x \in A \text{ und } x \in B \}$ = **Durchschnitt** von A und B, also die Menge aller Elemente, die sowohl in A als auch in B enthalten sind.
- $A \setminus B := \{ x \mid x \in A \text{ und } x \notin B \}$ = **Differenzmenge** oder **Differenz** von A und B, also die Menge aller Elemente, die in A, aber nicht in B enthalten sind.
- $A \times B = \{ (x, y) \mid x \in A \text{ und } y \in B \}$ = **Produktmenge** oder oder **Produkt** oder **kartesisches Produkt** von A und B, also die Menge aller **Paare** (x, y) die mit Elementen $x \in A$ und $y \in B$ möglich sind.

Ist I eine beliebige Indexmenge, so definieren wir für eine Familie von Mengen $(A_i)_{i \in I}$ die folgenden Verknüpfungen:

- $\bigcup_{i \in I} A_i = \{ x \mid \exists i \in I \text{ mit } x \in A_i \}$ = **Vereinigung** von $(A_i)_{i \in I}$.
- $\bigcap_{i \in I} A_i = \{ x \mid x \in A_i \text{ für alle } i \in I \}$ = **Durchschnitt** von $(A_i)_{i \in I}$.
- $\prod_{i \in I} A_i = \{ (a_i)_{i \in I} \mid a_i \in A_i \text{ für alle } i \in I \}$ = **kartesisches Produkt** von $(A_i)_{i \in I}$.

Ist speziell $I = \{1, 2, \ldots, n\}$, so schreiben wir

- $\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n$.
- $\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n$.
- $\prod_{i=1}^{n} A_i = A_1 \times A_2 \times \ldots A_n = \{ (a_1, \ldots, a_n) \mid a_i \in A_i \text{ für alle } i = 1, \ldots, n \}$.

Ist sogar $A = A_1 = A_2 = \ldots = A_n$, so schreiben wir

$$A^n := \prod_{i=1}^{n} A_i = A \times A \times \ldots A = \{ (a_1, \ldots, a_n) \mid a_i \in A \text{ für alle } i = 1, \ldots, n \} .$$

Zum Beispiel erhalten wir wie üblich

$$\mathbb{R}^n := \{ (x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ für alle } i = 1, \ldots, n \} .$$

0.2 Logik und elementare Beweise

Mathematische Aussagen und quantifizierte Aussagen

Definition Eine **(mathematische) Aussage** ist eine Kombination von Wörtern und mathematischen Symbolen, der man eindeutig den Wahrheitswert **wahr** (w, true) oder **falsch** (f, false) zuordnen kann. Wir schreiben P , Q , R etc. für mathematische Aussagen.

Definition Ist P eine Aussage, so auch ihre **Negation**, geschrieben $\neg P$. Bei $\neg P$ wird der Wahrheitswert von P umgedreht. Man sagt auch **nicht** P und schreibt nicht(P) .
Ist P wahr, so ist $\neg P$ falsch und umgekehrt.

Definition Eine **Aussagenform** ist eine Kombination von Wörtern und mathematischen Symbolen, die eine oder mehrere Variablen enthält. Wir fordern, dass durch Einsetzen von Werten für jede der auftretenden Variablen die Aussagenform zu einer Aussage wird. Wir schreiben $P(X)$, $Q(X, Y)$, $R(X, Y, n)$ etc. für Aussagenformen.
Sei S eine Menge, welche die erlaubten Werte für die Variable X enthält, und sei $P(X)$ eine Aussagenform in der Variablen X .

(a) **Allquantor/Universalquantor:** $\forall$, „für alle”, „Falls $X \in S$”

„Für alle $X \in S$ ist $P(X)$ wahr.” liest sich in Kurzform

$$\forall X \in S : P(X) \quad \text{bzw.} \quad (\forall X \in S)\, P(X)$$

(b) **Existenzquantor:** $\exists$, „es gibt”, „mindestens ein $x \in S$”

„Es gibt ein $X \in S$, so dass $P(X)$ wahr ist.” liest sich in Kurzform

$$\exists X \in S : P(X) \quad \text{bzw.} \quad (\exists X \in S)\, P(X)$$

Die Klammerung wird oft unterdrückt.

Kombination von quantifizierten Aussagen: Sei S eine Menge, welche die erlaubten Werte für die Variable X enthält, sei T eine Menge, welche die erlaubten Werte für die Variable Y enthält, und sei $Q(X, Y)$ eine Aussagenform in den Variablen X , Y .

$\forall X \in S \;\; \exists Y \in T : Q(X, Y)$ Für alle $X \in S$ gibt es ein $Y \in T$, so dass $Q(X, Y)$ wahr ist.

$\exists X \in S \;\; \forall Y \in T : Q(X, Y)$ Es gibt ein $X \in S$, so dass für alle $Y \in T$ $Q(X, Y)$ wahr ist.

Wir schreiben auch

$$(\forall X \in S)(\exists Y \in T)\, Q(X, Y) \text{ und } (\exists X \in S)(\forall Y \in T)\, Q(X, Y) \; .$$

Wir versuchen, folgender Konvention zu folgen:

- Quantoren erscheinen geordnet am Anfang einer Aussage bzw. einer Aussagenform. Speziell in Kurzformen versuchen wir, diese Regel strikt einzuhalten.
- In einer Kombination von Quantoren gehen wir von links nach rechts vor. Dabei werden weiter links stehende Variablen danach als Konstanten behandelt.
- Die Aussage „$\exists X \in S \;\; \forall Y \in T : \; Q(X, Y)$" ist als stärker anzusehen als die Aussage „$\forall X \in S \;\; \exists Y \in T : Q(X, Y)$".

Negation von Quantoren:

- Die Negation des Allquantors in Kurzform, also $\neg(\forall X \in S : \; P(X))$ ist gleichbedeutend mit
$$\exists X \in S : \neg P(X) \quad \text{bzw.} \quad (\exists X \in S)\,(\neg P(X)) \;.$$
- Die Negation des Existenzquantors in Kurzform, also $\neg(\exists X \in S : \; P(X))$ ist gleichbedeutend mit
$$\forall X \in S : \neg P(X) \quad \text{bzw.} \quad (\forall X \in S)\;(\neg P(X)) \;.$$
- In einer Kombination von Quantoren gehen wir von links nach rechts vor. Für $\neg(\forall X \in S \;\; \exists Y \in T : \; Q(X, Y))$ erhalten wir
$$\exists X \in S \;\; \forall Y \in T : \neg Q(X, Y) \;.$$

Verknüpfung mit Junktoren

Wie in der deutschen Sprache kann man logische Aussagen auch verknüpfen, man muss jedoch mit der Bedeutung vorsichtig sein. Z.B. hat das mathematische "oder" eine andere Bedeutung, da es nicht ausschließlich ist. Zur Erinnerung hier die wichtigsten Konstrukte. Seien P und Q mathematische Aussagen. Abhängig vom Wahrheitswert dieser Aussagen, kann man die Verknüpfungen in einer Wahrheitstafel angeben.

Konjunktion: $P \wedge Q$, „P und Q" ist wahr, wenn beide wahr sind.

Disjunktion: $P \vee Q$, „P oder Q" ist wahr, wenn mindestens eine der beiden Aussagen wahr ist.

Implikation: $P \Rightarrow Q$, „P impliziert Q" ist nur falsch, wenn P wahr und Q falsch ist.

Äquivalenz: $P \Leftrightarrow Q$, „P ist äquivalent zu Q" ist wahr, wenn P und Q denselben Wahrheitswert besitzen.

P	Q	$\neg P$	$P \vee Q$	$P \wedge Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
w	w	f	w	w	w	w
w	f	f	w	f	f	f
f	w	w	w	f	w	f
f	f	w	f	f	w	w

Wichtige Tautologien: Mit Wahrheitstafeln erhalten wir folgende bekannte **Tautologien**, die sich durch Betrachtung der Wahrheitswerte elementarer Aussagen ergeben:

(a) $P \quad \Leftrightarrow \quad \neg(\neg P)$

(b) $(P \Rightarrow Q) \quad \Leftrightarrow \quad (\neg Q \Rightarrow \neg P)$

(c) $(P \Rightarrow Q) \quad \Leftrightarrow \quad \neg(P \wedge (\neg Q))$

(d) $(P \Leftrightarrow Q) \quad \Leftrightarrow \quad ((P \Rightarrow Q) \wedge (Q \Rightarrow P))$

(e) De Morgansche Gesetze:

   (i) $\neg(P \wedge Q) \quad \Leftrightarrow \quad (\neg P) \vee (\neg Q)$

   (ii) $\neg(P \vee Q) \quad \Leftrightarrow \quad (\neg P) \wedge (\neg Q)$

(f) Seien A und B zwei Mengen. Dann gilt immer:

$$A = B \quad \Leftrightarrow \quad [(A \subseteq B) \wedge (B \subseteq A)] \ .$$

Elementare Beweistechniken für $P \Rightarrow Q$

Zum Beweis der Aussage $P \Rightarrow Q$ können alle der drei folgenden Strategien eingeschlagen werden, da sie logisch äquivalent sind:

(a) **Direkter Beweis:** $P \Rightarrow Q$. **Idee:** Wir nehmen die Voraussetzung P an und beweisen mit einer Folge von Implikationen, dass Q wahr ist.

(b) **Kontraposition:** $(\neg Q) \Rightarrow (\neg P)$. **Idee:** Wir nehmen die Voraussetzung $\neg Q$ an und beweisen mit einer Folge von Implikationen, dass $\neg P$ wahr ist.

(c) **Indirekter Beweis/Widerspruchsbeweis:** $\neg(P \wedge \neg Q)$. **Idee:** Wir nehmen die Voraussetzung P sowie die Negation von Q an und führen diese Annahmen zu einem Widerspruch.

Elementare Beweistechniken „Gegenbeispiel“

Sei S eine Menge, welche die erlaubten Werte für die Variable x enthält, und sei $P(x)$ eine Aussagenform in der Variablen x. Die Negation der allquantifizierten Aussage $\forall x \in S : P(x)$ ist wie gesehen $\exists x \in S : \neg P(x)$.

Idee: Um zu zeigen, dass die Aussage $\forall X \in S : P(X)$ falsch ist, müssen wir lediglich einen Wert $x_0 \in S$ finden, so dass $P(x_0)$ falsch ist.

0.3 Abbildungen

Wir definieren hier, was eine Abbildung ist. Dies ist nicht zu verwechseln mit dem wesentlich stärkeren Begriff der **linearen Abbildung**. Eine lineare Abbildung ist insbesondere eine Abbildung.

Definition Eine **Abbildung** oder **Funktion** von einer Menge A in eine Menge B ist eine Vorschrift

$$\begin{aligned} f: \ A &\to B \\ a &\mapsto f(a) = b \ , \end{aligned}$$

derart dass jedem $a \in A$ genau ein $b \in B$ zugeordnet wird. In Kurzform:

$$\forall a \in A \ \exists! b \in B : \ f(a) = b \ .$$

Mit Klammern:

$$(\forall a \in A)(\exists! b \in B) \ [f(a) = b] \ .$$

A heißt der **Definitionsbereich von** f und B der **Bild-** oder **Wertebereich** von f. Eine Funktion $f : A \to B$ ist also in folgendem Sinne **wohldefiniert**:

(i) Für jedes $a \in A$ ist $f(a)$ definiert und $f(a) \in B$.

(ii) $\forall a_1, a_2 \in A : (a_1 = a_2 \Rightarrow f(a_1) = f(a_2))$.

Definition Sei $f : A \to B$ eine Abbildung.

(a) Dann heißt

$$\mathrm{Bild}(f) = f(A) = \mathrm{Im}(f) := \{b \in B \mid \exists a \in A \ \text{ mit } \ f(a) = b\}$$

das **Bild von** f und A heißt das **Urbild von** f.

(b) Für $A' \subseteq A$ heißt

$$f(A') := \{b \in B \mid \exists a' \in A' \text{ derart dass } f(a') = b\}$$

das **Bild von** A'.

(c) Für $B' \subseteq B$ heißt

$$f^{-1}(B') = \{a \in A \mid f(a) = b' \in B'\}$$

das **Urbild von** B'.

(d) Wir schreiben für die Menge aller Abbildungen von A nach B

$$\mathrm{Abb}(A, B) = B^A := \{f : A \to B \mid f \text{ Abbildung}\} \ .$$

Definition Sei $f : A \to B$ eine Abbildung.

(a) f ist genau dann **injektiv**, wenn zu jedem Element $b \in B$ höchstens ein Element $a \in A$ existiert, so dass $f(a) = b$. Wir nennen f dann eine **Injektion**. Kurzform:

$$\forall a_1, a_2 \in A : (f(a_1) = f(a_2) \Rightarrow a_1 = a_2) .$$

Dies ist logisch äquivalent zu

$$\forall a_1, a_2 \in A : (a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)) .$$

(b) f ist genau dann **surjektiv**, wenn es zu jedem $b \in B$ ein Element $a \in A$ gibt, so dass $f(a) = b$. Wir nennen f dann eine **Surjektion**. Kurzform:

$$\forall b \in B \ \exists a \in A : f(a) = b .$$

(c) f ist genau dann **bijektiv**, wenn f injektiv und surjektiv ist. D.h. zu jedem $b \in B$ gibt es genau ein $a \in A$, so dass $f(a) = b$. Wir nennen f eine **Bijektion**. Kurzform:

$$\forall b \in B \ \exists! a \in A : f(a) = b .$$

Wir erinnern nun an die Verknüpfung von kompatiblen Abbildung. Sind $f : A \to B$ und $g : B' \to C$ zwei Abbildungen, so dass $B \subseteq B'$, so sehen wir die Abbildungen als kompatibel an. Der Funktionswert $g(f(a))$von g ist dann für jeden Wert von $a \in A$ definiert. Der Einfachheit halber betrachten wir nur den Fall, dass $B' = B$.

Definition Seien $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Wir definieren die **Komposition** oder die **Verknüpfung** von g mit f, geschrieben $g \circ f$, als die Abbildung

$$\begin{aligned} g \circ f : \ A &\to C \\ a &\mapsto g \circ f(a) = g(f(a)) . \end{aligned}$$

Bemerkung 0.1. *Sei $f : A \to B$ eine Abbildung. Dann gelten die folgenden Regeln:*

(a) $f \circ \mathrm{id}_A = f$ *und* $\mathrm{id}_B \circ f = f$.

(b) *f ist injektiv $\Leftrightarrow$ $\exists$ Abbildung $h : B \to A$, derart dass $h \circ f = \mathrm{id}_A$.*

(c) *f ist surjektiv $\Leftrightarrow$ $\exists$ Abbildung $g : B \to A$, derart dass $f \circ g = \mathrm{id}_B$.*

(d) *f ist bijektiv $\Leftrightarrow$ $\exists$ Abbildung $f^{-1} : B \to A$, so dass $f \circ f^{-1} = \mathrm{id}_B \wedge f^{-1} \circ f = \mathrm{id}_A$.*

(e) *Sind A und B **endliche** Mengen mit gleich vielen Elementen, so sind die Begriffe äquivalent, d.h.*

$$f \ \textit{injektiv} \ \Leftrightarrow \ f \ \textit{surjektiv} \ \Leftrightarrow \ f \ \textit{bijektiv} .$$

Eigenschaften von Bijektionen:

(a) Falls $f : A \to B$ eine Bijektion ist, dann ist auch $f^{-1} : B \to A$ eine Bijektion, genannt die Inverse von f. Für alle $a \in A$, $b \in B$ gilt:

$$f^{-1}(b) = a \quad \Leftrightarrow \quad f(a) = b \ .$$

(b) Sei $f : A \to B$ eine Abbildung. Falls es eine Abbildung $g : B \to A$ gibt, so dass $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$, so ist f eine Bijektion und g die Inverse von f.

(c) **Abzählen durch Bijektion:** Sei A eine endliche Menge. Falls es eine endliche Menge B und eine Bijektion $f : A \to B$ gibt, so ist

$$|A| = |B| \ .$$

0.4 Lineare Gleichungssysteme – Einführende Beispiele

Eine **lineare Gleichung in n Variablen** $x_1, \ldots, x_n$ ist von der Form

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b \, , \tag{0.1}$$

wobei $a_1, \ldots, a_n, b$ gegebene Elemente sind. In dieser Vorlesung sind dies Elemente eines Körpers, so dass man dividieren darf. Körper werden wir im nächsten Kapitel genau einführen.

Die Gleichung (0.1) besitzt über den reellen und komplexen Zahlen immer eine Lösung, es sei denn $a_1 = a_2 = \ldots = a_n = 0$ und $b \neq 0$. Diesen Fall nennt man manchmal entartet. Sucht man nach Lösungen über den reellen Zahlen und ist man nicht im entarteten Fall, so beschreibt (0.1) für $n = 2$ eine Gerade in der xy-Ebene.
Ein **lineares Gleichungssystem von m Gleichungen in n Variablen** $x_1, \ldots, x_n$ besteht aus m linearen Gleichungen wie in (0.1), ist also von der Form

$$\left\{ \begin{array}{cccccccccccc} a_{11} & x_1 & + & a_{12} & x_2 & + & \ldots & + & a_{1n} & x_n & = & b_1 \\ a_{21} & x_1 & + & a_{22} & x_2 & + & \ldots & + & a_{2n} & x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & x_1 & + & a_{m2} & x_2 & + & \ldots & + & a_{mn} & x_n & = & b_m \end{array} \right\} \tag{0.2}$$

wobei a_{ij}, b_i für alle $i = 1, \ldots, m$ und $j = 1, \ldots, n$ gegebene feste Körperelemente sind.

Eine **Lösung** von (0.2) ist ein n-dimensionaler Spaltenvektor

$$k = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} \, ,$$

wobei eingesetzt für $x_1 = k_1, \ldots, x_n = k_n$ sämtliche m Gleichungen erfüllt sind.

Wir schreiben auch einfach $x = (x_1, \ldots, x_n)$ für eine Lösung, wenn klar ist, dass damit nicht die Variablen gemeint sind.

Satz 0.2. *Seien $a, b, c, d, r, s \in \mathbb{R}$ gegeben, die das folgende lineare Gleichungssystem von 2 Gleichungen in den beiden Variablen x, y bestimmen:*

$$\left\{ \begin{array}{ccccccc} a & x & + & b & y & = & r \\ c & x & + & d & y & = & s \end{array} \right\} \tag{0.3}$$

Dann gilt:

(a) *Falls $ad - bc \neq 0$ ist, so existiert eine eindeutige Lösung von* (0.3)*.*

(b) *Falls $ad - bc = 0$ ist, so gibt es entweder unendlich viele Lösungen von* (0.3) *oder* (0.3) *ist unlösbar.*

Beweis. Übungsblatt bzw. selbst nachrechnen mit Fallunterscheidung $a = b = c = d = 0$ oder o.B.d.A. $d \neq 0$. ■

Dieses Resultat kann für einen beliebigen Körper K anstatt $\mathbb{R}$ verallgemeinern.

1 Gruppen, Ringe, Körper

Wir werden in diesem Kapitel die Grundstrukturen Gruppen, Ringe und Körper präzise einführen. Gruppen werden genauer in der Vorlesung Algebra II untersucht, während Ringe besonders in der Vorlesung Algebra I diskutiert werden. Für den Rest dieser Vorlesung werden Körper eine wichtige Rolle spielen. Der Vollständigkeit halber werden wir auch Halbgruppen und Monoide einführen, welche jedoch für den Rest der Vorlesung nur eine untergeordnete Rolle spielen werden.

1.1 Gruppen

Definition Eine nichtleere Menge G zusammen mit einer Verknüpfung $*$, d.h. einer Abbildung

$$\textbf{(G0)} \qquad \begin{aligned} * : G \times G &\to G \\ (a, b) &\mapsto a * b \ , \end{aligned}$$

mit anderen Worten: $a * b \in G$ für alle $a, b \in G$, heißt eine **Halbgruppe**, wenn folgende Bedingung erfüllt ist:

(G1) Assoziativität: $(a * b) * c = a * (b * c)$ für alle $a, b, c \in G$.

Wir schreiben $(G, *)$ für eine Halbgruppe. Ein **Monoid** ist eine Halbgruppe $(G, *)$, für die folgende Bedingung erfüllt ist:

(G2) Neutrales Element (zweiseitig): Es gibt ein Element $e \in G$, so dass

$$e * a = a = a * e \text{ für alle } a \in G \ .$$

Eine **Gruppe** ist ein Monoid $(G, *)$, für das folgende Bedingung erfüllt ist:

(G3) Inverses (Invertierbares Element): Für alle $a \in G$ gibt es ein $b \in G$, so dass

$$a * b = e = b * a \ .$$

Eine der drei Strukturen $(G, *)$ nennt man **abelsch** oder **kommutativ**, wenn zusätzlich Folgendes gilt:

(G4) Kommutativität: $a * b = b * a$ für alle $a, b \in G$.

Definition Ein Element $e \in G$ wie in **(G2)** heißt (zweiseitiges) **neutrales Element** von G. Ein Element $b \in G$ wie in **(G3)** heißt **Inverses von** a. Existiert ein solches Element für $a \in G$, so heißt a **invertierbar**.

Notation: Wir schreiben $(G, *)$ oder $(G, \circ)$ für Halbgruppen / Monoide / Gruppen. Ist die binäre Operation bekannt, so schreiben wir einfach G und $a\,b$ für $a * b$. Unsere Schreibweise einer Gruppe wird also multiplikativ sein. Da es additive und multiplikative Strukturen gibt, verwenden wir zur expliziten Unterscheidung die folgende Notation:

(A) Additive Notation: $(G, +)$, $a * b = a + b$, $e = 0$, $b = -a$.

(M) Multiplikative Notation: $(G, \cdot)$, $a * b = a \cdot b = a\,b$, $e = 1$, $b = a^{-1}$.

Die Eindeutigkeit von e, 0, 1, $-a$, a^{-1} folgt später.

Um die Gruppenaxiome **(G0)**-**(G3)** schneller nachzurechnen, genügen folgende Sätze.

Satz 1.1. (**Schwaches Axiomengesetz**) *Sei $(G, *)$ eine Halbgruppe, es seien für $(G, *)$ also* **(G0)** *und* **(G1)** *erfüllt.*

(a) *$(G, *)$ ist genau dann eine Gruppe, wenn neben* **(G0)** *und* **(G1)** *auch die folgenden Bedingungen erfüllt sind:*

(G2l) *Linksneutrales Element: Es gibt ein Element $e \in G$, so dass*

$$e * a = a \quad \text{für alle } a \in G \text{ gilt.}$$

(G3l) *Linksinverses: Für alle $a \in G$ gibt es ein Element $b \in G$, so dass*

$$b * a = e \ .$$

(b) *$(G, *)$ ist genau dann eine Gruppe, wenn neben* **(G0)** *und* **(G1)** *auch die folgenden Bedingungen erfüllt sind:*

(G2r) *Rechtsneutrales Element: Es gibt ein Element $e \in G$, so dass*

$$a * e = a \quad \text{für alle } a \in G \text{ gilt.}$$

(G3r) *Rechtsinverses: Für alle $a \in G$ gibt es ein Element $b \in G$, so dass*

$$a * b = e \ .$$

Es genügt also, eine der folgenden Kombinationen nachzuprüfen:

- entweder **(G0)**, **(G1)**, **(G2l)**, **(G3l)**
- oder **(G0)**, **(G1)**, **(G2r)**, **(G3r)**

Bemerkung 1.2.

(a) *Es ist verboten die Linksaxiome mit den Rechtsaxiomen zu mischen. Würde man neben* **(G0)** *und* **(G1)** *für eine Halbgruppe* **(G2l)** *und* **(G3r)** *oder* **(G2r)** *und* **(G3l)** *fordern, so ist dies nicht ausreichend, um eine Gruppenstruktur nachzuweisen.*

(b) *Satz 1.1 besagt nicht, dass* **(G2l)** *und* **(G2r)** *äquivalent sind und auch nicht, dass* **(G3l)** *und* **(G3r)** *äquivalent sind. Zum Beispiel impliziert die Existenz eines Linksinversen nicht die Existenz eines Rechtsinversen einer Halbgruppe $\neq$ Gruppe und umgekehrt.*

Wir betrachten nun elementare Eigenschaften und Vereinfachungen.

Offensichtliche Verallgemeinerungen durch vollständige Induktion:

(a) **Allgemeines Assoziativgesetz**:
Sei $(G, *)$ eine Halbgruppe und seien $a_1, \ldots, a_n \in G$, $n \in \mathbb{N}$. Dann hängt das Produkt

$$a_1 * a_2 * \ldots * a_n := (\ldots((a_1 * a_2) * a_3) * \ldots) * a_n$$

nicht von der Wahl der Klammerung ab. Jede andere Wahl der Klammerung liefert denselben Wert. Die Aussage gilt insbesondere für Monoide und Gruppen.

(b) **Allgemeines Kommutativgesetz**:
Sei $(G, *)$ eine abelsche Halbgruppe und seien $a_1, \ldots, a_n \in G$. Dann gilt:

$$a_1 * a_2 * \ldots * a_n = a_{\sigma(1)} * a_{\sigma(2)} * \ldots * a_{\sigma(n)} \quad \text{für alle } \sigma \in S_n .$$

Dies gilt insbesondere für abelsche Monoide und abelsche Gruppen.

Von hier an verwenden wir $a\,b$ anstelle von $a * b$ und verzichten auf die Klammerung. Also ist für eine Halbgruppe G und $a_1, \ldots, a_n \in G$:

$$\prod_{j=1}^{n} a_j = a_1 a_2 \cdots a_n .$$

Beziehungsweise

$$\sum_{j=1}^{n} a_j = a_1 + a_2 + \ldots + a_n ,$$

wenn $* = +$ und die Notation additiv ist.

Im Falle $a_j = a$ für $1 \leq j \leq n$ schreibt man

$$a^n := \prod_{j=1}^{n} a_j = \underbrace{a\,a \cdots a}_{n\text{-mal}} .$$

Satz 1.3.

(a) *In einem Monoid ist das neutrale Element eindeutig bestimmt.*

(b) *In einer Gruppe ist das Inverse zu einem Element eindeutig bestimmt.*

Für das eindeutige Inverse b eines Elements $a \in G$ in einer Gruppe G schreiben wir ab sofort $a^{-1} := b$. Das Inverse erfüllt dann

$$a^{-1}a = a\,a^{-1} = e \ .$$

Wir schreiben außerdem für $n \in \mathbb{N}$:

$$a^{-n} := (a^{-1})^n \ .$$

Es gelten in einer Halbgruppe / einem Monoid / einer Gruppe die folgenden Potenzgesetze für alle $a \in G$ und für alle $m, n \in \mathbb{N} / \mathbb{N}_0 / \mathbb{Z}$:

(1) $a^m\, a^n = a^{m+n}$.

(2) $(a^m)^n = a^{mn}$.

In additiver Notation bedeutet dies:

(1) $m\,a + n\,a = (m+n)\,a$.

(2) $n\,(m\,a) = m\,n\,a$.

Satz 1.4. *Sei G eine Gruppe und seien $a, b, c \in G$. Dann gilt:*

(a) $(a^{-1})^{-1} = a$.

(b) $(ab)^{-1} = b^{-1} a^{-1}$.

(c) *Kürzungsregeln:*

$$ac = bc \quad \Rightarrow \quad a = b \ ,$$
$$ca = cb \quad \Rightarrow \quad a = b \ .$$

Definition Sei G eine Gruppe. Eine nichtleere Teilmenge U ist eine **Untergruppe** von G, geschrieben $U \leq G$, wenn U unter der selben Verknüpfung wie G eine Gruppe ist.

Proposition 1.5. (**Untergruppentest I**) *Eine Teilmenge $U \subseteq G$ einer Gruppe G ist genau dann eine Untergruppe, wenn die folgenden Bedingungen erfüllt sind:*

(U1) *Das neutrale Element e von G liegt in U.*

(U2) $a\,b \in U$ *für alle* $a, b \in U$.

(U3) $a^{-1} \in U$ *für alle* $a \in U$.

Proposition 1.6. (**Untergruppentest II**) *Eine Teilmenge $U \subseteq G$ einer Gruppe G ist genau dann eine Untergruppe, wenn die folgenden Bedingungen erfüllt sind:*

(U1) *Das neutrale Element e von G liegt in U.*

(U2') $a\,b^{-1} \in U$ *für alle* $a, b \in U$.

Der Vollständigkeit halber listen wir noch den folgenden Test für endliche Gruppen.

Proposition 1.7. (**Endlicher Untergruppentest**) *Sei G eine Gruppe und sei $U \subseteq G$ eine endliche Teilmenge. Es gilt:*

U ist genau dann eine Untergruppe von G, wenn **(U1)** *und* **(U2)** *erfüllt sind.*

1.2 Die ganzen Zahlen modulo m

Axiom (**Wohlordnungsaxiom**) Jede nichtleere Teilmenge von $\mathbb{N}_0$ (bzw. $\mathbb{N}$) besitzt ein kleinstes Element.

Satz 1.8. (**Division mit Rest in** $\mathbb{Z}$) *Für alle $a \in \mathbb{Z}$ und $b \in \mathbb{N}$, gibt es eindeutige Zahlen $q, r \in \mathbb{Z}$, so dass*

$$a = q\,b + r \quad \textit{und} \quad 0 \leq r < b \ .$$

Sei $m \in \mathbb{N}$, $m > 1$. Wir definieren die **ganzen Zahlen modulo** m durch

$$\mathbb{Z}_m := \{0, 1, \ldots, m-1\} \subseteq \mathbb{Z} \ .$$

Wir definieren im Folgenden zwei Operationen auf $\mathbb{Z}_m$: Seien $a, b \in \mathbb{Z}_m$. Dann existieren nach obigem Satz eindeutige $q_1, q_2, r_1, r_2 \in \mathbb{Z}$, so dass

$$\begin{aligned} a + b &= q_1 m + r_1 \quad \text{und} \quad 0 \leq r_1 \leq m-1 \ , \\ a \cdot b &= q_2 m + r_2 \quad \text{und} \quad 0 \leq r_2 \leq m-1 \ . \end{aligned}$$

Wir setzen

$$\begin{aligned} a +_m b &:= r_1 \in \mathbb{Z}_m \ , \\ a \cdot_m b &:= r_2 \in \mathbb{Z}_m \ . \end{aligned}$$

Mittels Äquivalenzrelation wird in Büchern oft $\mathbb{Z}/m\mathbb{Z}$ anstelle von $\mathbb{Z}_m$ definiert. Prinzipiell erhält man jedoch dasselbe Objekt. Wir operieren direkt auf dem Vertretersystem.

(a) $(\mathbb{Z}_m, +_m)$ ist eine abelsche Gruppe.

(b) $(\mathbb{Z}_m, \cdot_m)$ ist ein abelsches Monoid.

(c) $(\mathbb{Z}_m \setminus \{0\}, \cdot_m)$ ist genau dann eine Gruppe, wenn $m \in \mathbb{P}$. Siehe dazu die Einheitentheorie bei Ringen.

1.3 Ringe

Zu beachten ist, dass wir hauptsächlich kommutative Ringe mit Eins betrachten, d.h. abweichend von einigen Büchern haben unsere Ringe immer eine Eins.

Definition Eine nichtleere Menge R zusammen mit zwei binären Operationen (oder Verknüpfungen), $+$ für Addition und $\cdot$ für Multiplikation, auf R heißt ein **Ring**, wenn folgende Bedingungen erfüllt sind:

(A) $(R, +)$ ist eine abelsche Gruppe.

(M) $(R, \cdot)$ ist ein Monoid, d.h. es gelten die folgenden Axiome:

- **(M0) Abgeschlossenheit:** $a \cdot b \in R$ für alle $a, b \in R$.
- **(M1) Assoziativität:** $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ für alle $a, b, c \in R$.
- **(M2) Einselement/Eins:** Es gibt ein Element $1 \in R$, genannt das Einselement oder die Eins, so dass $1 \cdot a = a \cdot 1 = a$ für alle $a \in R$.

(D) Es gelten die **Distributivgesetze**, d.h. für alle $a, b, c \in R$ gelten:

$$(a+b) \cdot c = a \cdot c + b \cdot c$$

$$a \cdot (b+c) = a \cdot b + a \cdot c \ .$$

Ist $(R, \cdot)$ kommutativ, so heißt $(R, +, \cdot)$ ein **kommutativer Ring**.

Notation: Wir schreiben $(R, +, \cdot)$ oder einfach R und $a \cdot b = a\,b$. Wir schreiben 0 bzw. 1 für das neutrale Element bzgl. der Addition bzw. der Multiplikation.

Proposition 1.9. *Für alle $m \in \mathbb{N}$ ist $(\mathbb{Z}_m, +_m, \cdot_m)$ ein kommutativer Ring.*

Proposition 1.10. (**Elementare Eigenschaften von Ringen**) Sei $(R, +, \cdot)$ ein Ring. Dann gelten die folgenden elementaren Eigenschaften.

(a) In R sind die 0 und die 1 eindeutig bestimmt. Falls $1 = 0$ in R gilt, so ist dies genau dann möglich, wenn $R = \{0\}$. Andernfalls gilt $1 \neq 0$, welches nur möglich ist, wenn $R \neq \{0\}$.

(b) $0 \cdot a = a \cdot 0 = 0$ für alle $a \in R$.

(c) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$ für alle $a, b \in R$.

(d) $(-a) \cdot (-b) = a \cdot b$ für alle $a, b \in R$.

(e) **Verallgemeinertes Distributivgesetz:** Für alle $a_1, \ldots, a_n, b_1, \ldots, b_m \in R$ gilt

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{m} b_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \cdot b_j \ .$$

Insbesondere gilt für alle $n \in \mathbb{N}_0$ und $a, b \in R$:

$$(n\,a)\,b = a\,(n\,b) \ .$$

Hierbei ist $n\,a := a + a + \ldots a$, genau n Mal.

Definition Ein Element $a \in R \setminus \{0\}$ eines Ringes R heißt ein **Nullteiler** in R, wenn ein Element $b \in R \setminus \{0\}$ existiert, so dass

$$a \cdot b = 0 \quad \text{oder} \quad b \cdot a = 0 \ .$$

R heißt **nullteilerfrei**, wenn es keine Nullteiler gibt, d.h. wenn Folgendes für alle $a, b \in R$ gilt:

$$a \cdot b = 0 \quad \Rightarrow \quad (a = 0) \text{ oder } (b = 0) \ .$$

1.4 Körper

Wir formalisieren nun, was es bedeutet, dass in einem Ring Elemente multiplikative Inverse besitzen. Neben Kommutativität ist dies, was Ringen zum Körper fehlt.

Definition Sei R ein Ring.

(a) Ein Element $a \in R$ heißt eine **Einheit** von R, wenn ein $b \in R$ existiert, so dass

$$ab = ba = 1 \ ,$$

wobei 1 das Einselement in R bezeichnet.

(b) Wir definieren die Menge der Einheiten von R als

$$R^* := \{ a \in R \mid a \text{ Einheit von } R \} \ .$$

(c) Ein kommutativer Ring $R \neq \{0\}$ heißt ein **Körper**, wenn $R^* = R \setminus \{0\}$.

Satz 1.11. *Sei K ein Körper.*

(a) *Dann ist $(K^* = K \setminus \{0\}, \cdot)$ eine Gruppe, genannt die **Einheitengruppe** von K. Insbesondere ist für jedes $a \in K^*$ das Inverse a^{-1} von a eindeutig bestimmt.*

(b) *Jeder Körper ist nullteilerfrei, d.h. für alle $a, b \in K$ gilt:*

$$a \cdot b = 0 \quad \Rightarrow \quad a = 0 \quad \textit{oder} \quad b = 0 \ .$$

Korollar 1.12. *Eine nichtleere Menge K zusammen mit zwei binären Verknüpfungen $+$ und $\cdot$ ist genau dann ein Körper, wenn die folgenden Bedingungen erfüllt sind:*

(A) *$(K, +)$ ist eine abelsche Gruppe.*

(M) *$(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.*

(D) *Es gilt das **Distributivgesetz**, d.h. für alle $a, b, c \in K$ gilt:*

$$a \cdot (b + c) = a \cdot b + a \cdot c \ .$$

Wegen der Kommutativität eines Körpers gilt auch das zweite Distributivgesetz.

2 Lineare Gleichungssysteme und Matrizen

Die Koeffizienten eines linearen Gleichungssystems und deren Anordnung bestimmen das System. Man kann die Koeffizienten als eine Struktur – eine Matrix – auffassen. Wir formalisieren zuerst den Begriff einer Matrix und erklären die Rechenregeln von Matrizen im Detail. Matrizen spielen in vielen Gebieten der Mathematik und in den Naturwissenschaften eine grundlegende Rolle.

In diesem Kapitel sei immer K ein Körper mit den Operationen $+$ und $\cdot$. Insbesondere besitzen alle von 0 verschiedene Elemente ein Inverses. In vielen Körpern schreibt man statt $a \cdot b^{-1}$ auch a/b. Außerdem seien $m, n \in \mathbb{N}$.

Ein **lineares Gleichungssystem von** m **Gleichungen in** n **Variablen** $x_1, \ldots, x_n$ **über dem Körper** K oder kurz ein $m \times n$ **lineares Gleichungssystem** über K ist gegeben durch

$$\left\{ \begin{array}{cccccccccccc} a_{11} & x_1 & + & a_{12} & x_2 & + & \ldots & + & a_{1n} & x_n & = & b_1 \\ a_{21} & x_1 & + & a_{22} & x_2 & + & \ldots & + & a_{2n} & x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & x_1 & + & a_{m2} & x_2 & + & \ldots & + & a_{mn} & x_n & = & b_m \end{array} \right\}, \tag{2.1}$$

wobei $a_{ij}, b_i \in K$ für alle $i = 1, \ldots, m$ und $j = 1, \ldots, n$. Dies ist äquivalent zu

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \quad (1 \leq i \leq m) . \tag{2.2}$$

Die a_{ij} heißen die **Koeffizienten** und die b_i die **rechten Seiten** von (2.1). Ist mindestens ein $b_i \neq 0$, so heißt (2.1) **inhomogen**, ansonsten **homogen** ($b_1 = \ldots = b_m = 0$).

Eine elegantere Schreibweise von (2.1) ist in Matrixform:

$$Ax = b \qquad (A \in K^{m \times n}, x \in K^{n \times 1}, b \in K^{m \times 1}) . \tag{2.3}$$

Dazu benötigen wir zuerst Matrizen und befassen uns dann mit den folgenden Fragen:

(a) Existiert eine Lösung von (2.3)? Was ist eine Lösung?

(b) Ist die Lösung von (2.3) eindeutig? Wenn nicht, wieviele Lösungen gibt es?

(c) Wie bestimmt man die Lösung(en) von (2.3)?

2.1 Matrizenrechnung

Definition Sei K ein Körper und seien $m\,,\, n \in \mathbb{N}$. Eine $m \times n$ **Matrix** A über K ist ein rechteckiges Schema mit m Zeilen und n Spalten der Form

$$\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1,i-1} & a_{1i} & a_{1,i+1} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2,i-1} & a_{2i} & a_{2,i+1} & \cdots & a_{2,j-1} & a_{2j} & a_{2,j+1} & \cdots & a_{2n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,i-1} & a_{i-1,i} & a_{i-1,i+1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
a_{i1} & a_{i2} & \cdots & a_{i,i-1} & a_{ii} & a_{i,i+1} & \cdots & a_{i,j-1} & a_{ij} & a_{i,j+1} & \cdots & a_{in} \\
a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,i-1} & a_{i+1,i} & a_{i+1,i+1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,i-1} & a_{j-1,i} & a_{j-1,i+1} & \cdots & a_{j-1,j-1} & a_{j-1,j} & a_{j-1,j+1} & \cdots & a_{j-1,n} \\
a_{j1} & a_{j2} & \cdots & a_{j,i-1} & a_{ji} & a_{j,i+1} & \cdots & a_{j,j-1} & a_{jj} & a_{j,j+1} & \cdots & a_{jn} \\
a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,i-1} & a_{j+1,i} & a_{j+1,i+1} & \cdots & a_{j+1,j-1} & a_{j+1,j} & a_{j+1,j+1} & \cdots & a_{j+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
a_{m-1,1} & a_{m-1,2} & \cdots & a_{m-1,i-1} & a_{m-1,i} & a_{m-1,i+1} & \cdots & a_{m-1,j-1} & a_{m-1,j} & a_{m-1,j+1} & \cdots & a_{m-1,n} \\
a_{m1} & a_{m2} & \cdots & a_{m,i-1} & a_{mi} & a_{m,i+1} & \cdots & a_{m,j-1} & a_{mj} & a_{m,j+1} & \cdots & a_{mn}
\end{pmatrix}$$

wobei $a_{ij} \in K$ für alle $i = 1, \ldots, m$ und $j = 1, \ldots, n$. Wir schreiben für diesen so definierten Ausdruck

$$A = (a_{ij})_{i,j} = (a_{ij}) = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \qquad \text{wobei alle } \; a_{ij} \in K \;.$$

(a) Wir nennen i den **Zeilenindex von** A und m die Anzahl der Zeilen von A.

(b) Wir nennen j den **Spaltenindex von** A und n die Anzahl der Spalten von A.

(c) Die **Größe von** A definieren wir als $m \times n$.

(d) Der Eintrag a_{ij} in der i-ten Zeile und j-ten Spalte wird mit **ij-ter Koeffizient von** A bezeichnet.

(e) Die Menge aller $m \times n$ Matrizen über K bezeichnen wir mit

$$K^{m\times n} = \left\{ A \,\middle|\, A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}, \text{ wobei } a_{ij} \in K \text{ für } i = 1, \ldots, m,\, j = 1, \ldots, n \right\} .$$

Oft schreibt man auch $\mathrm{Mat}(m, n, K)$ für $K^{m\times n}$.

(f) Die Matrix

$$\begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix} \in K^{1\times n}$$

nennt man **i-te Zeile** oder **i-ten Zeilenvektor von** A.

(g) Die Matrix

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^{m\times 1}$$

nennt man **j-te Spalte** oder **j-ten Spaltenvektor von** A.

(h) Die $m\times n$ **Nullmatrix** ist die Matrix $0_{m,n} = 0_{mn} = (0)$, wobei alle Einträge 0 sind. Wenn die Größe klar ist, schreiben wir einfach 0 für 0_{mn}.

(i) Man nennt $A \in K^{m\times n}$ eine **quadratische Matrix**, wenn $m = n$ gilt.

(j) Zwei Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ sind **gleich**, wenn sie dieselbe Größe $m\times n$ haben und $a_{ij} = b_{ij}$ für alle $i = 1, \dots, m$ und $j = 1, \dots, n$ gilt.

Wir definieren nun Zeilen- und Spaltenvektoren (auch ohne eine gegebene Matrix A).

- Ein **Zeilenvektor** mit n Komponenten über K ist eine $1\times n$ Matrix

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \in K^{1\times n} \quad \text{oder} \quad \begin{pmatrix} a_1, & a_2, & \dots, & a_n \end{pmatrix} \in K^{1\times n} .$$

- Ein Spaltenvektor mit m Komponenten über K ist eine $m\times 1$ Matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \in K^{m\times 1} .$$

- Wir identifizieren K mit $K^{1\times 1}$.

Wir definieren nun 3 Operationen auf Matrizen und untersuchen, welche Eigenschaften diese haben.

Matrizenaddition

Seien $A = (a_{ij})_{i,j} \in K^{m\times n}$ und $B = (b_{ij})_{i,j} \in K^{m\times n}$ gegeben. Dann definieren wir die **Summe von** A **und** B als die Matrix, die durch komponentenweise Addition entsteht:

$$A + B := (a_{ij} + b_{ij})_{i,j} \in K^{m\times n} .$$

Anders ausgedrückt haben wir

$$A + B := C := (c_{ij})_{i,j} \in K^{m\times n} ,$$

wobei die Koeffizienten c_{ij} von C gegeben sind durch

$$c_{ij} = a_{ij} + b_{ij} \quad \text{für alle } i = 1, \ldots, m \text{ und } j = 1, \ldots, n .$$

Die Summe von Matrizen unterschiedlicher Größe ist undefiniert.

Satz 2.1. (Rechenregeln der Matrizenaddition) *Seien $m, n \in \mathbb{N}$ und sei K ein Körper. Dann ist $(K^{m\times n}, +)$ eine abelsche Gruppe. Insbesondere gelten die folgenden Eigenschaften für alle A, B, $C \in K^{m\times n}$:*

(a) $A + (B + C) = (A + B) + C$.

(b) $A + 0_{mn} = 0_{mn} + A = A$.

(c) *Es existiert ein additives Inverses $-A \in K^{m\times n}$ von A, so dass gilt:*

$$A + (-A) = 0_{mn} = (-A) + A .$$

Ist $A = (a_{ij})_{i,j}$, so ist $-A := (-a_{ij}) \in K^{m\times n}$.

(d) $A + B = B + A$.

Skalarmultiplikation

Sei $A = (a_{ij}) = (a_{ij})_{i,j} \in K^{m\times n}$ und sei $\lambda \in K$ gegeben. Dann definieren wir die **Skalarmultiplikation des Skalars λ mit der Matrix** A wie folgt komponentenweise:

$$\lambda \cdot A := \lambda \cdot (a_{ij})_{i,j} := (\lambda a_{ij})_{i,j} \in K^{m\times n} .$$

Anders ausgedrückt haben wir

$$\lambda \cdot A := C := (c_{ij})_{i,j} \in K^{m\times n} ,$$

wobei die Koeffizienten c_{ij} von C gegeben sind durch

$$c_{ij} = \lambda a_{ij} \quad \text{für alle } i = 1, \ldots, m \text{ und } j = 1, \ldots, n .$$

Bemerkung 2.2. *Wir werden später sehen, dass $K^{m\times n}$ zusammen mir der Matrizenaddition und der Skalarmultiplikation zu einem K-Vektorraum wird. Hierbei ist Vorsicht geboten, da die oben definierte Multiplikation nur für (Skalar)$\times$(Matrix) gilt.*

(a) *Man kann Matrizenaddition und Skalarmultiplikation auch zusammenfassen, indem man für $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j} \in K^{m\times n}$ und $\lambda, \mu \in K$ rechnet:*

$$\lambda \cdot A + \mu \cdot B = (\lambda a_{ij} + \mu b_{ij})_{i,j} .$$

Dies ist die Kurzschreibweise für das Folgende: Es ist

$$\lambda \cdot A + \mu \cdot B := C := (c_{ij})_{i,j} \in K^{m\times n} ,$$

wobei die Koeffizienten c_{ij} von C gegeben sind durch

$$c_{ij} = \lambda a_{ij} + \mu b_{ij} \quad \text{für alle } i = 1, \ldots, m \text{ und } j = 1, \ldots, n .$$

(b) *Für das additive Inverse zu $A \in K^{m\times n}$ gilt: $-A = (-1) \cdot A$.*

(c) *Für die Subtraktion von $A = (a_{ij})_{i,j}$ und $B = (b_{ij})_{i,j} \in K^{m\times n}$ erhalten wir:*

$$A - B := A + (-B) = (a_{ij} - b_{ij})_{i,j} \in K^{m\times n} .$$

Satz 2.3. (Rechenregeln der Skalarmultiplikation) *Seien $m, n \in \mathbb{N}$ und sei K ein Körper. Für alle A, $B \in K^{m\times n}$ und $\lambda, \mu \in K$ gelten die folgenden Eigenschaften:*

(a) $1 \cdot A = A$.

(b) $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$.

(c) $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$.

(d) $(\lambda \mu) \cdot A = \lambda \cdot (\mu \cdot A)$.

Matrizenmultiplikation

Wir führen hier das Produkt zweier Matrizen ein. Im Gegensatz zur Matrizenaddition und Skalarmultiplikation wird die Matrizenmultiplikation nicht komponentenweise definiert.

Wir gehen nun in zwei Schritten vor. Zuerst definieren wir das Produkt eines Zeilenvektors mit einem Spaltenvektor gleicher Länge wie folgt:

$$\underbrace{\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}}_{\in K^{1\times n}} \cdot \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{\in K^{n\times 1}} := \underbrace{\sum_{i=1}^{n} a_i b_i}_{\in K = K^{1\times 1}} = a_1 b_1 + a_2 b_2 + \dots a_n b_n \ .$$

Allgemein müssen wir eine derartige Rechnung für jede Komponente des Produktes zweier Matrizen durchführen. Seien $A = (a_{ij}) \in K^{m\times n}$ und $B = (b_{jk}) \in K^{n\times p}$ gegeben. Dann definieren wir das **Produkt von** A **und** B als

$$A \cdot B := (c_{ik}) \in K^{m\times p} \ ,$$

wobei wir für alle $i = 1, \dots, m$ und $k = 1, \dots, p$ setzen

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix} \ .$$

Zur Illustration betrachten wir die folgende Situation:

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}}_{(m\times n)} \cdot \underbrace{\begin{pmatrix} b_{11} & \dots & b_{1k} & \dots & b_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & \dots & b_{nk} & \dots & b_{np} \end{pmatrix}}_{(n\times p)} = \underbrace{\begin{pmatrix} c_{11} & \dots & c_{1k} & \dots & c_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{i1} & \dots & c_{ik} & \dots & c_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{m1} & \dots & c_{mk} & \dots & c_{mp} \end{pmatrix}}_{(m\times p)}$$

Bemerkung 2.4. *Wir beachten folgende Eigenschaften.*

(a) *Das Produkt $A \cdot B$ zweier Matrizen A und B ist nur definiert, wenn die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist.*

(b) *Für $A \in K^{m \times n}$ Und $B \in K^{n \times p}$ ist zwar immer $A \cdot B$ definiert, aber $B \cdot A$ nur dann, wenn $m = p$ gilt. Ansonsten ist $B \cdot A$ undefiniert. Ist $m = p$, so sind $B \cdot A \in K^{n \times n}$ und $A \cdot B \in K^{m \times m}$.*

(c) *Für quadratische Matrizen $A, B \in K^{n \times n}$ sind $A \cdot B$ und $B \cdot A$ immer definiert.*

(d) *Im Allgemeinen ist die Matrizenmultiplikation nicht kommutativ und $K^{n \times n}$ besitzt Nullteiler bzgl. der oben definierten Multiplikation. Wir werden sehen, dass $(K^{n \times n}, +, \cdot)$ ein Ring ist.*

Satz 2.5. (Rechenregeln der Matrizenmultiplikation) *Sei K ein Körper. Für alle Matrizen A, B, C, D, für die die folgenden Summen und Produkte definiert sind, und für alle $\lambda \in K$ gilt:*

(a) $A(BC) = (AB)C$.

(b) $A(B+C) = AB + AC$.

(c) $(A+B)C = AC + BC$.

(d) $\lambda \cdot (AB) = (\lambda \cdot A)B = A(\lambda \cdot B)$.

(e) $(A+B)(C+D) = AC + AD + BC + BD$. *Insbesondere haben wir*

$$(A+B)^2 = A^2 + AB + BA + B^2 \quad .$$

Wir definieren nun das **Kroneckersymbol** δ_{ij} für $i\,,\, j \in \mathbb{N}$ als

$$\delta_{ij} \;:=\; \begin{cases} 1 & \text{, wenn } i = j \\ 0 & \text{, wenn } i \neq j \end{cases} \;.$$

Definition Für jedes $n \in \mathbb{N}$ ist die n-te **Einheitsmatrix** E_n in $K^{n \times n}$ die Matrix

$$E_n \;:=\; (\,\delta_{ij}\,) \;=\; \begin{array}{cc} & \begin{matrix} & & & & i & & & & & j & & & \end{matrix} \\ \begin{matrix} \\ \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & 0 & 0 & \ldots 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 & 0 & \ldots 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 1 & \ldots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots 1 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots 0 & 1 \end{pmatrix} \end{array} \;\in\; K^{n \times n} \;.$$

Satz 2.6. *Sei $A \in K^{m \times n}$, wobei K Körper ist und $m\,,\, n \in \mathbb{N}$. Dann gilt:*

(a) $A\,E_n \;=\; A\,.$

(b) $E_m\,A \;=\; A\,.$

(c) *Falls $m \;=\; n$ ist, so ist $A\,E_n \;=\; E_n\,A \;=\; A\,.$*

Korollar 2.7. *Sei K ein Körper und sei $n \in \mathbb{N}$. Dann ist $(K^{n\times n}, +, \cdot)$ ein Ring, genannt der **Matrizenring der** $n \times n$ **Matrizen über** K.*

Es ist leicht zu sehen, dass multiplikative Inverse von quadratischen Matrizen nicht immer existieren müssen. Wir führen den Begriff einer Inversen explizit für Matrizen ein.

Definition Wir sagen, dass $A \in K^{n\times n}$ **invertierbar** oder **regulär** ist, wenn eine Matrix $B \in K^{n\times n}$ existiert, so dass

$$AB = E_n = BA \ .$$

Die Matrix B wird dann ein **Inverses von** A genannt. Wir definieren

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

Diese Menge heißt **allgemeine lineare Gruppe**. Der Term Gruppe wird im nächsten Satz gerechtfertigt.

Satz 2.8. *Sei K ein Körper und sei $n \in \mathbb{N}$. Dann ist $(\mathrm{GL}(n, K), \cdot)$ eine Gruppe mit neutralem Element E_n. Außerdem gilt:*

(a) *Ist A invertierbar, so ist das Inverse von A eindeutig bestimmt, und wir bezeichnen es mit A^{-1}.*

(b) $(A^{-1})^{-1} = A$ *für alle* $A \in \mathrm{GL}(n, K)$.

(c) $(AB)^{-1} = B^{-1}A^{-1}$ *für alle* $A, B \in \mathrm{GL}(n, K)$.

(d) $(\lambda A)^{-1} = \lambda^{-1}A^{-1}$ *für alle* $A \in \mathrm{GL}(n, K)$ *und für alle* $\lambda \in K^*$.

Wir werden später besprechen, wie man Inverse berechnet, sofern diese existieren.

Transposition

Zum Schluss dieses Abschnitts führen wir noch die Transponierte einer Matrix A ein. Diese entsteht aus A, indem man Zeilen und Spalten vertauscht. Sie wird unter anderem benötigt, wenn man über Spaltenoperationen und über den Spaltenraum spricht.

Definition Sei $A = (a_{ij}) \in K^{m\times n}$. Dann definieren wir die **Transponierte von** A als die $n \times m$ Matrix

$$A^T := (a_{ji}) = (a_{ji})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \in K^{n\times m} .$$

Wir definieren also die Transponierte einer Matrix $A = (a_{ij}) \in K^{m\times n}$ als die Matrix

$$A^T := (c_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} ,$$

wobei die Koeffizienten c_{ij} von A^T gegeben sind durch

$$c_{ij} = a_{ji} \qquad \text{für alle } i = 1, \dots, n \text{ und } j = 1, \dots, m .$$

Wir illustrieren dies wie folgt:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in K^{m\times n} \quad \Rightarrow \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} \in K^{n\times m}$$

Satz 2.9. (Rechenregeln der Transponierten) *Seien $m, n \in \mathbb{N}$ und sei K ein Körper. Für alle Matrizen A, B, für die die folgenden Summen und Produkte definiert sind, und für alle $\lambda \in K$ gilt:*

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

(b) $(A+B)^T = A^T + B^T = (B+A)^T$.

(c) $(AB)^T = B^T A^T$.

(d) $(\lambda A)^T = \lambda A^T$.

(e) *Falls $m = n$ und A invertierbar ist, so ist auch A^T invertierbar, und wir haben*

$$(A^T)^{-1} = (A^{-1})^T .$$

Matrizenmultiplikation ausgedrückt mit Vektoren

Seien $A = (a_{ij}) \in K^{m \times n}$ und $B = (b_{jk}) \in K^{n \times p}$ gegeben. Dann können wir auch A mittels Zeilenvektoren und B mittels Spaltenvektoren schreiben:

$$A = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} \quad , \quad B = \begin{pmatrix} s_1 & s_2 & \dots & s_p \end{pmatrix} \ ,$$

wobei

$$z_i := \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \quad , \quad s_k := \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix}$$

Folglich ist

$$A\,B := (c_{ik}) := (z_i \cdot s_k) \in K^{m \times p} \quad (1 \le i \le m, 1 \le k \le p) \ .$$

In der Schreibweise der Transponierten erhalten wir sofort, dass

$$z_i \cdot s_k = \sum_{j=1}^{n} a_{ij}\, b_{jk} = \sum_{j=1}^{n} b_{jk}\, a_{ij} = s_k^T \cdot z_i^T \quad (1 \le i \le m, 1 \le k \le p) \ .$$

Damit lassen sich viele Rechenregeln einfach wie z.B. Teil (c) aus Satz 2.9 beweisen.

Satz 2.10. *Es seien $A \in K^{m \times n}$ und $m, n, p, r \in \mathbb{N}$.*

(a) *Falls A eine Nullzeile besitzt, so auch $A \cdot B$ für alle $B \in K^{n \times p}$.*

(b) *Falls A eine Nullspalte besitzt, so auch $C \cdot A$ für alle $C \in K^{r \times m}$.*

(c) *Falls A eine Nullzeile oder Nullspalte besitzt, so ist A nicht invertierbar.*

2.2 Elementare Zeilenumformungen

Sei K ein Körper und seien $m, n, p, r, q \in \mathbb{N}$. Wir schreiben wie gewohnt das lineare Gleichungssystem (2.1) in Matrixform

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{2.4}$$

oder kurz

$$A x = b \qquad (A \in K^{m \times n}, x \in K^{n \times 1}, b \in K^{m \times 1}) .$$

Man nennt dann

$$\text{Lös}(A, b) := \{ x \in K^{n \times 1} \mid A x = b \}$$

die **Lösungsmenge** von (2.4).

Definition Wir definieren die **erweiterte Koeffizientenmatrix** $(A \mid b) \in K^{m \times (n+1)}$ von (2.4) als

$$(A \mid b) := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right) \in K^{m \times (n+1)} .$$

Der vertikale Strich dient zur Orientierung und wird manchmal weggelassen. Die Matrix A nennt man die **Koeffizientenmatrix** von (2.4).

Ziel:

- Überführe die erweiterte Koeffizientenmatrix in eine einfachere Form, so dass die Lösungsmenge abgelesen werden kann und die Lösbarkeit bestimmt werden kann.
- Wir formalisieren zuerst, was "überführen" bedeutet und warum dies erlaubt ist.
- Außerdem untersuchen wir, welche "einfacheren Formen" es gibt.

Definition Sei $A \in K^{m\times n}$. Unter einer **elementaren Zeilenumformung** versteht man eine der folgenden Operationen:

(I) Multiplikation der k-ten Zeile von A mit einem Skalar $\lambda \in K^* = K \setminus \{0\}$, wobei $k \in \{1, \ldots, m\}$.
Schreibweise: $A \to A_I$ und die Operation als $Z_k \to \lambda Z_k$.

(II) Addition des μ-fachen der ℓ-ten Zeile von A zur k-ten Zeile, wobei $\mu \in K$ beliebig und $k, \ell \in \{1, \ldots, m\}$ mit $k \neq \ell$.
Schreibweise: $A \to A_{II}$ und die Operation als $Z_k \to Z_k + \mu Z_\ell$.

(III) Vertauschen der k-ten und der ℓ-ten Zeile von A, wobei $k, \ell \in \{1, \ldots, m\}$.
Schreibweise: $A \to A_{III}$ und die Operation als $Z_k \leftrightarrow Z_\ell$.

Man hat also folgende elementaren Zeilenumformungen:

Typ I: Multiplikation einer Zeile von A mit einem von 0 verschiedenen Skalar.

Typ II: Addition des Vielfachen einer Zeile zu einer anderen.

Typ III: Vertauschen zweier Zeilen.

In Typ I ist zu beachten, dass der Skalar unbedingt $\neq 0$ sein muss. Dann wird jeder Eintrag der k-ten Zeile mit λ multipliziert. Für Typ II erlauben wir ein beliebiges Vielfaches μ, insbesondere $\mu = 0$. Die andere Zeile muss tatsächlich eine verschiedene Zeile sein. In Typ III ist dieselbe Zeile erlaubt, was dann aber zu keiner Veränderung führt.

Ein wichtiges Resultat für die algorithmische Berechnung der Lösung von (2.4) ist der folgende Satz, den man direkt anhand der Gleichungen verifiziert oder aus einem späteren wichtigen Ergebnis im Zusammenhang mit Elementarmatrizen folgert (siehe Satz 2.20).

Satz 2.11. *Sei $A \in K^{m\times n}$ und sei $b \in K^{m\times 1}$. Ist $(\tilde{A} \mid \tilde{b})$ die erweiterte Koeffizientenmatrix, die aus der erweiterten Koeffizientenmatrix $(A \mid b)$ von $Ax = b$ durch eine elementare Zeilenumformung hervorgegangen ist, so gilt:*

$$\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b}) .$$

Vollkommen analog definieren wir folgende elementare Spaltenumformungen:

Typ I: Multiplikation einer Spalte von A mit einem von 0 verschiedenen Skalar.

Typ II: Addition des Vielfachen einer Spalte zu einer anderen.

Typ III: Vertauschen zweier Spalten.

Definition Sei $A \in K^{m\times n}$. Unter einer **elementaren Spaltenumformung** versteht man eine der folgenden Operationen:

(I) Multiplikation der k-ten Spalte von A mit einem Skalar $\lambda \in K^* = K\setminus\{0\}$, wobei $k \in \{1, \ldots, n\}$.
Schreibweise: $A \to A^I$ und die Operation als $S_k \to \lambda \cdot S_k$.

(II) Addition des μ-fachen der ℓ-ten Spalte von A zur k-ten Spalte, wobei $\mu \in K$ beliebig und $k, \ell \in \{1, \ldots, m\}$ mit $k \neq \ell$.
Schreibweise: $A \to A^{II}$ und die Operation als $S_k \to S_k + \mu S_\ell$.

(III) Vertauschen der k-ten und der ℓ-ten Spalte von A, wobei $k, \ell \in \{1, \ldots, n\}$.
Schreibweise: $A \to A^{III}$ und die Operation als $S_k \leftrightarrow S_\ell$.

Bemerkung 2.12. *Es ist **Vorsicht** mit Spaltenoperationen geboten.*

(a) ***Nicht-Invarianz des Lösungsraumes bei elementaren Spaltenumformungen:*** *Sei $A \in K^{m\times n}$ und sei $b \in K^{m\times 1}$. Ist $(\tilde{A} \mid \tilde{b})$ die erweiterte Koeffizientenmatrix, die aus der erweiterten Koeffizientenmatrix $(A \mid b)$ von $Ax = b$ durch eine elementare Spaltenumformung hervorgegangen ist, so gilt im Allgemeinen $\text{Lös}(A, b) \neq \text{Lös}(\tilde{A}, \tilde{b})$. Die Lösungen stimmen zwar "im Prinzip" überein, aber nur, wenn man die Variablen richtig umnummeriert. Deswegen werden wir uns bei den Algorithmen auf Zeilenumformungen beschränken.*

(b) *Die elementaren Spaltenumformungen entsprechen elementaren Zeilenumformungen angewandt auf A^T. Zum Beispiel:*

$$A \to A^I \quad \Leftrightarrow \quad \left[A \to A^T \to (A^T)_I \to ((A^T)_I)^T = A^I\right] \ .$$

2.3 Elementarmatrizen

Sei K ein Körper und seien $m, n \in \mathbb{N}$. Elementare Zeilenumformungen sind ultimativ mit den drei Typen von Elementarmatrizen verbunden.

Definition Eine **Elementarmatrix** ist eine Matrix, die aus einer elementaren Umformung angewandt auf die Einheitsmatrix entsteht.

(a) Wendet man eine elementare Zeilenumformung speziell auf E_m an, so erhält man die **zu dieser elementaren Zeilenumformung gehörige Elementarmatrix** $Q \in K^{m\times m}$.

(b) Wendet man eine elementare Spaltenumformung auf E_n an, so erhält man die **zu dieser elementaren Spaltenumformung gehörige Elementarmatrix** $P \in K^{n\times n}$.

Für beliebige $k, \ell \in \mathbb{N}$ definieren wir nun

$$E_{k\ell} := \begin{array}{cc} & \begin{array}{ccccccccccccc} & & & & k & & & & \ell & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ k \\ \\ \\ \\ \ell \\ \\ \\ \\ \\ \end{array} & \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \end{array} \in K^{m\times m} ,$$

mit deren Hilfe wir die zu elementaren Zeilenumformungen gehörigen Elementarmatrizen einfach beschreiben können.

Typ I: Multiplikation der k-ten Zeile von E_m mit einem Skalar $\lambda \in K^*$.

$$E_m \xrightarrow{Z_k \to \lambda Z_k} S_k(\lambda) := \begin{array}{cc} & \begin{array}{ccccccccccccc} & & & & k & & & & \ell & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ k \\ \\ \\ \\ \ell \\ \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \lambda & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \end{array} \in K^{m \times m} .$$

Typ II: Addition des μ-fachen der ℓ-ten Zeile von E_m zur k-ten Zeile, wenn $k \neq \ell$.

$$E_m \xrightarrow{Z_k \to Z_k + \mu Z_\ell} Q_{k\ell}(\mu) := \begin{array}{c@{}c} & \begin{array}{ccccccccccccc} & & & & k & & & & & \ell & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ k \\ \\ \\ \\ \ell \\ \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \mu & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \end{array} \in K^{m \times m} .$$

Typ III: Vertauschen der k-ten und der ℓ-ten Zeile von E_m, wenn $k \neq \ell$.

$$E_m \xrightarrow{Z_k \leftrightarrow Z_\ell} P_{k\ell} := \begin{array}{c@{}c} & \begin{array}{ccccccccccccc} & & & & k & & & & \ell & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ k \\ \\ \\ \\ \ell \\ \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \end{array} \in K^{m \times m} .$$

Vollkommen analog kann man die elementaren Spaltenumformungen anwenden bzw. elementare Zeilenumformungen auf A^T anwenden.

Typ I: Multiplikation der k-ten Spalte von E_n mit einem Skalar $\lambda \in K^*$.

$$E_n \overset{S_k \to \lambda S_k}{\longrightarrow} S_k(\lambda) \in K^{n \times n} .$$

Typ II: Addition des μ-fachen der ℓ-ten Spalte von E_n zur k-ten Spalte, wenn $k \neq \ell$..

$$E_n \overset{S_k \to S_k + \mu S_\ell}{\longrightarrow} Q_{\ell k}(\mu) \in K^{n \times n} .$$

Typ III: Vertauschen der k-ten und der ℓ-ten Spalte von E_n, wenn $k \neq \ell$..

$$E_n \overset{S_k \leftrightarrow S_\ell}{\longrightarrow} P_{k\ell} \in K^{n \times n} .$$

Proposition 2.13. *Für alle $k, \ell \in \mathbb{N}$ mit $k \neq \ell$ sowie $\lambda \in K^*$ und $\mu \in K$ gilt:*

(a) $S_k(\lambda) = E_m + (\lambda - 1) E_{kk} = (\sigma_{ri}) \in K^{m \times m}$, *wobei*

$$\sigma_{ri} := \left\{ \begin{array}{ll} 1 & \text{, falls } r = i \text{ und } r \neq k \\ \lambda & \text{, falls } r = i = k \\ 0 & \text{, sonst} \end{array} \right\} \qquad \text{für alle } r, i \in \{1, \ldots, m\} .$$

(b) $Q_{k\ell}(\mu) = E_m + \mu E_{k\ell} = (\tau_{ri}) \in K^{m \times m}$, *wobei*

$$\tau_{ri} := \left\{ \begin{array}{ll} 1 & \text{, falls } r = i \\ \mu & \text{, falls } r = k \text{ und } i = \ell \\ 0 & \text{, sonst} \end{array} \right\} \qquad \text{für alle } r, i \in \{1, \ldots, m\} .$$

(c) $P_{k\ell} = E_m - E_{kk} + E_{k\ell} - E_{\ell\ell} + E_{\ell k} = (\pi_{ri}) \in K^{m \times m}$, *wobei*

$$\pi_{ri} := \left\{ \begin{array}{ll} 1 & \text{, falls } r = i \neq k, \ell \text{ oder } r = k, i = \ell \text{ oder } r = \ell, i = k \\ 0 & \text{, sonst} \end{array} \right\}$$

für alle $r, i \in \{1, \ldots, m\}$.

Proposition 2.14. (**Transposition von Elementarmatrizen**)
Für alle $k, \ell \in \mathbb{N}$ mit $k \neq \ell$ sowie $\lambda \in K^*$ und $\mu \in K$ gilt:

(a) $S_k(\lambda)^T = S_k(\lambda)$.

(b) $Q_{k\ell}(\mu)^T = Q_{\ell k}(\mu)$.

(c) $P_{k\ell}^T = P_{k\ell}$.

Satz 2.15. *Sei $A \in K^{m\times n}$. Ist $Q \in K^{m\times m}$ die zu einer elementaren Zeilenumformung gehörige Elementarmatrix, dann ist $Q\,A \in K^{m\times n}$ die Matrix, welche aus A bei dieser Umformung entsteht. Genauer gelten die folgenden Aussagen.*

(a) *Sei $\lambda \in K^* = K \setminus \{0\}$ und $k \in \{1, \dots, m\}$.*

$$A \overset{Z_k \to \lambda\cdot Z_k}{\longrightarrow} A_I \qquad \Leftrightarrow \qquad A_I = S_k(\lambda)\,A\ .$$

(b) *Sei $\mu \in K$ und seien $k, \ell \in \{1, \dots, m\}$ mit $k \neq \ell$.*

$$A \overset{Z_k \to Z_k + \mu Z_\ell}{\longrightarrow} A_{II} \qquad \Leftrightarrow \qquad A_{II} = Q_{k\ell}(\mu)\,A\ .$$

(c) *Seien $k, \ell \in \{1, \dots, m\}$ mit $k \neq \ell$.*

$$A \overset{Z_k \leftrightarrow Z_\ell}{\longrightarrow} A_{III} \qquad \Leftrightarrow \qquad A_{III} = P_{k\ell}\,A\ .$$

Satz 2.16. *Sei $A \in K^{m\times n}$. Ist $P \in K^{n\times n}$ die zu einer elementaren Spaltenumformung gehörige Elementarmatrix, dann ist $A\,P \in K^{m\times n}$ die Matrix, welche aus A bei dieser Umformung entsteht. Genauer gelten die folgenden Aussagen.*

(a) *Sei $\lambda \in K^* = K \setminus \{0\}$ und $k \in \{1, \dots, n\}$.*

$$A \overset{S_k \to \lambda S_k}{\longrightarrow} A^I \qquad \Leftrightarrow \qquad A^I = A\,S_k(\lambda)\ .$$

(b) *Sei $\mu \in K$ und seien $k, \ell \in \{1, \dots, n\}$ mit $k \neq \ell$.*

$$A \overset{S_k \to S_k + \mu S_\ell}{\longrightarrow} A^{II} \qquad \Leftrightarrow \qquad A^{II} = A\,Q_{\ell k}(\mu)\ .$$

(c) *Seien $k, \ell \in \{1, \dots, n\}$ mit $k \neq \ell$.*

$$A \overset{S_k \leftrightarrow S_\ell}{\longrightarrow} A^{III} \qquad \Leftrightarrow \qquad A^{III} = A\,P_{k\ell}\ .$$

Satz 2.17. *Jede Elementarmatrix ist invertierbar und ihre Inversen sind wieder Elementarmatrizen. Genauer gilt für alle $k, \ell \in \mathbb{N}$ mit $k \neq \ell$ sowie $\lambda \in K^*$ und $\mu \in K$:*

(a) $S_k(\lambda)^{-1} = S_k(\lambda^{-1})$.

(b) $Q_{k\ell}(\mu)^{-1} = Q_{k\ell}(-\mu)$.

(c) $P_{k\ell}^{-1} = P_{k\ell}$.

Korollar 2.18. *Informell gelten die folgenden Aussagen:*

(a) *Elementare Zeilenumformungen entsprechen der Multiplikation mit den zugehörigen Elementarmatrizen von links.*

(b) *Elementare Spaltenumformungen entsprechen der Multiplikation mit den zugehörigen Elementarmatrizen von rechts.*

Wir können damit ein zentrales Ergebnis dieses Kapitels formulieren, welches eine Zusammenfassung der bisherigen Resultate darstellt.

Satz 2.19. *Sei $A \in K^{m\times n}$ und sei $B \in K^{m\times n}$ eine Matrix, welche aus A durch q elementare Zeilenumformungen mit zugehörigen Elementarmatrizen $Q_1, \ldots, Q_q$ und p elementaren Spaltenumformungen mit zugehörigen Elementarmatrizen $P_1, \ldots, P_p$ entsteht. Dann existieren Matrizen $Q \in \mathrm{GL}(m, K)$ und $P \in \mathrm{GL}(n, K)$, so dass*

$$Q\,A\,P = B\ .$$

Satz 2.20. *Es seien $A \in K^{m\times n}$, $b \in K^{m\times 1}$ und sei $Q \in \mathrm{GL}(m, K)$. Dann sind die linearen Gleichungssysteme $A\,x = b$ und $(Q\,A)\,x = Q\,b$ äquivalent. Insbesondere bedeutet dies, dass*

$$\text{Lös}\,(A, b) = \text{Lös}\,(Q\,A, Q\,b)\ .$$

Der folgende Satz ist die Wiederholung der Behauptung aus Satz 2.11 und wird mit wesentlich eleganteren Methoden bewiesen.

Satz 2.21. *Es seien $A \in K^{m\times n}$ und $b \in K^{m\times 1}$. Sei $(\tilde{A} \mid \tilde{b})$ die erweiterte Koeffizientenmatrix, die aus der erweiterten Koeffizientenmatrix $(A \mid b)$ des linearen Gleichungssytems $A\,x = b$ durch eine endliche Folge von elementaren Zeilenumformungen hervorgegangen ist. Dann sind die linearen Gleichungssysteme $A\,x = b$ und $\tilde{A}\,x = \tilde{b}$ äquivalent. Zudem gilt:*

$$\text{Lös}\,(A, b) = \text{Lös}\,(\tilde{A}, \tilde{b})\ .$$

2.4 Gaußscher Algorithmus und Gauß-Jordan-Verfahren

Wir haben das theoretische Fundament für diesen Abschnitt bereits gelegt und die wichtigsten Aussagen bewiesen. Nun betrachten wir die wesentlichen Algorithmen und stellen zwei explizite Verfahren vor.

Sei K ein Körper und seien $m, n \in \mathbb{N}$. Sei $A \in K^{m\times n}$ und sei $b \in K^{m\times 1}$. Die Idee ist wie folgt: Wende elementare Zeilenumformungen auf die erweiterte Koeffizientenmatrix $(A \mid b)$ des linearen Gleichungssystems $Ax = b$ an, um ein einfacheres, äquivalentes Gleichungssystem zu erhalten. Das neue System soll ungefähr obere Dreiecksform haben. In der Literatur gibt es mehrere unterschiedliche Definitionen für dieselben Objekte. Wir betrachten die Zeilenstufenform, die Gaußsche Normalform und die reduzierte Gaußsche Normalform in unseren Definitionen. Die Zeilenstufenform taucht in der Algebra I auf und spielt eine wichtige Rolle in der Elementarteilertheorie. Das ist insbesondere relevant, wenn der zugrundeliegende Koeffizientenbereich ein Ring und kein Körper ist.

2.4.1 Zeilenstufenform

Definition Wir sagen, eine Matrix $A = (a_{ij}) \in K^{m\times n}$ sei in **Zeilenstufenform**, mit r Stufen, wobei $r \in \{0, \ldots, m\}$, und mit Stufenindizes $j_1, j_2, \ldots, j_r$, wenn gilt:

(i) $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ und $r \leq \min(n, m)$.

(ii) $a_{ij} = 0$ für alle $i > r$ und $1 \leq j \leq n$.

(iii) Für alle $i = 1, \ldots, r$ gilt: $a_{ij_i} \neq 0$ und $a_{ij} = 0$ für alle $1 \leq j < j_i$.

Man nennt die Stufenpositionen $(1, j_1), (2, j_2), \ldots, (r, j_r)$ **Pivotpositionen** und $a_{1j_1}, \ldots, a_{rj_r}$ die **Pivotelemente**.

Informell bedeutet dies, eine Matrix $A = (a_{ij}) \in K^{m\times n}$ ist in Zeilenstufenform, wenn Folgendes gilt:

- Nullzeilen (wenn vorhanden) tauchen am Ende der Matrix auf.
- Für zwei von der Nullzeile verschiedene, aufeinanderfolgende Zeilen ist der erste von 0 verschiedene Eintrag in der oberen Zeile weiter links positioniert als der erste von 0 verschiedene Eintrag in der unteren Zeile.

Offensichtlich kann man eine Matrix mittels elementaren Zeilenumformungen in eine Zeilenstufenform bringen. Das wird später beschrieben und in den Übungen vertieft.

2.4.2 Gaußsche Normalform

Definition Eine Matrix $A = (a_{ij}) \in K^{m\times n}$ ist in **Gaußscher Normalform** mit r Stufen, wobei $r \in \{0, \dots, m\}$, und mit Stufenindizes $j_1, j_2, \dots, j_r$, wenn sie in Zeilenstufenform ist und die Pivotelemente alle gleich 1 sind. D.h. es gilt:

(i) $1 \le j_1 < j_2 < \dots < j_r \le n \quad$ und $\quad r \le \min(n, m)$

(ii) $a_{ij} = 0$ für alle $i > r$ und $1 \le j \le n$.

(iii) Für alle $i = 1, \dots, r$ gilt: $a_{ij_i} = 1$ und $a_{ij} = 0$ für alle $1 \le j < j_i$.

Informell bedeutet dies, eine Matrix $A = (a_{ij}) \in K^{m\times n}$ ist in Gaußscher Normalform, wenn Folgendes gilt:

- Nullzeilen (wenn vorhanden) tauchen am Ende der Matrix auf.
- Für jede von der Nullzeile verschiedene Zeile ist der erste Eintrag $\neq 0$ von links betrachtet eine 1. Diesen Eintrag nennt man dann die **führende** 1.
- Für zwei von der Nullzeile verschiedene, aufeinanderfolgende Zeilen ist die führende 1 der oberen Zeile weiter links positioniert als die führende 1 in der unteren Zeile.

Satz 2.22. *Sei $A \in K^{m\times n}$.*

(a) *A lässt sich durch elementare Zeilenumformungen vom Typ II und III in Zeilenstufenform bringen.*

(b) *A lässt sich durch elementare Zeilenumformungen vom Typ I, II und III in Gaußsche Normalform bringen.*

Algorithmus 2.23. (**Gaußsche Normalform**)

Input: *$A \in K^{m\times n}$, wobei K eine Körper ist und $m, n \in \mathbb{N}$.*

Output: *Eine Matrix $B \in K^{m\times n}$ in Gaußscher Normalform mit r Stufen, welche aus A durch elementare Zeilenumformungen hervorgeht.*

(0) *Ist $A = (0)$, so ist $r = 0$ und STOP.*

(1) *$\ell := 1$. In jedem Schritt ersetzen wir die alte Matrix A durch eine neue Matrix, welche wir wiederum A nennen.*

(2) *Existieren Nullzeilen, so führe Vertauschungen von Zeilen (Typ III) durch, so dass alle Nullzeilen am Ende der Matrix sind.*

(3) *Hauptschritt: Finde rekursiv die Stufenindizes j_ℓ für $1 \leq \ell \leq r$ in Untermatrizen.*

(3.1) *Sei S_{j_ℓ} die erste Spalte von A, die ab der ℓ-ten Zeile nicht nur Komponenten gleich 0 hat. [oder wähle geschickt]*

(3.2) *Zu dieser Spalte S_{j_ℓ} in (3.1) gehört eine erste Zeile Z_k mit Einträgen*

$$(a_{k1}, \ldots, a_{kj_\ell}, \ldots, a_{kn}) \in K^{1\times n} \quad , \quad \textit{wobei } k \geq \ell \textit{ und } a_{kj_\ell} \neq 0 \quad .$$

(3.3) *Vertausche die k-te und die ℓ-te Zeile (Typ III). Danach ist $a_{\ell j_\ell} \neq 0$ und der ℓ-te Stufenindex j_ℓ ist gefunden.*

(3.4) *Multipliziere die ℓ-te Zeile mit $a_{\ell j_\ell}^{-1}$ (Typ I), welches existiert, da $a_{\ell j_\ell} \neq 0$. In der neuen Matrix A ist demnach $a_{\ell j_\ell} = 1$.*

(3.5) *Lösche alle Einträge in der ℓ-ten Spalte unterhalb von $a_{\ell j_\ell}$ mittels Typ II Zeilenumformungen. Für jedes $k > \ell$ ersetze Z_k durch $Z_k + (-a_{kj_\ell})\, Z_\ell$.*

(3.6) *Existieren Nullzeilen, so führe Vertauschungen von Zeilen (Typ III) durch, so dass alle Nullzeilen am Ende der Matrix sind.*

(3.7) *Gibt es in der entstandenen Matrix A noch eine Spalte, die ab der $(\ell+1)$-ten Zeile nicht nur Nullen als Komponenten hat, so ersetze ℓ durch $\ell + 1$ und gehe zurück zu (3.1). Andernfalls $r := \ell$ und STOP.*

Die Korrektheit des Algorithmus folgt unmittelbar nach Konstruktion und kann durch vollständige Induktion bewiesen werden. (a) folgt nun, indem man den Schritt (3.4) weglässt und in (3.5) Z_k durch $Z_k + \left(-a_{kj_\ell}\, a_{\ell j_\ell}^{-1}\right) Z_\ell$ ersetzt.

Satz 2.24. (Lösbarkeitskriterium) Sei $A \in K^{m\times n}$ eine Matrix in Zeilenstufenform bzw. in Gaußscher Normalform mit r Stufen und sei $b \in K^{m\times 1}$. Dann gilt Folgendes: Das lineare Gleichungssystem System $Ax = b$ ist genau dann lösbar, wenn

$$b_{r+1} = \ldots = b_m = 0 \; .$$

2.4.3 Reduzierte Gaußsche Normalform, reduzierte Zeilenstufenform

Definition Eine Matrix $A = (a_{ij}) \in K^{m\times n}$ ist in **reduzierter Gaußscher Normalform** mit r Stufen, wobei $r \in \{0, \ldots, m\}$, und mit Stufenindizes $j_1, j_2, \ldots, j_r$, wenn sie in Gaußscher Normalform ist und wenn für alle $i = 1, \ldots, r$ die j_i-te Spalte von A gleich der i-ten Spalte von E_m ist. D.h. es gilt:

(i) $1 \leq j_1 < j_2 < \ldots < j_r \leq n \quad \text{und} \quad r \leq \min(n, m)$

(ii) $a_{ij} = 0$ für alle $i > r$ und $1 \leq j \leq n$.

(iii) Für alle $i = 1, \ldots, r$ gilt: $a_{ij_i} = 1$ und $a_{ij} = 0$ für alle $j \neq j_i$.

Informell bedeutet dies, eine Matrix $A = (a_{ij}) \in K^{m\times n}$ ist in reduzierter Gaußscher Normalform, wenn gilt:

- Nullzeilen (wenn vorhanden) tauchen am Ende der Matrix auf.
- Für jede von der Nullzeile verschiedene Zeile ist der erste Eintrag $\neq 0$ von links betrachtet eine 1. Diesen Eintrag nennt man dann die **führende** 1.
- Für zwei von der Nullzeile verschiedene, aufeinanderfolgende Zeilen ist die führende 1 der oberen Zeile weiter links positioniert als die führende 1 in der unteren Zeile.
- Jede Spalte, welche die führende 1 einer Zeile enthält, hat in allen anderen Komponenten nur Nullen.

Satz 2.25. *Jede Matrix $A \in K^{m\times n}$ lässt sich durch elementare Zeilenumformungen vom Typ I, II und III in reduzierte Gaußsche Normalform bringen.*

Algorithmus 2.26. (**Reduzierte Gaußsche Normalform**)

Input: $A \in K^{m\times n}$, *K Körper.*

Output: *Eine Matrix* $B \in K^{m\times n}$ *in reduzierter Gaußscher Normalform mit r Stufen, welche aus A durch elementare Zeilenumformungen hervorgeht.*

(0) *Ist* $A = (0)$, *so ist* $r = 0$ *und STOP.*

(1) $\ell := 1$. *In jedem Schritt ersetzen wir die alte Matrix A durch eine neue Matrix, welche wir wiederum A nennen.*

(2) *Existieren Nullzeilen, so führe Vertauschungen von Zeilen (Typ III) durch, so dass alle Nullzeilen am Ende der Matrix sind.*

(3) *Hauptschritt: Finde rekursiv die Stufenindizes* j_ℓ *für* $1 \le \ell \le r$ *in Untermatrizen.*

(3.1) *Sei* S_{j_ℓ} *die erste Spalte von* A, *die ab der* ℓ*-ten Zeile nicht nur Komponenten gleich* 0 *hat. [oder wähle geschickt]*

(3.2) *Zu dieser Spalte* S_{j_ℓ} *in (3.1) gehört eine erste Zeile* Z_k *mit Einträgen*

$$(a_{k1}, \dots, a_{kj_\ell}, \dots, a_{kn}) \in K^{1\times n}, \quad \text{wobei } k \ge \ell \text{ und } a_{kj_\ell} \ne 0 \ .$$

(3.3) *Vertausche die* k*-te und die* ℓ*-te Zeile (Typ III). Danach ist* $a_{\ell j_\ell} \ne 0$ *und der* ℓ*-te Stufenindex* j_ℓ *ist gefunden.*

(3.4) *Multipliziere die* ℓ*-te Zeile mit* $a_{\ell j_\ell}^{-1}$ *(Typ I), welches existiert, da* $a_{\ell j_\ell} \ne 0$. *In der neuen Matrix A ist demnach* $a_{\ell j_\ell} = 1$.

(3.5') *Lösche alle Einträge in der* ℓ*-ten Spalte unterhalb und oberhalb von* $a_{\ell j_\ell}$ *mittels Typ II Zeilenumformungen. Für jedes* $k > \ell$ *ersetze* Z_k *durch* $Z_k + (-a_{kj_\ell})\, Z_\ell$.

(3.6) *Existieren Nullzeilen, so führe Vertauschungen von Zeilen (Typ III) durch, so dass alle Nullzeilen am Ende der Matrix sind.*

(3.7) *Gibt es in der entstandenen Matrix A noch eine Spalte, die ab der* $(\ell+1)$*-ten Zeile nicht nur Nullen als Komponenten hat, so ersetze* ℓ *durch* $\ell+1$ *und gehe zurück zu (3.1). Andernfalls* $r := \ell$ *und STOP.*

Die Korrektheit des Algorithmus folgt nach Konstruktion und durch vollständige Induktion. Alternativ: Wende Algorithmus 2.23 und am Ende führt man aus: Für alle $k = 1, \dots, r$ und $\ell = k+1, \dots, m$ ersetze Z_k durch $Z_k + (-a_{kj_\ell})Z_\ell$.

Satz 2.27. *Sei* $A \in K^{m \times n}$ *und sei* $b \in K^{m \times 1}$. *Angenommen, das (inhomogene) lineare Gleichungssytem* $A\,x = b$ *sei lösbar. Dann gilt:*

$$\text{Lös}\,(A\,,\,b) = \{\, z + y \mid y \in \text{Lös}\,(A\,,\,0)\,\} \quad ,$$

wobei z *eine spezielle Lösung des (inhomogenen) linearen Gleichungssytems* $A\,x = b$ *ist, d.h.* $z \in \text{Lös}\,(A\,,\,b)$.

Satz 2.28. (Parameterdarstellung der Lösungsmenge) Sei $A \in K^{m \times n}$ eine Matrix in reduzierter Gaußscher Normalform mit r Stufen und sei $b \in K^{m \times 1}$. Angenommen, das (inhomogene) lineare Gleichungssytem $A\,x = b$ sei lösbar. Dann gilt:

$$\text{Lös}\,(A\,,\,b) = \left\{ z + y \;\middle|\; y = \sum_{k=1}^{n-r} \lambda_k\, y_k \;\text{ mit }\; \lambda_1\,,\ldots,\lambda_{n-r} \in K \right\} \quad ,$$

wobei $z\,,\,y_1\,,\,\ldots\,,\,y_{n-r}$ wie folgt gegeben sind: Sei

$$\{\,\ell_1\,,\ldots,\ell_{n-r}\,\} := \{\,1\,,\,\ldots\,,\,n\,\} \backslash \{\,j_1\,,\ldots,j_r\,\} \quad \text{mit} \quad 1 \le \ell_1 < \ell_2 < \ldots < \ell_{n-r} \le n \quad .$$

- z ist die Spalte, die an den Stellen j_i für $1 \le i \le r$ den Eintrag b_i hat und an allen anderen Stellen nur 0, d.h. für $z = (\,z_1\,z_2\ldots z_n\,)^T$ gilt:

$$z_i = \begin{cases} b_i\,, & \text{falls } i \in \{\,j_1\,,\ldots,j_r\,\} \\ 0\,, & \text{falls } i \in \{\,\ell_1\,,\ldots,\ell_{n-r}\,\} \end{cases} \quad .$$

- Wir definieren y_k für alle $k = 1\,,\,\ldots\,,\,n-r$ als die Spalte, welche an den Stellen j_i für $1 \le i \le r$ den Eintrag $-a_{i\ell_k}$ hat, zudem nur an der Stelle ℓ_k den Eintrag 1 und sonst 0.

Der Spezialfall, dass $j_i = i$ für $1 \le i \le r$, wird in der Vorlesung behandelt.

Wir sind nun in der Lage, die beiden Algorithmen zur Lösung eines linearen Gleichungssystems $A\,x = b$ zu lösen, wobei A und b gegeben sind.

Gaußscher Algorithmus: Das 1. Verfahren arbeitet mit Rücksubsitution.

(1) Schreibe A und b als erweiterte Koeffizientenmatrix $(\,A \mid b\,)$.

(2) Wende elementare Zeilenumformungen vom Typ I, II und III auf $(\,A \mid b\,)$ an, um A in Gaußsche Normalform zu bringen. Damit erhält man $(\,\tilde{A} \mid \tilde{b}\,)$ und ein äquivalentes lineares Gleichungssystem $\tilde{A}\,x = \tilde{b}$.

(3) Bestimme anhand der Gaußschen Normalform, ob das lineare Gleichungssystem $\tilde{A}\,x = \tilde{b}$ lösbar ist. Verwende dazu Satz 2.24. Sollte das lineare Gleichungssystem $\tilde{A}\,x = \tilde{b}$ unlösbar sein, so STOP.

(4) Wende Rücksubstitution der Variablen an, um die Lösung zu finden. Dabei setze gegebenenfalls Parameter ein.

(5) Teste das Ergebnis auf Korrektheit durch Einsetzen in die ursprüngliche Gleichung. (Hinweis: Das sollte JEDER sowieso tun.)

Gauß-Jordan-Verfahren: Das 2. Verfahren wendet im Prinzip Satz 2.28 an.

(a) Schreibe A und b als erweiterte Koeffizientenmatrix $(\,A \mid b\,)$.

(b) Wende elementare Zeilenumformungen vom Typ I, II und III auf $(\,A \mid b\,)$ an, um A in reduzierte Gaußsche Normalform zu bringen. Damit erhält man $(\,A^* \mid b^*\,)$ und ein äquivalentes lineares Gleichungssystem $A^*\,x = b^*$.

(c) Bestimme anhand der reduzierten Gaußschen Normalform, ob das lineare Gleichungssystem $A^*\,x = b^*$ lösbar ist. Verwende dazu Satz 2.24. Sollte das lineare Gleichungssystem $A^*\,x = b^*$ unlösbar sein, so STOP.

(d) Bestimme die Parameterdarstellung der Lösung durch Anwendung von Satz 2.28 oder einfach durch Rücksubstitution.

(e) Teste das Ergebnis auf Korrektheit durch Einsetzen in die ursprüngliche Gleichung. (Hinweis: Das sollte JEDER sowieso tun.)

2.5 Theoretische Resultate zu linearen Gleichungssystemen

Zum Abschluss dieses Kapitels formulieren wir einige wichtige Resultate, welche von großem theoretischem Interesse sind.

Satz 2.29. *Sei* $b \in K^{n\times 1}$. *Für jede Matrix* $A \in \mathrm{GL}(n, K)$ *hat das lineare Gleichungssystem* $A\,x = b$ *eine eindeutige Lösung.*

Das folgende Lemma ist ein Hilfslemma.

Lemma 2.30. *Sei* $A \in K^{n\times n}$ *eine quadratische Matrix.*

(a) *Existiert eine invertierbare Matrix* $B \in \mathrm{GL}(n, K)$ *mit* $B\,A = E_n$, *so ist auch* A *invertierbar und* $B = A^{-1}$.

(b) *Existiert eine invertierbare Matrix* $B \in \mathrm{GL}(n, K)$ *mit* $A\,B = E_n$, *so ist auch* A *invertierbar und* $B = A^{-1}$.

Satz 2.31. *Es seien* $A \in K^{n\times n}$ *und sei* $b \in K^{n\times 1}$. *Dann sind die folgenden Aussagen äquivalent:*

(a) A *ist invertierbar, also* $A \in \mathrm{GL}(n, K)$.

(b) *Ist* A^* *eine reduzierte Gaußsche Normalform von* A *mit* r *Stufen, so gilt* $r = n$ *und damit* $A^* = E_n$.

(c) *Das lineare Gleichungssystem* $A\,x = 0$ *hat nur die triviale Lösung* $x = 0$.

(d) *Das lineare Gleichungssystem* $A\,x = b$ *hat eine eindeutige Lösung, nämlich die Lösung* $x_0 := A^{-1}\,b$.

(e) A *ist ein Produkt von Elementarmatrizen.*

Satz 2.32. *Es seien $A, B \in K^{m\times n}$ Matrizen in reduzierter Gaußscher Normalform. Existiert ein $Q \in \mathrm{GL}(m, K)$ mit $A = QB$, so gilt $A = B$. Insbesondere ist die reduzierte Gaußsche Normalform einer Matrix eindeutig.*

Nun sind wir in der Lage, ein einfacheres Kriterium für die Invertierbarkeit anzugeben. Zu beachten ist, dass hier nur noch verlangt wird, dass B quadratisch ist und dass nur entweder Multiplikation von links oder von rechts (nicht beide) benötigt wird.

Satz 2.33. *Sei $A \in K^{n\times n}$ eine quadratische Matrix.*

(a) *Existiert eine quadratische Matrix $B \in K^{n\times n}$ mit $BA = E_n$, so ist auch A invertierbar und $B = A^{-1}$.*

(b) *Existiert eine quadratische Matrix $B \in K^{n\times n}$ mit $AB = E_n$, so ist auch A invertierbar und $B = A^{-1}$.*

Korollar 2.34. *Sei $A \in K^{m\times n}$. Gilt $m < n$, so hat das homogene lineare Gleichungssystem $Ax = 0$ mindestens 2 Lösungen.*

Algorithmus zum Invertieren

Gegeben sei eine quadratische Matrix $A \in K^{n \times n}$, für die wir Satz 2.31 anwenden.

(1) Schreibe A und E_n als erweiterte Koeffizientenmatrix $(\, A \mid E_n \,)$.

(2) Wende elementare Zeilenumformungen vom Typ I, II und III auf $(\, A \mid E_n \,)$ an, um A in reduzierte Gaußsche Normalform A^* zu bringen. Damit erhält man $(\, A^* \mid Q \,)$, wobei $A^* = Q\,A$ und $Q \in \mathrm{GL}(\, n\,, K\,)$.

(3) Verwende Satz 2.31: Gilt $A^* = E_n$, so ist A invertierbar und $Q = A^{-1}$. Gilt $A^* \neq E_n$, so ist A nicht invertierbar.

(4) Falls A invertierbar ist, so teste das Ergebnis auf Korrektheit durch Multiplikation: $A\,A^{-1} = E_n$ bzw. $A^{-1}\,A = E_n$?

Der Algorithmus kann auch leicht modifizert werden, sofern die Invertierbarkeit unbekannt ist. Man könnte A einfach in Gaußsche Normalform $\tilde{A}$ bringen. Gilt $r < n$ bzw. taucht eine Nullzeile in der Matrix $\tilde{A}$ auf, so ist A nicht invertierbar. Ist $r = n$ und damit $\tilde{A}$ in oberer Dreiecksgestalt, so ist A invertierbar. Man führt dabei auch Operationen auf $(\, A \mid E_n \,)$ aus. Im Falle von Invertierbarkeit führt man weitere elementare Zeilenumformungen vom Typ I, II und III auf die erweiterte Matrix aus, um A in reduzierte Gaußsche Normalform A^* zu bringen. Im Falle von Invertierbarkeit erhalten wir $A^* = E_n$ und $(\, A^* \mid Q \,) = (E_n \mid A^{-1}\,)$.

Satz 2.35. *Sei $A \in K^{m \times n}$. Dann kann man A mit einer endlichen Folge von elementaren Zeilen- und Spaltenumformungen in die Form*

$$B = \begin{pmatrix} E_r & 0_{r\,,(n-r)} \\ 0_{(m-r)\,,r} & 0_{(m-r)\,,(n-r)} \end{pmatrix}$$

überführen, wobei E_r die r-te Einheitsmatrix ist und $0_{r \times (n-r)}$, $0_{(m-r)\,,r}$, $0_{(m-r)\,,(n-r)}$ die Nullmatrizen mit den angegebenen Größen sind. Außerdem existieren invertierbare Matrizen $Q \in \mathrm{GL}(\, m\,, K\,)$ und $P \in \mathrm{GL}(\, n\,, K\,)$, so dass

$$Q\,A\,P = B \ .$$

3 Vektorräume

3.1 Definition und elementare Eigenschaften von Vektorräumen

Definition Sei K ein Körper. Eine nichtleere Menge V heißt ein K**-Vektorraum** (oder **Vektorraum über** K), wenn es eine innere Verknüpfung

$$\begin{array}{rccc} + : & V \times V & \to & V \\ & (v, w) & \mapsto & v + w \end{array},$$

die so genannte Addition (oder Vektoraddition) und eine äußere Verknüpfung

$$\begin{array}{rccc} \cdot : & K \times V & \to & V \\ & (\lambda, v) & \mapsto & \lambda \cdot v \end{array},$$

die so genannte Skalarmultiplikation gibt, so dass die folgenden Axiome gelten:

(A) $(V, +)$ ist eine abelsche Gruppe.

- **(A0)** **Abgeschlossenheit:** Sind $v, w \in V$, so ist auch $v + w \in V$.
- **(A1)** **Assoziativität:** $(u + v) + w = u + (v + w)$ für alle $u, v, w \in V$.
- **(A2)** **Neutrales Element der Addition:** Es gibt ein Element $0 \in V$, so dass $0 + v = v = v + 0$ für alle $v \in V$.
- **(A3)** **Inverses Element der Addition:** Zu jedem $v \in V$ gibt es ein $-v \in V$, so dass $(-v) + v = v + (-v) = 0$.
- **(A4)** **Kommutativität:** $v + w = w + v$ für alle $v, w \in V$.

(S) Regeln der Skalarmultiplikation und der Addition.

- **(S0))** **Abgeschlossenheit:** Für beliebige $\lambda \in K$ und $v \in V$ ist auch $\lambda \cdot v \in V$.
- **(S1)** **Unität:** $1 \cdot v = v$ für alle $v \in V$.
- **(S2)** **Assoziativität:** $(\lambda \mu) \cdot v = \lambda \cdot (\mu \cdot v)$ für alle $\lambda, \mu \in K$ und $v \in V$.
- **(S3)** **Distributivität 1:** $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ für alle $\lambda \in K$ und $v, w \in V$.
- **(S4)** **Distributivität 2:** $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$ für alle $\lambda, \mu \in K$ und $v \in V$.

Wenn die Situation klar ist, schreiben wir λv anstelle von $\lambda \cdot v$. Die Elemente aus V heißen **Vektoren** und die Elemente aus K heißen **Skalare**.

Satz 3.1. *Es seien K ein Körper und V ein K-Vektorraum. Dann gilt für alle $\lambda \in K$ und für alle $v, w \in V$:*

(a) $0 \cdot v = 0$.

(b) $\lambda \cdot 0 = 0$.

(c) $(-\lambda) \cdot v = \lambda \cdot (-v) = -(\lambda \cdot v)$. *Insbesondere ist* $(-1) \cdot v = -v$.

(d) *Falls* $\lambda \cdot v = 0$, *dann ist entweder* $\lambda = 0$ *oder* $v = 0$.

(e) *Kürzungsregel: Falls* $\lambda \cdot v = \lambda \cdot w$, *dann ist entweder* $\lambda = 0$ *oder* $v = w$.

(f) *Der Vektor* $0 \in V$ *ist eindeutig bestimmt, und zu jedem* $v \in V$ *ist* $-v \in V$ *eindeutig bestimmt.*

Exkurs: Polynome

Wir führen hier die aus der Schule und der Analysis bekannten Polynome ein. In der Algebra I werden wir Polynome noch formaler über Ringen einführen und auf die Bedeutung der Unbestimmten eingehen. Hier benötigen wir vorerst nur die Koeffizienten, den Grad und die Tatsache, dass wir einsetzen dürfen. Wir unterscheiden zwischen formalen **Polynomen** und **Polynomabbildungen**.

Polynomabbildungen

Sei K ein Körper. Wir nennen $f \in \mathrm{Abb}(K, K)$ eine **Polynomabbildung** bzw. **Polynomfunktion** auf K, wenn ein $n \in \mathbb{N}_0$ und **Koeffizienten** $a_0, a_1, \ldots, a_n \in K$ von f existieren, so dass

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \qquad (\forall x \in K) .$$

Bemerkung 3.2. *Sei K ein Körper.*

(a) *Für $K = \mathbb{R}$ gilt immer, dass zwei reelle Polynomabbildungen nur dann dieselben Funktionswerte liefern, wenn sie dieselben Koeffizienten haben. Dies gilt zum Beispiel nicht in $\mathbb{Z}_p$, wenn p prim ist. Zwei Elemente $f, g \in \mathrm{Abb}(K, K)$ sind genau dann gleich, wenn $f(x) = g(x)$ für alle $x \in K$. Beispielsweise gilt für die Funktion $f : \mathbb{Z}_p \to \mathbb{Z}_p$ mit $f(x) := x^p - x$: Für alle $x \in \mathbb{Z}_p$ ist $f(x) = 0$, aber es ist sicherlich $f(x)$ nicht die Nullfunktion.*

(b) *Die Menge der Polynomabbildungen auf K bildet einen Ring unter punktweiser Addition und Multiplikation.*

Formale Polynome

Sei K ein Körper. Für eine allgemeinere Definition von Polynomen verweisen wir auf die Algebra oder die Literatur. Ein **Polynom f in der Unbestimmten t über K** ist ein formaler Ausdruck der Form

$$f = a_0 + a_1 t + \ldots + a_{n-1} t^{n-1} + a_n t^n = \sum_{j=0}^{n} a_j t^j \ ,$$

wobei $n \in \mathbb{N}_0$ und $a_j \in K$ für alle $j = 0, \ldots, n$. Manchmal schreibt man auch $f(t)$ anstelle von f.

- Die Elemente $a_0, a_1, \ldots, a_n \in K$ heißen die **Koeffizienten** von f.
- Zwei Polynome $f = \sum_{j=0}^{n} a_j t^j$ und $g = \sum_{j=0}^{n} b_j t^j$ sind genau dann gleich, wenn $a_j = b_j$ für alle $j \in \{0, \ldots, n\}$.

- Die Menge aller Polynome über K in der Unbestimmten t bezeichnen wir mit
$$K[t] := \left\{ f = a_0 + a_1 t + \ldots + a_n t^n = \sum_{j=0}^{n} a_j t^j \mid n \in \mathbb{N}_0, a_j \in K, 0 \leq j \leq n \right\} .$$
- Das **Nullpolynom** ist das Polynom $f = 0$, dessen Koeffizienten alle gleich 0 sind.
- Für ein Polynom $f = \sum_{j=0}^{n} a_j t^j \in K[t]$ definieren wir den **Grad** von f als
$$\deg(f) = \deg f = \begin{cases} \max\{ j \in \{0, \ldots, n\} \mid a_j \neq 0\} & \text{, falls } f \neq 0 \\ -\infty & \text{, falls } f = 0 \end{cases}$$
D.h. ist z.B. $a_n \neq 0$, so ist $\deg(f) = n$.
- Ein Polynom $f = \sum_{j=0}^{n} a_j t^j \in K[t]$ vom Grad n heißt **normiert**, wenn $a_n = 1$.

Wir definieren nun drei Operationen auf $K[t]$. Es seien
$$f = \sum_{j=0}^{n} a_j t^j \quad , \quad g = \sum_{j=0}^{m} b_j t^j \in K[t] .$$

- **Addition:** Wir nehmen ohne Beschränkung der Allgemeinheit $m = n$ an.
[Ist zum Beispiel $m < n$, so setzen wir $b_{m+1} = b_{m+2} = \ldots = b_n = 0$.]
$$f + g = \sum_{j=0}^{n} a_j t^j + \sum_{j=0}^{n} b_j t^j := \sum_{j=0}^{n} (a_j + b_j) t^j$$
Man führt komponentenweise Addition wie folgt aus:
$$(a_0 + a_1 t + \ldots + a_n t^n) + (b_0 + b_1 t + \ldots + b_n t^n) = (a_0 + b_0) + (a_1 + b_1) t + \ldots + (a_n + b_n) t^n .$$
- **Multiplikation:**
$$f \cdot g := c_0 + c_1 t + \ldots + c_{n+m} t^{n+m} = \sum_{k=0}^{n+m} c_k t^k ,$$
wobei wir $c_k \in K$ für alle $k \in \mathbb{N}_0$ wie folgt definieren:
$$c_k := a_0 b_k + a_1 b_{k-1} + \ldots + a_{k-1} b_1 + a_k b_0 = \sum_{i+j=k} a_i b_j .$$
Zum Beispiel sind
$$c_0 = a_0 b_0 \quad , \quad c_1 = a_0 b_1 + a_1 b_0 \quad , \quad c_{n+m} = a_n b_m .$$
- **Skalarmultiplikation:** Für jedes $\lambda \in K$ definieren wir
$$\lambda \cdot f := (\lambda a_0) + (\lambda a_1) t + \ldots + (\lambda a_n) t^n = \sum_{j=0}^{n} (\lambda a_j) t^j .$$

Proposition 3.3. *Sei K ein Körper. Dann gilt:*

(a) *$K[t]$ ist unter der oben definierten Addition und Multiplikation ein kommutativer, nullteilerfreier Ring, aber kein Körper. Man nennt deshalb $K[t]$ auch den **Polynomring in der Unbestimmten t über K**.*

(b) *Es seien $f, g \in K[t]$. Dann ist*

$$\deg(f+g) \leq \max\{\deg(f), \deg(g)\} \quad , \quad \deg(fg) = \deg(f) + \deg(g) \; .$$

Für $f = 0$ oder $g = 0$ rechnen wir $n - \infty = -\infty + m = -\infty + (-\infty) = -\infty$.

(c) *$K[t]$ ist unter Polynomaddition und Skalarmultiplikation ein K-Vektorraum.*

Wichtig ist, dass wir in Polynome aus $K[t]$ Elemente aus K **einsetzen** können. Wir werden in der Algebra I genauer sehen, was das bedeutet. Dort werden wir den wichtigen Einsetzungshomomorphismus verfolgen. Hier gehen wir wie folgt vor: Sei $\alpha \in K$ und sei $f = \sum_{j=0}^{n} a_j t^j \in K[t]$. Dann definieren wir die Einsetzung von α in f als

$$f(\alpha) := \sum_{j=0}^{n} a_j \alpha^j = a_0 + a_1 \alpha + \ldots + a_n \alpha^n \quad \in K \; .$$

Gilt $f(\alpha) = 0$ für ein $\alpha \in K$, so heißt α eine **Nullstelle** von f.
Außerdem rechnen wir für alle $f, g \in K[t]$ und $\lambda, \alpha \in K$:

(a) $(f+g)(\alpha) = f(\alpha) + g(\alpha)$.

(b) $(f \cdot g)(\alpha) = f(\alpha) \cdot g(\alpha)$.

(c) $(\lambda f)(\alpha) = \lambda f(\alpha)$.

Einsetzen liefert für jedes $f \in K[t]$ eine Polynomabbildung $\tilde{f} \in \mathrm{Abb}(K, K)$, wobei

$$\begin{array}{rccl} \tilde{f}: & K & \to & K \\ & \alpha & \mapsto & \tilde{f}(\alpha) := f(\alpha) = \sum_{j=0}^{n} a_j \alpha^j \; . \end{array}$$

Die Abbildung $\tilde{f} \in \mathrm{Abb}(K, K)$ wird die zum Polynom f gehörige Polynomabbildung genannt. Insgesamt erhält man damit eine im Allgemeinen nicht bijektive Abbildung

$$\begin{array}{rccl} \Phi: & K[t] & \to & \mathrm{Abb}(K, K) \\ & f & \mapsto & \Phi(f) := \tilde{f} \; . \end{array}$$

Proposition 3.4. (**Division mit Rest in $K[t]$**) *Sei K ein Körper. Sind $f, g \in K[t]$ mit $g \neq 0$, dann gibt es eindeutige Polynome $q, r \in K[t]$, derart dass*

$$f = qg + r \quad \textit{und} \quad \deg(r) < \deg(g) \; .$$

Den Beweis der letzten Proposition überlassen wir als Übung.

Untervektorräume

Nach diesem Exkurs in die wichtigsten Aussagen über Polynome betrachten wir die Theorie der Untervektorräume. Ähnlich wie bei Gruppen und Untergruppen wird uns damit das Rechnen wesentlich vereinfacht.

Definition Es seien K ein Körper und V ein K-Vektorraum. Eine nichtleere Teilmenge $U \subseteq V$ heißt ein K-**Untervektorraum** oder K-**Unterraum** von V, wenn U zusammen mit der von V induzierten Addition und Skalarmultiplikation ein K-Vektorraum ist.

Satz 3.5. (**Untervektorraum-Test I**) *Sei K ein Körper und sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ ist genau dann ein Untervektorraum von V, wenn die folgenden Bedingungen erfüllt sind:*

(UV1) $U \neq \varnothing$.

(UV2) ***Abgeschlossenheit bzgl. der Addition:*** *Für alle $u, v \in U$ gilt $u+v \in U$.*

(UV3) ***Abgeschlossenheit bzgl. der Skalarmultiplikation:*** *Für beliebige $u \in U$ und $\lambda \in K$ ist auch $\lambda \cdot u \in U$.*

Satz 3.6. (**Untervektorraum-Test II**) *Sei K ein Körper und sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ ist genau dann ein Untervektorraum von V, wenn die folgenden Bedingungen erfüllt sind:*

(UV1) $U \neq \varnothing$.

(UV2') *Für alle $u, v \in U$ und $\lambda, \mu \in K$ ist auch $\lambda \cdot u + \mu \cdot v \in U$.*

3.2 Erzeugendensysteme, Lineare Unabhängigkeit und Basen

Die Begriffe Erzeugendensysteme, Spann, lineare Unabhängigkeit, Basen sowie Dimension sind ultimativ wichtig in der Theorie der Vektorräume.

Definition Sei V ein K-Vektorraum über einem Körper K und sei $(v_i)_{i\in I}$ eine Familie von Vektoren aus V bzgl. einer Indexmenge I. Dann heißt ein Ausdruck der Form

$$\sum_{i\in I}\lambda_i v_i \qquad \text{mit } \lambda_i \in K \text{ und } \lambda_i = 0 \text{ für fast alle } i \in I$$

eine **Linearkombination** von $(v_i)_{i\in I}$. Die Körperelemente λ_i, $i \in I$, heißen die **Koeffizienten** dieser Linearkombination. Die Menge aller solcher Linearkombinationen ist

$$\mathrm{Spann}_K(v_i)_{i\in I} := \left\{ \sum_{i\in I}\lambda_i v_i \,\middle|\, \lambda_i \in K, \lambda_i = 0 \text{ für fast alle } i \in I \right\} .$$

Eine solche Linearkombination heißt **trivial**, wenn $\lambda_i = 0$ für alle $i \in I$ gilt.
Schreibweise: $\mathrm{Spann}_K(v_i)$ oder $\langle (v_i)_{i\in I}\rangle$ anstelle von $\mathrm{Spann}_K(v_i)_{i\in I}$.

Bemerkung 3.7. *Sei V ein K-Vektorraum über einem Körper K.*

(a) *Für eine beliebige Familie $(v_i)_{i\in I}$ von Vektoren in V und $v \in \mathrm{Spann}_K(v_i)_{i\in I}$ existiert ein Index r, nämlich $r = \max\{i \mid \lambda_i \neq 0\}$, so dass*

$$v = \lambda_1 v_{i_1} + \lambda_2 v_{i_2} + \ldots + \lambda_r v_{i_r} = \sum_{j=1}^{r} \lambda_j v_{i_j} \qquad (\lambda_j \in K, v_{i_j} \in V, 1 \le j \le r) .$$

(b) *Im Spezialfall $I = \{1, 2\ldots, n\}$ nennen wir dann*

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i \qquad (\lambda_1, \lambda_2, \ldots, \lambda_n \in K)$$

eine Linearkombination von $(v_1, v_2, \ldots, v_n)$ mit Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$.

$$\begin{aligned}\mathrm{Spann}_K(v_1, v_2, \ldots, v_n) &= \left\{ \sum_{i=1}^{n} \lambda_i v_i \,\middle|\, \lambda_i \in K, 1 \le i \le n \right\} \\ &= \left\{ u \in V \,\middle|\, \exists \lambda_1, \ldots, \lambda_n \in K, \text{ so dass } u = \sum_{i=1}^{n} \lambda_i v_i \right\} .\end{aligned}$$

(c) *Ist $I = \{1\}$, so ist $\mathrm{Spann}_K(v_1) = K v_1$.*

(d) *Ist $I = \varnothing$, so setzen wir $\mathrm{Spann}_K(v_i)_{i\in I} := \{0\}$ mit dem Nullvektor $0 \in V$.*

Satz 3.8. *Sei V ein K-Vektorraum über einem Körper K und sei $(v_i)_{i \in I}$ eine Familie von Vektoren aus V. Dann gilt:*

(a) $\mathrm{Spann}_K(v_i)_{i \in I}$ *ist ein K-Untervektorraum von V.*

(b) $\mathrm{Spann}_K(v_i)_{i \in I}$ *ist der kleinste K-Untervektorraum von V, der alle v_i enthält, also*

$$\mathrm{Spann}_K(v_i)_{i \in I} = \bigcap_{\substack{U \text{ Untervektorraum von } V \\ \text{und } v_i \in U \ \forall i \in I}} U$$

Dies ist gleichbedeutend mit der folgenden Aussage: Ist $U \subseteq V$ ein Untervektorraum von V und gilt $v_i \in U$ für alle $i \in I$, so ist $\mathrm{Spann}_K(v_i) \subseteq U$.

Korollar 3.9. *Es seien V ein K-Vektorraum über einem Körper K und $v_1, \ldots, v_n \in V$. Dann ist $\mathrm{Spann}(v_1, \ldots, v_n)$ ein Untervektorraum von V. Dies ist der kleinste Untervektorraum von V, der $v_1, \ldots, v_n$ enthält, d.h.*

$$\mathrm{Spann}_K(v_1, \ldots, v_n) = \bigcap_{\substack{U \text{ Untervektorraum von } V \\ \text{und } v_1, \ldots, v_n \in U}} U$$

Definition Sei V ein K-Vektorraum über einem Körper K und sei $(v_i)_{i \in I}$ eine Familie von Vektoren aus V. Wir nennen $(v_i)_{i \in I}$ ein **Erzeugendensystem** von V, wenn

$$V = \mathrm{Spann}_K(v_i)_{i \in I} .$$

Der Vektorraum V heißt **endlich erzeugt**, wenn es ein endliches Erzeugendensystem von V gibt, d.h. wenn endlich viele Vektoren $v_1, \ldots, v_n \in V$ existieren, so dass

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

Wir sagen in dem Fall, das Erzeugendensystem $\mathcal{B} = (v_1, \ldots, v_n)$ habe eine **endliche Länge** n, d.h. $\text{Länge}(\mathcal{B}) := n$.

Die Vereinigung von Vektoren eines Vektorraumes V ist i.A. kein Vektorraum. Man kann jedoch Vektoren zu einem Vektorraum abschließen, indem man den kleinsten Untervektorraum des Vektorraumes betrachtet, der diese Vektoren enthält.

Satz 3.10. *Sei V ein K-Vektorraum über einem Körper K und seien $v, v_1, \ldots, v_n \in V$. Gilt $V = \mathrm{Spann}_K(v_1, \ldots, v_n)$, so gilt auch $V = \mathrm{Spann}_K(v, v_1, \ldots, v_n)$.*

Satz 3.11. *Sei V ein K-Vektorraum über einem Körper K. Seien $v_1, \ldots, v_m \in V$, so dass $V = \mathrm{Spann}_K(v_1, \ldots, v_m)$. Sind $w_1, \ldots, w_n \in V$ mit*

$$v_1, \ldots, v_m \in \mathrm{Spann}_K(w_1, \ldots, w_n) \ ,$$

so gilt auch

$$V = \mathrm{Spann}_K(w_1, \ldots, w_n) \ .$$

Definition Sei V ein K-Vektorraum über einem Körper K und seien $v_1, \ldots, v_n \in V$. Die endliche Familie $(v_1, \ldots, v_n)$ heißt **linear unabhängig über** K bzw. K**-linear unabhängig**, wenn sich die 0 von V nur als triviale Linearkombination der Vektoren $v_1, \ldots, v_n \in V$ darstellen lässt. D.h.

$$(\forall \lambda_1, \ldots, \lambda_n \in K) \quad \left[\sum_{i=1}^{n} \lambda_i v_i = \lambda_1 v_1 + \ldots + \lambda_n v_n = 0 \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_n = 0 \right] \ .$$

Andernfalls heißt die Familie $(v_1, \ldots, v_n)$ **linear abhängig über** K bzw. K**-linear abhängig**. Die Familie $(v_1, \ldots, v_n)$ ist also linear abhängig über K, wenn $\lambda_1, \ldots, \lambda_n \in K$ existieren, die nicht alle gleich 0 sind, so dass

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i = 0 \ .$$

Definition Sei V ein K-Vektorraum über einem Körper K. Eine beliebige Familie $(v_i)_{i \in I}$ von Vektoren in V heißt **linear unabhängig über** K bzw. K**-linear unabhängig**, wenn jede endliche Teilfamilie von $(v_i)_{i \in I}$ linear unabhängig über K ist.

$$(\forall \lambda_i \in K \text{ mit } \lambda_i = 0 \text{ für fast alle } i \in I) \left[\sum_{i \in I} \lambda_i v_i = 0 \Rightarrow \lambda_i = 0 \text{ für alle } i \in I \right] \ .$$

Ansonsten heißt die Familie $(v_i)_{i \in I}$ **linear abhängig über** K bzw. K**-linear abhängig**.

Bemerkung 3.12. *Sei V ein K-Vektorraum über einem Körper K und sei $(v_i)_{i \in I}$ eine beliebige Familie von Vektoren in V.*

(a) *Die Familie $(v_i)_{i \in I}$ heißt linear unabhängig über K, wenn für jede endliche Teilfamilie $(v_{i_1}, v_{i_2}, \ldots, v_{i_r})$ gilt:*

$$(\forall \lambda_1, \ldots, \lambda_n \in K) \quad [\lambda_1 v_{i_1} + \ldots + \lambda_r v_{i_r} = 0 \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_r = 0] \ .$$

(b) *Die Familie $(v_i)_{i \in I}$ heißt linear abhängig über K, wenn es eine endliche Teilfamilie $(v_{i_1}, v_{i_2}, \ldots, v_{i_r})$ und $\lambda_1, \ldots, \lambda_r \in K$ gibt, die nicht alle gleich 0 sind, so dass*

$$\lambda_1 v_{i_1} + \lambda_2 v_{i_2} + \ldots + \lambda_r v_{i_r} = 0 \ .$$

Satz 3.13. *Sei V ein K-Vektorraum über einem Körper K und sei $(v_i)_{i\in I}$ eine Familie von Vektoren in V. Sei $v \in V$ ein beliebiger Vektor. Dann gilt:*

(a) *Die einelementige Familie (v) ist genau dann K-linear unabhängig, wenn $v \neq 0$.*

(b) *Ist $v_j = 0$ für ein $j \in I$, so ist $(v_i)_{i\in I}$ stets linear abhängig. Dasselbe gilt, wenn sich ein Vektor wiederholt, es also Indizes $i, j \in I$, $i \neq j$, gibt, so dass $v_i = v_j$.*

(c) *Ist $(v_i)_{i\in I}$ linear unabhängig über K, so ist $(v_i)_{i\in J}$ K-linear unabhängig für jede Indexmenge $J \subseteq I$. Ist. z.B. $(v_1, \dots, v_n)$ K-linear unabhängig, so auch $(v_{i_1}, v_{i_2}, \dots, v_{i_r})$ für jede paarweise verschiedenen Indizes $i_1, \dots, i_r \in \{1, \dots, n\}$.*

(d) *Sei $k \in \mathbb{N}_{\geq 2}$. Die Familie $(v_1, \dots, v_k)$ ist genau dann linear abhängig über K, wenn einer dieser Vektoren eine Linearkombination der anderen ist, d.h. wenn ein Index $j \in \{1, \dots, k\}$ existiert, so dass $v_j \in \mathrm{Spann}_K(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_k)$.*

Satz 3.14. *Sei V ein K-Vektorraum über einem Körper K und sei $(v_i)_{i\in I}$ eine Familie von Vektoren in V. Dann sind die folgenden Aussagen äquivalent:*

(a) *$(v_i)_{i\in I}$ ist linear unabhängig über K.*

(b) *Jeder Vektor $v \in \mathrm{Spann}_K(v_i)_{i\in I}$ lässt sich eindeutig als Linearkombination von $(v_i)_{i\in I}$ darstellen. D.h. für alle $v \in \mathrm{Spann}_K(v_i)_{i\in I}$ existieren eindeutige $\lambda_i \in K$ $(i \in I)$, so dass*

$$v = \sum_{i\in I} \lambda_i v_i \quad \textit{und} \quad \lambda_i = 0 \textit{ für fast alle } i \in I .$$

Anwendung auf $V = K^{n\times 1}$

Wir betrachten als Beispiel den Vektorraum $V = K^{n\times 1}$ über einem Körper K unter den Standardoperationen. Dann sind wir in der Lage, einige algorithmische Probleme sofort mittels Gaußschem Algorithmus oder Gauß-Jordan Algorithmus zu lösen. Mit leichter Modifikation kann man andere endlich-erzeugte K-Vektorräume betrachten.

Problem 1 Gegeben $v_1, \ldots, v_k \in K^{n\times 1}$ und $b \in K^{n\times 1}$; liegt $b \in \mathrm{Spann}_K(v_1, \ldots, v_k)$?

Wir fragen also, ob $\lambda_1, \ldots, \lambda_k \in K$ existieren, so dass

$$b = \lambda_1 v_1 + \ldots + \lambda_k v_k \ ?$$

Lösung: Wir schreiben die Vektoren in eine Matrix A und suchen eine Lösung $x \in K^{k\times 1}$ von $Ax = b$, wobei

$$A := \begin{pmatrix} v_1 & v_2 & \ldots & v_k \end{pmatrix} \in K^{n\times k} \quad , \quad x := \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_k \end{pmatrix}^T \in K^{k\times 1} \ .$$

Auf die Gaußsche Normalform $(\tilde{A} \mid \tilde{b})$ oder die reduzierte Gaußsche Normalform $(A^* \mid b^*)$ der erweiterten Koeffizientenmatrix $(A \mid b)$ von $Ax = b$ wenden wir nun das Lösbarkeitskriterium (Satz 2.24) an und stellen einfach fest, ob $b \in \mathrm{Spann}_K(v_1, \ldots, v_k)$ oder nicht. Falls $b \in \mathrm{Spann}_K(v_1, \ldots, v_k)$, so können wir eine Linearkombination sofort angeben.

Problem 2 Gegeben $v_1, \ldots, v_k \in K^{n\times 1}$; gilt $K^{n\times 1} = \mathrm{Spann}_K(v_1, \ldots, v_k)$?

Wir fragen also, ob für einen beliebigen Vektor $b \in K^{n\times 1}$ Skalare $\lambda_1 = \lambda_1(b), \ldots, \lambda_k = \lambda_k(b) \in K$ existieren, so dass

$$b = \lambda_1 v_1 + \ldots + \lambda_k v_k \ ?$$

Lösung: Wir schreiben die Vektoren in eine Matrix A und suchen eine Lösung $x \in K^{k\times 1}$ von $Ax = b$, wobei

$$A := \begin{pmatrix} v_1 & v_2 & \ldots & v_k \end{pmatrix} \in K^{n\times k} \quad , \quad x := \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_k \end{pmatrix}^T \in K^{k\times 1} \ .$$

Auf die Gaußsche Normalform $(\tilde{A} \mid \tilde{b})$ oder die reduzierte Gaußsche Normalform $(A^* \mid b^*)$ der erweiterten Koeffizientenmatrix $(A \mid b)$ von $Ax = b$ wenden wir nun das Lösbarkeitskriterium (Satz 2.24) an.

Problem 3 Gegeben $v_1, \ldots, v_k \in K^{n\times 1}$; ist die Familie $(v_1, \ldots, v_k)$ linear unabhängig?

Wir fragen also, ob die 0 nur als triviale Linearkombination dieser Vektoren darstellbar ist. Seien $\lambda_1, \ldots, \lambda_k \in K$, so dass

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = 0 \ .$$

Lösung: Wir schreiben die Vektoren in eine Matrix A und suchen eine Lösung $x \in K^{k\times 1}$ von $Ax = 0$, wobei

$$A := \begin{pmatrix} v_1 & v_2 & \ldots & v_k \end{pmatrix} \in K^{n\times k} \quad , \quad x := \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_k \end{pmatrix}^T \in K^{k\times 1} \ .$$

Wir bringen die erweiterte Koeffizientenmatrix $(A \mid 0)$ von $Ax = 0$ in Gaußsche Normalform $(\tilde{A} \mid 0)$ oder in reduzierte Gaußsche Normalform $(A^* \mid 0)$ und können dann entscheiden, ob eine nichttriviale Lösung existiert oder nicht.

Im Spezialfall $k = n$ gilt sogar: Existiert in einer dieser Normalformen eine Nullzeile, so liegt lineare Abhängigkeit vor (man hat dann mehr als eine Lösung, weil dann $r < k = n$ gilt). Andernfalls, wenn es keine Nullzeile gibt (also $r = n = k$), so existiert eine eindeutige Lösung, nämlich $x = 0$. In diesem Fall haben wir dass $\lambda_1 = \ldots = \lambda_k = 0$ und die Vektoren sind linear unabhängig.

Definition Sei V ein K-Vektorraum über einem Körper K. Wir nennen ein linear unabhängiges Erzeugendensystem von V eine **Basis** von K oder eine **K-Basis**. Genauer heißt eine Familie $\mathcal{B} = (v_i)_{i \in I}$ eine Basis von K, wenn

(i) $V = \operatorname{Spann}_K \mathcal{B}$ und

(ii) $\mathcal{B}$ ist linear unabhängig über K.

Gilt sogar $I = \{1, \ldots, n\}$, so heißt $\mathcal{B} = (v_1, \ldots, v_n)$ eine **endliche Basis** von V, und man sagt, V habe eine **endliche Basis der Länge** n, d.h. $\text{Länge}(\mathcal{B}) := n$.

Wir erhalten aus Satz 3.14 folgende Korollare. Das zweite Korollar wird für den endlichen Fall formuliert. Zudem werden wir sehen, dass die Länge einer Basis eindeutig ist.

Korollar 3.15. *Sei V ein K-Vektorraum über einem Körper K und sei $\mathcal{B} = (v_i)_{i \in I}$ eine Familie von Vektoren aus V. Dann sind die folgenden Aussagen äquivalent:*

(a) *$\mathcal{B}$ ist eine Basis von V.*

(b) *Jeder Vektor $v \in V$ lässt sich in eindeutiger Weise als Linearkombination aus B darstellen. D.h. für alle $v \in V$ existieren eindeutige $\lambda_i \in K$ $(i \in I)$, so dass*

$$v = \sum_{i \in I} \lambda_i v_i \qquad (\lambda_i = 0 \text{ für fast alle } i \in I) .$$

Korollar 3.16. (Koeffizientenvergleich) *Sei V ein K-Vektorraum über einem Körper K und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Familie von Vektoren aus V. Dann sind die folgenden Aussagen äquivalent:*

(a) *$\mathcal{B}$ ist eine Basis von V.*

(b) *Jeder Vektor $v \in V$ lässt sich in eindeutiger Weise als Linearkombination aus B darstellen. D.h. jeder Vektor $v \in V$ lässt sich als Linearkombination der Vektoren $v_1, \ldots, v_n$ darstellen und falls $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in K$ existieren, so dass*

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n = \mu_1 v_1 + \ldots \mu_n v_n \ ,$$

so gilt $\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n$.

3.3 Zentrale Aussagen der Vektorraumtheorie und Dimension

Sei V ein K-Vektorraum über einem Körper K. Dann tauchen die folgenden Fragen auf:

(a) Existiert immer eine Basis von V?

(b) Wenn eine Basis existiert, ist diese immer endlich und ist sie eindeutig?

(c) Wenn eine endliche Basis existiert, existiert dann für jeden Untervektorraum von V auch eine endliche Basis?

Der folgende Satz bildet die Grundlage für die Antworten auf diese Fragen. Oft wird Teil (b) dieses Satzes in der Literatur als Basisauswahlsatz und Teil (c) auch als Basisergänzungssatz bezeichnet. Wir konzentrieren uns hierbei zuerst auf den endlichen Fall.

Satz 3.17. *Sei V ein K-Vektorraum über einem Körper K und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Familie von Vektoren aus V. Dann sind die folgenden Aussagen äquivalent:*

(a) *$\mathcal{B}$ ist eine Basis von V.*

(b) *$\mathcal{B}$ ist minimales Erzeugendensystem von V. D.h. $V = \mathrm{Spann}_K \mathcal{B}$, und für jedes $j \in \{1, \ldots, n\}$ gilt*

$$V \neq \mathrm{Spann}_K (v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n) .$$

(c) *$\mathcal{B}$ ist ein maximal linear unabhängiges System von Vektoren aus V. D.h. $\mathcal{B}$ ist linear unabhängig, und für jedes $v \in V$ ist $(v_1, \ldots, v_n, v)$ linear abhängig.*

Satz 3.18. *Sei V ein K-Vektorraum über einem Körper K und sei $\mathcal{B} = (v_i)_{i \in I}$ eine Familie von Vektoren aus V. Dann sind die folgenden Aussagen äquivalent:*

(a) *$\mathcal{B}$ ist eine Basis von V.*

(b) *$\mathcal{B}$ ist minimales Erzeugendensystem von V. D.h. $V = \mathrm{Spann}_K \mathcal{B}$, und für jedes $j \in I$ gilt*

$$V \neq \mathrm{Spann}_K (v_i)_{\substack{i \in I \\ i \neq j}} .$$

(c) *$\mathcal{B}$ ist ein maximal linear unabhängiges System von Vektoren aus V. D.h. $\mathcal{B}$ ist linear unabhängig, und für jedes $v \in V$ ist $(v, (v_i)_{i \in I})$ linear abhängig.*

Der obige zentrale Satz spiegelt die Min-Max-Situation einer Basis fundamental wider. Deshalb müssen die direkten Konsequenzen genauer beleuchten.

Satz 3.19. (Basisauswahlsatz) *Sei V ein endlich erzeugter K-Vektorraum über einem Körper K. Dann lässt sich aus jedem endlichen Erzeugendensystem von V eine Basis von V über K auswählen.*

Damit ist die Existenz einer Basis für endlich erzeugte K-Vektorräume gesichert.

Satz 3.20. *Jeder Vektorraum über einem Körper K besitzt eine Basis. Insbesondere besitzt jeder endlich erzeugte Vektorraum über einem Körper K eine Basis aus endlich vielen Vektoren.*

Satz 3.21. *Sei V ein K-Vektorraum über einem Körper K. Ist $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V und $(w_1, \ldots, w_n)$ ein linear unabhängiges System von V, so gilt immer $m \geq n$.*

Satz 3.22. *Sei V ein endlich erzeugter K-Vektorraum über einem Körper K. Dann haben je zwei Basen von V die gleiche (endliche) Länge.*

Die Länge einer jeden Basis eines endlich erzeugten Vektorraumes ist einer der fundamentalen Invarianten der Vektorraumtheorie und hat deshalb natürlich einen Namen.

Definition Sei V ein K-Vektorraum über einem Körper K. Dann definieren wir die **Dimension von** V als

$$\dim_K V := \begin{cases} \infty & \text{, falls } V \text{ nicht endlich erzeugt ist,} \\ n & \text{, falls } V \text{ endlich erzeugt ist und eine Basis der Länge } n \text{ besitzt.} \end{cases}$$

Ein Vektorraum V mit $\dim_K V < \infty$ heißt **endlich-dimensional**.

Satz 3.23. *Sei V ein endlich erzeugter K-Vektorraum über einem Körper K der Dimension $\dim_K V = n$. Dann gilt:*

(a) *Jede Familie mit $m < n$ Vektoren ist kein Erzeugendensystem von V.*

(b) *Jede Familie mit $m > n$ Vektoren ist linear abhängig.*

(c) *Jedes Erzeugendensystem von n Vektoren aus V ist eine Basis von V.*

(d) *Jede linear unabhängige Familie von n Vektoren aus V ist eine Basis von V.*

Satz 3.24. (Basisergänzungssatz, endlich erzeugter Fall) *Sei V ein endlich-erzeugter K-Vektorraum über einem Körper K mit $\dim_K V = n$. Dann lässt sich jede linear unabhängige Familie von Vektoren aus V zu einer Basis von V ergänzen. Genauer gilt: Sei $(w_1, \ldots, w_m)$ eine Familie linear unabhängiger Vektoren aus V und sei $\mathcal{C}$ ein beliebiges Erzeugendensystem von V. Dann lässt sich aus $\mathcal{C}$ eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ auswählen, und es existieren paarweise verschiedene Indizes $i_{m+1}, \ldots, i_n \in \{1, \ldots, n\}$, so dass*

$$\left(w_1, \ldots, w_m, v_{i_{m+1}}, \ldots, v_{i_n}\right)$$

eine Basis von V ist.

Satz 3.25. (Basisergänzungssatz, allgemeiner Fall) *Sei V ein K-Vektorraum über einem Körper K und sei $(w_j)_{j \in J}$ eine Familie linear unabhängiger Vektoren aus V. Dann existiert eine Basis $\mathcal{B} = (v_i)_{i \in I}$ von V, die die Familie $(w_j)_{j \in J}$ umfasst.*

Wir sind auch in der Lage, die letzte noch offene Frage zu klären.

Satz 3.26. *Sei V ein endlich erzeugter K-Vektorraum der Dimension $\dim_K V = n$ über einem Körper K und sei $U \subseteq V$ ein Untervektorraum von V. Dann gilt:*

(a) *U ist auch endlich-dimensional und erfüllt*

$$\dim_K U \leq \dim_K V .$$

Insbesondere ist jeder Untervektorraum eines endlich erzeugten Vektorraums auch endlich erzeugt.

(b) $\dim_K U = \dim_K V \Leftrightarrow U = V$.

In der Literatur findet man zudem noch das **Austauschlemma** und den **Steinitzschen Austauschsatz**, aus dem man auch den Basisauswahlsatz, die Existenz einer Basis und den Basisergänzungssatz ableiten kann.

Satz 3.27. (Austauschlemma) *Sei V ein K-Vektorraum der Dimension* $\dim_K V = n$ *über einem Körper K und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Für jedes $w \in V$ mit $w \neq 0$ existieren $\lambda_1, \ldots, \lambda_n \in K$, die nicht alle gleich* 0 *sind, so dass*

$$w = \sum_{i=1}^{n} \lambda_i v_i = \lambda_1 v_1 + \ldots + \lambda_j v_j + \ldots + \lambda_n v_n .$$

Ist $\lambda_j \neq 0$ für ein $j \in \{1, \ldots, n\}$, so ist v_j in $\mathcal{B}$ gegen w austauschbar, d.h. dass

$$\mathcal{B}' = (v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n)$$

auch eine Basis von V.

Satz 3.28. (Steinitzscher Austauschsatz) *Sei V ein endlich erzeugter K-Vektorraum der Dimension* $\dim_K V = n$ *über einem Körper K und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Ist $(w_1, \ldots, w_m)$ eine beliebige linear unabhängige Familie von Vektoren aus V, so muss $m \leq n$ sein, und es gilt: Nach Umnummerierung ist*

$$(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$$

eine Basis von V. Genauer existieren Indizes $1 \leq i_1 < i_2 < \ldots < i_m \leq n$, so dass nach Austausch $v_{i_1} \mapsto w_1$, $v_{i_2} \mapsto w_1, \ldots, v_{i_m} \mapsto w_m$ die Familie

$$(v_1, \ldots, v_{i_1-1}, w_1, v_{i_1+1}, \ldots, v_{i_2-1}, w_2, v_{i_2+1}, \ldots, v_{i_m-1}, w_m, v_{i_m+1}, \ldots, v_n)$$

auch eine Basis von V ist.

4 Der Rang einer Matrix

Wir werden nun die Struktur von Matrizen mithilfe der neuen Begriffe Untervektorraum, Basis und Dimension untersuchen. Dabei können eine Reihe von rechnerischen Problemen für Untervektorräume gelöst werden.

Problem 1 Sei V ein K-Vektorraum über einem Körper K und seien $v_1, \ldots, v_m \in V$. Berechne eine Basis des von diesen Vektoren aufgespannten Untervektorraumes, d.h. berechne eine Basis von $U = \mathrm{Spann}_K(v_1, \ldots, v_m)$.

Problem 2 Sei V ein K-Vektorraum über einem Körper K und seien $v_1, \ldots, v_m \in V$. Bestimme die Dimension

$$\dim_K \mathrm{Spann}_K(v_1, \ldots, v_m) \ .$$

Wir werden sehen, dass dies in K^n äquivalent zu folgendem Problem ist: Bestimme den Rang einer Matrix, d.h. berechne den Zeilenrang oder Spaltenrang einer Matrix.

Problem 3 Sei V ein K-Vektorraum über einem Körper K und seien $m+p$ Vektoren $v_1, \ldots, v_m, w_1, \ldots, w_p \in V$ gegeben. Berechne eine Basis von

$$\mathrm{Spann}_K(v_1, \ldots, v_m) \cap \mathrm{Spann}_K(w_1, \ldots, w_p) \ .$$

Definition Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für eine Matrix $A \in K^{m\times n}$ heißt der Lösungsraum $\mathrm{NR}(A)$ des homogenen Gleichungssystems $Ax = 0$ der **Nullraum von** A bzw. der **Kern von** A:

$$\mathrm{NR}(A) = \mathrm{Lös}(A, 0) = \left\{ x \in K^{n\times 1} \mid Ax = 0 \right\} \subseteq K^{n\times 1} \ .$$

Dessen Dimension $\nu(A) := \dim_K \mathrm{NR}(A)$ nennt man die **Nullität von** A.

Zu beachten ist, dass der Nullraum in unserer Notation nicht der Name für den Untervektorraum $\{0\}$ ist. In der Literatur wird dieser manchmal auch mit Nullraum betitelt.

Wie man den Nullraum berechnet, haben wir bereits früher gesehen. Man wendet Gaußschen Algorithmus bzw. Gauß-Jordan Algorithmus an. Wir hatten den Raum nur nicht explizit benannt. Die linear unabhängigen Vektoren w_j, $1 \leq j \leq n-r$, aus Satz 2.28 bilden dann eine Basis des Nullraums und, mit demselben r, gilt dann $\nu(A) = n - r$. Dieses r ist also ganz wichtig, und wir werden dem r auch einen Namen geben. Offensichtlich wird auch Gaußscher Algorithmus bzw. Gauß-Jordan Algorithmus eine wichtige Rolle spielen. Die zentrale Frage wird jedoch sein, was mit den Spalten einer Matrix bei elementaren Zeilenumformungen geschieht.

Wir betrachten nun Zeilen und Spalten einer Matrix separat und kommen zu einigen erstaunlichen Resultaten. Sei $A = (a_{ij}) \in K^{m\times n}$. Sind $z_1, \ldots, z_m \in K^{1\times n}$ die Zeilen von A und $s_1, \ldots, s_n \in K^{m\times 1}$ die Spalten von A, so haben wir

$$A = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} s_1 & s_2 & \ldots & s_n \end{pmatrix} \in K^{m\times n} \,,$$

wobei

$$z_i := \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix} \in K^{1\times n} \quad , \quad s_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^{m\times 1} \,.$$

Definition Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für eine Matrix $A \in K^{m\times n}$ definieren wir mit der obigen Notation

(a) den **Zeilenraum von** A als $\mathrm{ZR}(A) := \mathrm{Spann}_K(z_1, \ldots, z_m) \subseteq K^{1\times n}$.

(b) den **Zeilenrang von** A als $\mathrm{zr}(A) := \dim_K \mathrm{ZR}(A) \le n < \infty$.

(c) den **Spaltenraum von** A als $\mathrm{SR}(A) := \mathrm{Spann}_K(s_1, \ldots, s_n) \subseteq K^{m\times 1}$.

(d) den **Spaltenrang von** A als $\mathrm{sr}(A) := \dim_K \mathrm{SR}(A) \le m < \infty$.

Zu beachten ist, dass der Zeilenraum bzw. Spaltenraum einer Matrix tatsächlich ein Untervektorraum von $K^{1\times n}$ bzw. von $K^{m\times 1}$ ist. Allerdings liegen Zeilenraum und Spaltenraum in **verschiedenen Vektorräumen**.

Wir wenden uns dem ersten erwähnten Problem zu und erklären nun, wie man dieses Problem lösen kann. Dazu beschränken wir uns auf den Zeilenraum und untersuchen den Effekt von elementaren Zeilenumformungen auf den Zeilenraum einer Matrix.

Satz 4.1. *Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für eine Matrix $A \in K^{m \times n}$ sind Zeilenraum und Zeilenrang invariant unter den elementaren Zeilenumformungen. D.h. $\mathrm{ZR}(A)$ und $\mathrm{zr}(A)$ verändern sich nicht unter Anwendung von elementaren Zeilenumformungen. Insbesondere erhalten wir die folgenden Aussagen:*

(a) *Geht $B \in K^{m \times n}$ aus A durch Anwendung einer endlichen Folge von elementaren Zeilenumformungen hervor, so gilt:*
$$\mathrm{ZR}(B) = \mathrm{ZR}(A) \quad \textit{und} \quad \mathrm{zr}(B) = \mathrm{zr}(A) \ .$$

(b) *Für alle $Q \in \mathrm{GL}(m, K)$ gilt:* $\mathrm{zr}(QA) = \mathrm{zr}(A)$.

(c) *Ist $\tilde{A}$ eine Gaußsche Normalform von A und A^* die reduzierte Gaußsche Normalform von A, so gilt:*
$$\mathrm{ZR}(\tilde{A}) = \mathrm{ZR}(A^*) = \mathrm{ZR}(A) \quad \textit{und} \quad \mathrm{zr}(\tilde{A}) = \mathrm{zr}(A^*) = \mathrm{zr}(A) \ .$$

Satz 4.2. *Sei $A = (a_{ij}) \in K^{m \times n}$ eine Matrix mit m Zeilen in Zeilenstufenform (z.B. in Gaußscher Normalform oder in reduzierter Gaußscher Normalform) mit r Stufen, wobei $r \in \{0, \ldots, m\}$, und mit Stufenindizes $j_1, j_2, \ldots, j_r$, wobei*
$$1 \leq j_1 < j_2 < \ldots < j_r \leq n \quad \textit{und} \quad r \leq \min(n, m) \ .$$
Dann gilt
$$\mathrm{zr}(A) = r \ ,$$
und das System der ersten r Zeilen von A bildet eine Basis von $\mathrm{ZR}(A)$.

Lösung von Problem 1 Wir betrachten nur den K-Vektorraum $V = K^{n\times 1}$ über einem Körper K. Der Allgemeinfall kann mit dem Koordinatenisomorphismus davon abgeleitet werden. Seien $v_1, \ldots, v_m \in K^{n\times 1}$. Das Ziel ist, eine Basis des von diesen Vektoren aufgespannten Untervektorraumes zu berechnen, d.h. wir bestimmen eine Basis des Untervektorraumes $U = \mathrm{Spann}_K(v_1, \ldots, v_m)$.

(1) Schreibe die Vektoren als Zeilen $v_1^T, \ldots, v_m^T \in K^{1\times n}$ in eine Matrix

$$A := \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_m^T \end{pmatrix} \in K^{m\times n} .$$

(2) Überführe A durch elementare Zeilenumformungen in Zeilenstufenform, z.B. führe Gaußschen bzw. Gauß-Jordan Algorithmus aus, um A in ein Gaußsche Normalform $\tilde{A}$ bzw. in reduzierte Gaußsche Normalform A^* zu bringen.

(3) Die erhaltene Zeilenstufenform besitzt r Stufen, $r \in \{0, \ldots, m\}$, und r Stufenindizes $j_1, j_2, \ldots, j_r$. Die ersten r Zeilen von $\tilde{A}$ bzw. A^* bilden dann nach Satz 4.2 eine Basis $\mathcal{B}'$ von $\mathrm{ZR}(A) \subseteq K^{1\times n}$.

(4) Die Transponierten dieser Basisvektoren aus $\mathcal{B}'$ bilden demnach eine Basis $\mathcal{B}$ von $U = \mathrm{Spann}_K(v_1, \ldots, v_m)$. Hat man A in eine Gaußsche Normalform $\tilde{A}$ bzw. in die reduzierte Gaußsche Normalform A^* gebracht, so ist für $A \neq 0$

$$\tilde{A} = \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \tilde{z}_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{bzw.} \quad A^* = \begin{pmatrix} z_1^* \\ z_2^* \\ \vdots \\ z_r^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Dann bilden die folgenden Familien Basen von $U = \mathrm{Spann}_K(v_1, \ldots, v_m)$:

$$\Big((\tilde{z}_1)^T, (\tilde{z}_2)^T, \ldots, (\tilde{z}_r)^T \Big) \quad \text{bzw.} \quad \Big((z_1^*)^T, (z_2^*)^T, \ldots, (z_r^*)^T \Big) .$$

VORSICHT: Bei elementaren Zeilenumformungen ist der Spaltenraum **nicht** invariant. Wir werden jedoch herleiten, dass der Spaltenrang invariant bleibt.

Bemerkung 4.3. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und sei $A \in K^{m\times n}$.*

(a) ***Warnung:*** *Bei elementaren Zeilenumformungen ist der Spaltenraum **nicht** invariant. Hier ist also Vorsicht angebracht. Wir werden jedoch sehen, dass der Spaltenrang invariant ist.*

(b) *Analog zu Satz 4.1 kann man zeigen, dass Spaltenraum und Spaltenrang invariant unter den elementaren **Spalten**umformungen sind. Zum Beispiel kann man dazu den Zeilenraum und den Zeilenrang der Transponierten $A^T \in K^{n\times m}$ betrachten. Für alle $P \in \mathrm{GL}(n, K)$ erhält man dann auch, dass $\mathrm{sr}(AP) = \mathrm{sr}(A)$.*

(c) *Ist A eine beliebige Matrix, so gilt*

$$\mathrm{zr}(A^T) = \mathrm{sr}(A) \qquad \textit{und} \qquad \mathrm{sr}(A^T) = \mathrm{zr}(A)\ .$$

In manchen Büchern identifiziert man sogar K^n mit $K^{n\times 1}$ und $K^{1\times n}$. Dann identifiziert man auch die entsprechenden Unterräume von K^n wie folgt:

$$\mathrm{ZR}(A^T) = \mathrm{SR}(A) \qquad \textit{und} \qquad \mathrm{SR}(A^T) = \mathrm{ZR}(A)\ .$$

Wir werden jedoch Spalten und Zeilen explizit unterscheiden.

Als zentrales und im ersten Moment überraschendes Resultat werden wir nun herleiten, dass Spaltenrang und Zeilenrang übereinstimmen. Nach den vorherigen Bemerkungen ist aber klar, dass Spaltenraum und Zeilenraum nicht gleich sind! Dazu benötigen wir folgendes Lemma.

Lemma 4.4. (Spaltenrang-Lemma) *Sei K ein Körper und seien $m, n \in \mathbb{N}$. Geht die Matrix B aus $A \in K^{m\times n}$ durch Anwendung einer endlichen Folge elementarer Zeilenumformungen hervor, so gilt:*

(a) *Die Spalten von A sind genau dann linear unabhängig, wenn die Spalten von B linear unabhängig sind.*

(b) *Der Spaltenrang von A bleibt unter elementaren Zeilenumformungen invariant, d.h.*

$$\mathrm{sr}(B) = \mathrm{sr}(A)\ .$$

Satz 4.5. *Sei K ein Körper und seien $m, n \in \mathbb{N}$. Sei $A \in K^{m \times n}$ eine Matrix und $A^* \in K^{m \times n}$ die reduzierte Gaußsche Normalform von A mit r Stufen, $r \in \{0, \ldots, m\}$, und mit Stufenindizes $j_1, j_2, \ldots, j_r$, wobei $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ und mit $r \leq \min(n, m)$ Dann gilt:*

(a) *Die Spalten $\left(s^*_{j_1}, s^*_{j_2}, \ldots, s^*_{j_r}\right)$ von A^* formen eine Basis von* $\mathrm{SR}(A^*)$.

(b) *Die Spalten $(s_{j_1}, s_{j_2}, \ldots, s_{j_r})$ von A formen eine Basis von* $\mathrm{SR}(A)$.

(c) **(Spaltenrang = Zeilenrang)**

$$\mathrm{sr}(A) = \mathrm{zr}(A) = r \ .$$

Definition Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für eine Matrix $A \in K^{m \times n}$ mit $A \neq 0$ definieren wir den **Rang von** A als

$$\mathrm{rang}(A) := \mathrm{sr}(A) = \mathrm{zr}(A) \ .$$

Außerdem setzen wir $\mathrm{rang}(0) := 0$.

Satz 4.6. *Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für eine Matrix $A \in K^{m \times n}$ und $b \in K^{m \times 1}$ sind die folgenden Aussagen äquivalent.*

(a) *Das lineare Gleichungssystem $A x = b$ ist lösbar.*

(b) $b \in \mathrm{Spann}_K(s_1, \ldots, s_n)$, *wobei $s_1, \ldots, s_n$ die Spalten von A sind.*

(c) $\mathrm{rang}(A \mid b) = \mathrm{rang}(A)$, *wobei $(A \mid b)$ die erweiterte Koeffizientenmatrix des linearen Gleichungsystems $A x = b$ ist.*

Korollar 4.7. *Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für $A \in K^{m \times n}$ gilt:*

(a) *Geht $B \in K^{m \times n}$ aus A durch Anwendung einer endlichen Folge von elementaren Zeilenumformungen hervor, so gilt:* $\mathrm{rang}(B) = \mathrm{rang}(A)$.

(b) $\mathrm{rang}(A^T) = \mathrm{rang}(A)$.

(c) *Für alle $Q \in \mathrm{GL}(m, K)$ gilt:* $\mathrm{rang}(Q A) = \mathrm{rang}(A)$.

(d) *Für alle $P \in \mathrm{GL}(n, K)$ gilt:* $\mathrm{rang}(A P) = \mathrm{rang}(A)$.

(e) **Ranggleichung:**

$$\mathrm{rang}(A) + \nu(A) = n \ .$$

Wir können nun unseren Hauptsatz über invertierbare Matrizen erweitern, d.h. für den Spezialfall von quadratischen Matrizen erhalten wir den folgenden Satz.

Satz 4.8. *Sei $A \in K^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:*

(a) *A ist invertierbar, also $A \in \mathrm{GL}(n, K)$.*

(b) *Ist A^* eine reduzierte Gaußsche Normalform von A mit r Stufen, so gilt $r = n$ und damit $A^* = E_n$.*

(c) *Das lineare Gleichungssystem $Ax = 0$ hat nur die triviale Lösung $x_0 := 0$.*

(d) *Das lineare Gleichungssystem $Ax = b$ hat eine eindeutige Lösung, nämlich die Lösung $x_0 := A^{-1} b$.*

(e) *A ist ein Produkt von Elementarmatrizen.*

(f) $\mathrm{rang}(A) = n$.

(g) $\nu(A) = 0$.

(h) *A besitzt genau n linear unabhängige Spalten.*

(i) *A besitzt genau n linear unabhängige Zeilen.*

5 Affine Unterräume (Optional)

Affine Unterräume bilden den Einstieg zur analytischen Geometrie. Dort werden diese immer wieder verwendet. An mehreren Beispielen kann man sich davon überzeugen, dass diese Unterräume im Prinzip die **richtigen** Räume sind. Beachte, dass affine Unterräume im Allgemeinen keine Untervektorräume sind. Wir werden sehen, dass man die Begriffe Dimension und Untervektorraum auf Teilmengen eines Vektorraumes ausdehnen kann, welche selbst keine Vektorräume sind. Solche affinen Unterräume entstehen durch punktweise Verschiebung von Untervektorräume.

Dieses Kapitel ist in der Vorlesung **Lineare Algebra** optional. Die meisten Aussagen werden für Zwei-Fächer-Bachelor in der Vorlesung **Geometrie** behandelt und die zugehörige Theorie wird teilweise in der **Algebra I** behandelt.

Definition Sei V ein K-Vektorraum über einem Körper K. Eine Teilmenge $X \subseteq V$ heißt ein **affiner Unterraum von** V, wenn es ein $v \in V$ und einen Untervektorraum U von V gibt, so dass

$$X = v + U := \{ v + u \mid u \in U \} \ .$$

X heißt auch **Nebenklasse von** U **durch** v und man schreibt auch $X = [v]$ bzw. $X = \overline{v}$. Man nennt U den **Richtungsraum von** X, und man definiert

$$\dim_K X := \dim_K U \ .$$

Bemerkung 5.1. *In obiger Notation gilt:*

(a) *Man muss den Term Unterraum hier mit Vorsicht verwenden. Im Allgemeinen ist ein affiner Unterraum nämlich kein Untervektorraum. Diese Begriffe bezeichnen unterschiedliche Objekte.*

(b) *Für einen affinen Unterraum $X = v + U$ ist U immer eindeutig bestimmt durch*

$$U = \{ x - y \mid x, y \in X \} \ .$$

Wir stellen nun einige wichtige Eigenschaften von affinen Unterräumen zusammen. Sei V ein K-Vektorraum über einem Körper K.

Seien v, $v' \in V$ und U, U' Untervektorräume von V.

(1) Für alle $x \in v + U$ gilt:
$$v + U = x + U \ .$$

(2) $v + U$ ist genau dann ein Untervektorraum von V, wenn $v \in U$ liegt.
In diesem Fall ist $v + U = U$.

(3) $v + U = v' + U \ \Leftrightarrow \ v - v' \in U$.

(4) Entweder ist $v + U = v' + U$ oder $(v + U) \cap (v' + U) = \varnothing$.

(5) Zu festem Untervektorraum U sind verschiedene affine Unterräume disjunkt. Nur der affine Unterraum $0 + U = U$ enthält die 0, kann also einen Untervektorraum von V sein. Die affinen Unterräume bilden folglich eine Partition von V.

(6) Die Relation
$$x \sim y \quad :\Leftrightarrow \quad x - y \in U$$
stellt eine Äquivalenzrelation auf V dar.

(7) $v + U = v' + U' \ \Rightarrow \ U = U'$ und $v - v' \in U = U'$.

(8) Entweder ist $(v + U) \cap (v' + U') = \varnothing$ oder
$$(v + U) \cap (v' + U') = w + (U \cap U') \quad \text{mit} \quad w \in (v + U) \cap (v' + U') \ .$$
Im letzteren Fall ist also $(v + U) \cap (v' + U')$ wieder ein affiner Unterraum mit neuem Richtungsraum $U \cap U'$. Außerdem ist genau dann $(v + U) \cap (v' + U') \neq \varnothing$, wenn $v - v' \in U + U'$.

Aus dem letzten Punkt leiten wir nun einen Lösungsweg für das folgende Problem 4 ab.

Problem 4 Seien v, v', $v_1, \ldots, v_m, w_1, \ldots, w_p \in K^{n \times 1}$ gegeben. Bestimme, ob
$$(v + \mathrm{Spann}_K(v_1, \ldots, v_m)) \cap (v' + \mathrm{Spann}_K(w_1, \ldots, w_p)) = \varnothing \ .$$

Lösung: Diese Frage ist äquivalent zur Frage, ob $\text{Lös}(A, v - v') = \varnothing$, wobei
$$A := \begin{pmatrix} v_1 & v_2 & \ldots & v_m & w_1 & w_2 & \ldots & w_p \end{pmatrix} \in K^{n \times (m+p)} \ .$$
Wende Gaußschen Algorithmus bzw. Gauß-Jordan Algorithmus auf die erweiterte Koeffizientenmatrix $(A \mid v - v')$ an, um A in Gaußsche Normalform bzw. reduzierte Gaußsche zu bringen. Das Lösbarkeitskriterium in Satz 2.24 liefert nun die Antwort.

6 Lineare Abbildungen

Wir werden sehen, dass Matrizen und lineare Abbildungen fest miteinander verknüpft sind. Lineare Abbildungen bilden eines der zentralen Themen der Linearen Algebra.

6.1 Definition und grundlegende Eigenschaften

Zur Wiederholung erinnern wir an die Definition von Abbildungen aus dem 1. Kapitel in Notation von Vektorräumen. Seien V, W Mengen. Eine **Abbildung** oder **Funktion** von V nach W ist eine Vorschrift

$$\begin{array}{rccl} F: & V & \to & W \\ & v & \mapsto & F(v) = w \ , \end{array}$$

derart dass jedem $v \in V$ genau ein $w \in W$ zugeordnet wird.

$$(\forall v \in V)(\exists !\, w \in W)\ [F(v) = w]\ .$$

Wir schreiben $\mathrm{Abb}(V, W)$ oder W^V für die Menge aller Abbildungen von V nach W.

Definition Sei K ein Körper und seien V, W zwei K-Vektorräume. Dann heißt eine Abbildung $F : V \to W$ eine **(K-)lineare Abbildung** bzw. **(K-)Vektorraum-homomorphismus**, wenn die folgenden Bedingungen erfüllt sind:

(i) $F(v + v') = F(v) + F(v')$ für alle $v, v' \in V$.

(ii) $F(\lambda v) = \lambda F(v)$ für alle $v \in V$ und $\lambda \in K$.

Die Idee ist, dass F strukturerhaltend ist.

Als wichtiges Beispiel erwähnen wir die **Standardinterpretation** einer Matrix. Sei dazu $A \in K^{m\times n}$, wobei K ein Körper ist und seien $m, n \in \mathbb{N}$. Dann ist die Abbildung

$$\begin{array}{rccl} F_A: & K^{n\times 1} & \to & K^{m\times 1} \\ & x & \mapsto & F_A(x) := A\,x \end{array}$$

eine lineare Abbildung. Aus jeder Matrix kann man also eine lineare Abbildung erzeugen. Die Umkehrung davon wird sich auch als richtig herausstellen.

Proposition 6.1. *Sei K ein Körper und seien V, W zwei K-Vektorräume. Eine Abbildung $F : V \to W$ ist genau dann eine lineare Abbildung, wenn die folgende Bedingung erfüllt ist:*

(i') $F(\lambda v + \lambda' v') = \lambda F(v) + \lambda' F(v')$ *für alle $v, v' \in V$ und $\lambda, \lambda' \in K$.*

Proposition 6.2. (Elementare Eigenschaften von linearen Abbildungen) *Sei K ein Körper und seien V, W zwei K-Vektorräume. Dann gelten für eine lineare Abbildung $F : V \to W$ die folgenden Eigenschaften:*

(a) $F(0_V) = 0_W$.

(b) $F(-v) = -F(v)$ *für alle $v \in V$.*

6.2 Kern und Bild von linearen Abbildungen

Wir werden jetzt lineare Abbildungen zerlegen und die wichtige Dimensionsformel beweisen, die wir im Prinzip in einem Spezialfall bereits bewiesen haben.

Definition Seien V, W zwei K-Vektorräume über einem Körper K und sei $F : V \to W$ eine lineare Abbildung. Wir bezeichnen mit

$$\mathrm{Kern}(F) := \{ v \in V \mid F(v) = 0_W \} = F^{-1}(\{0_W\})$$

den **Kern** von F und mit

$$\mathrm{Bild}(F) = \mathrm{Im}(F) = \{ w \in W \mid \text{Es gibt ein } v \in V \text{ mit } F(v) = w \} = \{ F(v) \mid v \in V \}$$

das **Bild** von F.

Satz 6.3. *Seien V, W zwei K-Vektorräume über einem Körper K und sei $F : V \to W$ eine lineare Abbildung. Dann gilt:*

(a) $\mathrm{Kern}(F)$ *ist ein Untervektorraum von V.*

(b) $\mathrm{Bild}(F)$ *ist ein Untervektorraum von W.*

Bemerkung 6.4. *Seien V, W zwei K-Vektorräume über einem Körper K Ist die Abbildung $F : V \to W$ linear, dann gilt noch allgemeiner:*

(a) *Ist $U' \subseteq W$ ein Untervektorraum von W, so ist*

$$F^{-1}(U') := \{ v \in V \mid F(v) \in U' \}$$

ein Untervektorraum von V. In dieser Notation bezeichnet $\mathrm{Kern}(F)$ *die* ***Faser*** $F^{-1}(\{0_W\})$, *also $U' = \{0_W\}$.*

(b) *Ist $U \subseteq V$ ein Untervektorraum von V, so ist*

$$F(U) := \{ F(u) \mid u \in U \}$$

ein Untervektorraum von W. In dieser Notation bezeichnet $\mathrm{Bild}(F) = F(V)$, *also $U = V$.*

Lemma 6.5. *Seien V und W zwei K-Vektorräume über einem Körper K und sei die Abbildung $F : V \to W$ eine lineare Abbildung. Ist $(v_1, \ldots, v_k)$ ein Erzeugendensystem von V, so ist $(F(v_1), \ldots, F(v_k))$ ein Erzeugendensystem von* $\mathrm{Bild}(F)$.

Außerdem haben wir in Kapitel 1 auch definiert, wann Abbildungen injektiv, surjektiv oder bijektiv heißen. Seien V, W Mengen. Eine Abbildung $F : V \to W$ ist genau dann **injektiv**, wenn zu jedem Element $w \in W$ höchstens ein Element $v \in V$ existiert, derart dass $F(v) = w$. In dem Fall nennt man f eine **Injektion**. In Kurzform: $F : V \to W$ ist genau dann injektiv, wenn

$$\forall v, v' \in V \ : \ (F(v) = F(v') \ \Rightarrow \ v = v') \ .$$

Dies ist logisch äquivalent zu

$$\forall v, v' \in V \ : \ (v \neq v' \ \Rightarrow \ F(v) \neq F(v')) \ .$$

Die Abbildung $F : V \to W$ ist genau dann **surjektiv**, wenn es zu jedem $w \in W$ ein Element $v \in V$ gibt, derart dass $F(v) = w$. In Kurzform: $F : V \to W$ ist genau dann surjektiv, wenn

$$\forall w \in W \ \ \exists v \in V \ : \ F(v) = w \ .$$

Eine Abbildung $F : V \to W$ ist genau dann **bijektiv**, wenn F injektiv und surjektiv ist. D.h. zu jedem $w \in W$ gibt es genau ein $v \in V$, so dass $F(v) = w$. In Kurzform

$$\forall w \in W \ \ \exists! v \in V \ : \ F(v) = w \ .$$

Definition Seien V, W zwei K-Vektorräume über einem Körper K. Eine lineare Abbildung $F : V \to W$ heißt

(a) ein **(Vektorraum-)Epimorphismus**, wenn F zudem surjektiv ist.

(b) ein **(Vektorraum-)Monomorphismus**, wenn F zudem injektiv ist.

(c) ein **(Vektorraum-)Isomorphismus**, wenn F zudem bijektiv ist, also injektiv und surjektiv. In dem Fall nennen wir V und W **isomorph** und schreiben $V \cong W$.

(d) ein **(Vektorraum-)Endomorphismus**, wenn $V = W$.

(e) ein **(Vektorraum-)Automorphismus**, wenn $V = W$ und F bijektiv ist.

Satz 6.6. *Seien V, W zwei K-Vektorräume über einem Körper K und sei $F : V \to W$ eine lineare Abbildung. Dann gilt:*

(a) *F ist genau dann ein Monomorphismus, wenn* $\mathrm{Kern}(F) = \{0\}$.

(b) *F ist genau dann ein Epimorphismus, wenn* $\mathrm{Bild}(F) = W$.

(c) *F ist genau dann ein Isomorphismus, wenn* $\mathrm{Kern}(F) = \{0\}$ *und* $\mathrm{Bild}(F) = W$.

Zur Erinnerung: Seien U, V, W Mengen. Wenn $F : U \to V$ und $G : V \to W$ zwei kompatible Abbildungen sind, so ist die Komposition $G \circ F$ von G mit F die Abbildung

$$\begin{aligned} G \circ F : \ U &\to W \\ u &\mapsto G \circ F(u) = G(F(u)) \ . \end{aligned}$$

Bezeichnet id_U die **Identität** von U und id_V die **Identität** von V, so gelten:

- $F \circ \mathrm{id}_U = F$ und $\mathrm{id}_V \circ F = F$.
- F ist injektiv $\Leftrightarrow$ $\exists\ H : V \to U$, derart dass $H \circ F = \mathrm{id}_U$.
- F ist surjektiv $\Leftrightarrow$ $\exists\ G : V \to U$, derart dass $F \circ G = \mathrm{id}_V$.
- F ist bijektiv $\Leftrightarrow$ $\exists\ F^{-1} : V \to U$, derart dass $F \circ F^{-1} = \mathrm{id}_V$ und $F^{-1} \circ F = \mathrm{id}_U$.

Folgende Proposition plus Beispiel sind echte Übungsaufgaben.

Proposition 6.7. (Komposition von linearen Abbildungen) *Seien U, V, W drei K-Vektorräume über einem Körper K und seien $F : U \to V$ sowie $G : V \to W$ zwei lineare Abbildungen. Dann gilt:*

(a) $G \circ F : U \to W$ *ist auch eine lineare Abbildung.*

(b) *Ist insbesondere $F : U \to V$ ein Isomorphismus, so ist auch $F^{-1} : V \to U$ ein Isomorphismus.*

(c) $\mathrm{id}_U : U \to U$ *ist ein Automorphismus.*

Beispiel: Sei V ein n−dimensionaler Vektorraum über einem Körper K mit Basis $\mathcal{B} = (v_1, \ldots, v_n)$. Dann ist die wichtige Abbildung

$$\begin{aligned} \phi_{\mathcal{B}} : \ K^{n\times 1} &\to V \\ x = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} &\mapsto \phi_{\mathcal{B}}(x) := \sum_{i=1}^{n} \lambda_i v_i \end{aligned}$$

ein Vektorraumisomorphismus mit der eindeutigen inversen Abbildung

$$\begin{aligned} I_{\mathcal{B}} := \phi_{\mathcal{B}}^{-1} : \ V &\to K^{n\times 1} \\ v = \sum_{i=1}^{n} \lambda_i v_i &\mapsto I_{\mathcal{B}}(v) := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \ , \end{aligned}$$

wobei $\lambda_1, \ldots, \lambda_n \in K$ die nach Satz 3.14 bzw. Korollar 3.16 eindeutigen **Koordinaten** der Darstellung von v als Linearkombination von $\mathcal{B}$ sind.

Satz 6.8. (Dimensionsformel) *Seien V, W zwei K-Vektorräume über einem Körper K und sei $F : V \to W$ eine lineare Abbildung. Dann gilt:*

$$\dim_K V = \dim_K \operatorname{Kern}(F) + \dim_K \operatorname{Bild}(F) \ .$$

Dieser Satz ist einer der zentralen Sätze der Vektorraumtheorie und der Theorie der linearen Abbildungen. Es gibt vielfältige Konsequenzen, und wir werden noch einige Anwendungen sehen. Das Resultat ist für alle Teilnehmer der Vorlesung fundamental wichtig.

Definition Seien V, W zwei K-Vektorräume über einem Körper K und sei $F : V \to W$ eine lineare Abbildung. Dann nennt man

$$\operatorname{rang}(F) := \dim_K \operatorname{Bild}(F)$$

den **Rang von** F.

Korollar 6.9. *Seien V, W zwei K-Vektorräume über einem Körper K und sei die Abbildung $F : V \to W$ eine lineare Abbildung. Dann gilt:*

$$\dim_K V = \dim_K \operatorname{Kern}(F) + \operatorname{rang}(F) \ .$$

Korollar 6.10. *Sei $A \in K^{m \times n}$, wobei K ein Körper ist und $m, n \in \mathbb{N}$. Dann erfüllt die Standardinterpretation von A*

$$\begin{array}{rccl} F_A : & K^{n \times 1} & \to & K^{m \times 1} \\ & x & \mapsto & F_A(x) = A\,x \end{array}$$

als lineare Abbildung die Beziehung

$$\operatorname{rang}(F_A) = \operatorname{rang}(A) \ .$$

6.3 Charakterisierung von linearen Abbildungen

Der folgende Satz liefert die zentrale Charakterisierung linearer Abbildungen. Damit ist eine lineare Abbildung durch ganz wenig Information schon vollständig beschrieben. Außerdem kann man immer eine lineare Abbildung zwischen zwei Vektorräumen mittels Basen konstruieren.

Satz 6.11. *Sei V ein endlich-dimensionaler K-Vektorraum über einem Körper K und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Ist $(w_1, \ldots, w_n)$ ein beliebiges System von Vektoren eines K-Vektorraumes W, dann existiert* ***genau eine*** *lineare Abbildung $F : V \to W$, so dass*

$$F(v_j) = w_j \qquad \text{für alle } j = 1, \ldots, n \;.$$

Lemma 6.12. (Eindeutigkeitsaussage) *Seien V, W zwei K-Vektorräume über einem Körper K und seien $F, G : V \to W$ zwei lineare Abbildungen. Ist $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V und gilt*

$$F(v_j) = G(v_j) \qquad \text{für alle } j = 1, \ldots, m \;,$$

so folgt $G = F$, d.h. die linearen Abbildungen sind identisch.

Satz 6.13. *Seien V, W zwei K-Vektorräume über einem Körper K und sei $F : V \to W$ eine lineare Abbildung. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und sei*

$$w_j := F(v_j) \quad \text{für alle } i = 1, \ldots, n \;.$$

Dann gilt:

(a) *F ist genau dann ein Monomorphismus, wenn $(w_1, \ldots, w_n)$ linear unabhängig über K ist.*

(b) *F ist genau dann ein Epimorphismus, wenn $(w_1, \ldots, w_n)$ ein Erzeugendensystem von W über K ist.*

(c) *F ist genau dann ein Isomorphismus, wenn $(w_1, \ldots, w_n)$ eine K-Basis von W ist.*

Satz 6.14. (Fundamentalsatz für endlich-dimensionale Vektorräume) *Seien V, W zwei endlich-dimensionale K-Vektorräume über einem Körper K. Es gilt: V und W sind genau dann isomorph, wenn $\dim_K V = \dim_K W$.*

Korollar 6.15. *Ist V ein n-dimensionaler K-Vektorraum, so ist V zu $K^{n\times 1}$ isomorph.*

Satz 6.16. *Seien V, W zwei endlich erzeugte K-Vektorräume über einem Körper K mit $\dim_K V = \dim_K W$, und sei $F : V \to W$ eine lineare Abbildung. Dann sind äquivalent:*

(a) *F ist ein Monomorphismus.*

(b) *F ist ein Epimorphismus.*

(c) *F ist ein Isomorphismus.*

Definition Seien V, W zwei K-Vektorräume über einem Körper K. Die Menge aller K-linearen Abbildungen von V nach W bezeichnen wir mit

$$\mathrm{Hom}_K(V, W) = \{ F : V \to W \mid F \text{ lineare Abbildung } \} \quad .$$

Proposition 6.17. *Seien V, W zwei K-Vektorräume über einem Körper K. Dann ist die Menge* $\mathrm{Hom}_K(V, W)$ *ein Untervektorraum von* $\mathrm{Abb}(V, W)$*, selbst also wieder ein K-Vektorraum unter den von* $\mathrm{Abb}(V, W)$ *induzierten Operationen.*

6.4 Lineare Abbildungen, Koordinaten und Matrizen

Wir erklären nun den Zusammenhang zwischen linearen Abbildungen und Matrizen. Insbesondere werden wir die Matrizenoperationen wiederfinden.

Satz 6.18. *Sei K ein Körper. Dann ist die Abbildung*

$$\begin{array}{rlcl} \tau: & K^{m\times n} & \to & K^{n\times m} \\ & A = (a_{ij}) & \mapsto & \tau(A) = A^T = (a_{ji}) \end{array}$$

ein Isomorphismus. Insbesondere gelten:

(a) $(A+B)^T = A^T + B^T$ *für alle* $A, B \in K^{m\times n}$.

(b) $(\lambda A)^T = \lambda A^T$ *für alle* $A \in K^{m\times n}$ *und* $\lambda \in K$.

Definition Sei $F : V \to W$ eine lineare Abbildung von endlich-dimensionalen K-Vektorräumen über einem Körper K. Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine K-Basis von V und $\mathcal{C} = (w_1, \dots, w_m)$ eine K-Basis von W, so existiert eine Matrix $A := (a_{ij}) \in K^{m\times n}$, welche durch folgende Relation eindeutig bestimmt ist:

$$F(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad \text{für alle } j \in \{1, \dots, n\} .$$

Die eindeutige Matrix A heißt die **darstellende Matrix von** F bzgl. der K-Basen $\mathcal{B}$ und $\mathcal{C}$ oder einfach die F **darstellende Matrix** bzgl. der K-Basen $\mathcal{B}$ und $\mathcal{C}$. Wir schreiben

$$M_{\mathcal{C}}^{\mathcal{B}}(F) := A = (a_{ij}) = (a_{ij})_{i,j} \in K^{m\times n} .$$

Für einen beliebigen Vektor $v = \sum_{j=1}^{n} \lambda_j v_j \in V$ mit $\lambda_1, \dots, \lambda_n \in K$ berechnet sich der Funktionswert $F(v)$ als

$$F(v) = F\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} \lambda_j a_{ij}\right) w_i .$$

Im Falle $V = W$ und $F = \mathrm{id}_V$ nennt man

$$M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_V) \in K^{m\times n}$$

auch die **Transformationsmatrix der Koordinaten** bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$ von V und schreibt auch $T_{\mathcal{C}}^{\mathcal{B}}$.

Aus den Beispielen in der Vorlesung geht hervor, dass wir die Koordinaten bzgl. der Basisdarstellungen benötigen. Wir betrachten die folgende Situation: Seien V, W zwei endlich-dimensionale K-Vektorräume über einem Körper K.

- Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V.
- Sei $\mathcal{C} = (w_1, \ldots, w_m)$ eine Basis von W.
- Sei $F : V \to W$ eine lineare Abbildung.
- Sei $A = (a_{ij}) := M_{\mathcal{C}}^{\mathcal{B}}(F) \in K^{m\times n}$, so dass also
$$F(v_j) = \sum_{i=1}^{m} a_{ij}\, w_i \qquad \text{für alle } j \in \{1, \ldots, n\} .$$
- Die folgende Abbildung $I_{\mathcal{B}}$ ist nach Satz 6.13 ein Vektorraumisomorphismus,
$$\begin{aligned} I_{\mathcal{B}} :\ V &\to K^{n\times 1} \\ v = \sum_{j=1}^{n} \lambda_j v_j &\mapsto I_{\mathcal{B}}(v) := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} , \end{aligned}$$
weil $\dim_K V = n$ und mit der Standardbasis $(e_1, \ldots, e_n)$ des $K^{n\times 1}$ gilt:
$$I_{\mathcal{B}}(v_j) = e_j \qquad \text{für alle } j \in \{1, \ldots, n\} .$$
$I_{\mathcal{B}}$ nennt man den **Koordinaten-Isomorphismus** von V bzgl. der Basis $\mathcal{B}$.
- Analog ist die folgende Abbildung $I_{\mathcal{C}}$ nach Satz 6.13 ein Vektorraumisomorphismus,
$$\begin{aligned} I_{\mathcal{C}} :\ W &\to K^{m\times 1} \\ w = \sum_{i=1}^{m} \mu_i w_i &\mapsto I_{\mathcal{C}}(w) := \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} , \end{aligned}$$
weil $\dim_K w = m$ und mit der Standardbasis $(e'_1, \ldots, e'_m)$ des $K^{m\times 1}$ gilt:
$$I_{\mathcal{C}}(w_i) = e'_i \qquad \text{für alle } i \in \{1, \ldots, m\} .$$
$I_{\mathcal{C}}$ nennt man den **Koordinaten-Isomorphismus** von W bzgl. der Basis $\mathcal{C}$.

Satz 6.19. (Matrix-Darstellung linearer Abbildungen) *In obiger Situation gilt:*

$$I_{\mathcal{C}} \circ F = F_A \circ I_{\mathcal{B}} \ .$$

Dabei ist F_A die Standardinterpretation von A, also

$$\begin{aligned} F_A : \ K^{n\times 1} &\to K^{m\times 1} \\ x &\mapsto F_A(x) = Ax \ . \end{aligned}$$

Hier ein Bild zu obigem Satz: Das folgende Diagramm ist **kommutativ**.

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[d, "I_{\mathcal{B}}"'] & W \arrow[d, "I_{\mathcal{C}}"] \\ K^{n\times 1} \arrow[r, "F_A"'] & K^{m\times 1} \end{tikzcd}$$

Wir erhalten die folgenden Konsequenzen des Satzes. Für alle $j \in \{1, \ldots, n\}$ gilt:

$$(I_{\mathcal{C}} \circ F)(v_j) = (F_A \circ I_{\mathcal{B}})(v_j) = F_A(I_{\mathcal{B}}(v_j)) = F_A(e_j) = A e_j \ .$$

Andererseits ist $A e_j$ genau die j-te Spalte von A. Folglich ist A genau so definiert, dass

$$A = \Big(I_{\mathcal{C}}(F(v_1)) \ \ldots \ I_{\mathcal{C}}(F(v_n)) \Big) \in K^{m\times n} \ . \tag{6.1}$$

Man schreibt also die Koordinatenvektoren aller $F(v_j)$ zur Basis $\mathcal{C} = (w_1, \ldots, w_m)$ als Spalten in eine Matrix, und diese Matrix ist genau die darstellende Matrix von F bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$.

Satz 6.20. *Jede lineare Abbildung $F : K^{n\times 1} \to K^{m\times 1}$ ist von der Form*

$$F = F_A \quad \textit{wobei } A \in K^{m\times n} \ .$$

Satz 6.21. (Hauptsatz für lineare Abbildungen) *Seien V, W zwei endlich-dimensionale K-Vektorräume über einem Körper K mit gegebenen Basen $\mathcal{B} = (v_1, \ldots, v_n)$ von V und $\mathcal{C} = (w_1, \ldots, w_m)$ von W. Dann ist die Abbildung*

$$\begin{aligned} \Psi = \Psi_{\mathcal{C}}^{\mathcal{B}} : \ \mathrm{Hom}_K(V, W) &\to K^{m\times n} \\ F &\mapsto \Psi(F) := M_{\mathcal{C}}^{\mathcal{B}}(F) \end{aligned}$$

ein Isomorphismus von K-Vektorräumen.

Damit induziert nicht nur jede Matrix eine lineare Abbildung, sondern jede lineare Abbildung F endlich-erzeugter K-Vektorräume induziert auch eine darstellende Matrix von F bzgl. der Basen $\mathcal{B}$ und $\mathcal{C}$.

Korollar 6.22. *Für zwei endlich-dimensionale K-Vektorräume V, W über einem Körper K gilt:*

$$\dim_K \operatorname{Hom}_K(V, W) = (\dim_K V) \cdot (\dim_K W) \quad .$$

Wir werden nun sehen, dass die Komposition von linearen Abbildungen dem Matrizenprodukt entspricht.

Satz 6.23. (Komposition von linearen Abbildungen) *Seien U, V, W endlich-dimensionale K-Vektorräume über einem Körper K mit gegebenen K-Basen $\mathcal{B}_1$ von U, $\mathcal{B}_2$ von V sowie $\mathcal{B}_3$ von W. Außerdem seien $F : U \to V$ und $G : V \to W$ lineare Abbildungen. Dann gilt:*

$$M_{\mathcal{B}_3}^{\mathcal{B}_1}(G \circ F) = M_{\mathcal{B}_3}^{\mathcal{B}_2}(G) \cdot M_{\mathcal{B}_2}^{\mathcal{B}_1}(F) \ .$$

Anders ausgedrückt: Sei $\dim_K U = p$, $\dim_K V = n$ und $\dim_K W = m$. Da F, G und $G \circ F$ lineare Abbildungen sind, existieren die folgenden darstellenden Matrizen

$$A = M_{\mathcal{B}_3}^{\mathcal{B}_2}(G) \in K^{m\times n} \ , B = M_{\mathcal{B}_2}^{\mathcal{B}_1}(F) \in K^{n\times p} \ , C = M_{\mathcal{B}_3}^{\mathcal{B}_1}(G \circ F) \in K^{m\times p} \ .$$

Diese Matrizen erfüllen dann die Definition der Matrizenmultiplikation

$$A \cdot B = (a_{ij}) \cdot (b_{jk}) = (c_{ik}) = C \ ,$$

$$c_{ik} := \sum_{j=1}^{n} a_{ij}\, b_{jk} \qquad \text{für alle } i = 1, \dots, m \text{ und } k = 1, \dots, p \ .$$

$$\begin{tikzcd}
U \arrow[rr, "G\circ F"] \arrow[dr, "F"'] \arrow[ddd, "I_{\mathcal{B}_1}"'] & & W \arrow[ddd, "I_{\mathcal{B}_3}"] \\
 & V \arrow[ur, "G"'] \arrow[d, "I_{\mathcal{B}_2}"] & \\
 & K^{n\times 1} \arrow[dr, "F_A"] & \\
K^{p\times 1} \arrow[ur, "F_B"'] \arrow[rr, "F_C"'] & & K^{m\times 1}
\end{tikzcd}$$

Dabei verwenden wir die folgenden Relationen:

$$\begin{aligned}
I_{\mathcal{B}_3} \circ G &= F_A \circ I_{\mathcal{B}_2} \ , \\
I_{\mathcal{B}_2} \circ F &= F_B \circ I_{\mathcal{B}_1} \ , \\
I_{\mathcal{B}_3} \circ G \circ F &= F_C \circ I_{\mathcal{B}_1} \ , \\
F_C &= F_A \circ F_B \ .
\end{aligned}$$

Hier ein weiteres Beispiel für die Matrizenmultiplikation (vergleiche mit Satz 2.5).

Proposition 6.24. *Sei K ein Körper.*

(a) *Es seien $A \in K^{m \times n}$, $B \in K^{n \times p}$ und $C \in K^{p \times q}$. Dann gilt:*

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C \ .$$

(b) *Sei $A, A' \in K^{m \times n}$ und $B, B' \in K^{n \times p}$. Dann gilt:*

(i) $A \cdot (B + B') = A \cdot B + A \cdot B'$.

(ii) $(A + A') \cdot B = A \cdot B + A' \cdot B$.

Satz 6.25. *Es seien V, W zwei endlich-dimensionale K-Vektorräume, welche von gleicher Dimension $\dim_K V = n = \dim_K W$ über einem Körper K mit gegebenen K-Basen $\mathcal{B} = (v_1, \ldots, v_n)$ von V und $\mathcal{C} = (w_1, \ldots, w_n)$ von W. Außerdem sei $F : V \to W$ eine lineare Abbildung. Dann gilt: Die F darstellende Matrix $A := M_{\mathcal{C}}^{\mathcal{B}}(F) \in K^{n \times n}$ bzgl. $\mathcal{B}$ und $\mathcal{C}$ ist genau dann invertierbar, wenn F ein Isomorphismus ist. In diesem Fall gilt insbesondere, dass*

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

Somit erkennen wir auch die früheren Sätze 2.31 , 2.33 , 6.13 und 2.8.

Satz 6.26. *Sei $A \in K^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:*

(a) $A \in \mathrm{GL}(n, K)$, *d.h. A invertierbar.*

(b) *Es existiert eine quadratische Matrix $B \in K^{n \times n}$ mit $B A = E_n$.*

(c) *Es existiert eine quadratische Matrix $B \in K^{n \times n}$ mit $A B = E_n$.*

(d) $\mathrm{rang}(A) = n$.

Proposition 6.27. *Sei K ein Körper und sei $n \in \mathbb{N}$.*

(a) *Für alle $A, B \in \mathrm{GL}(n, K)$ gilt: $AB \in \mathrm{GL}(n, K)$ und*

$$(AB)^{-1} = B^{-1} A^{-1} .$$

(b) *Für alle $A \in \mathrm{GL}(n, K)$ gilt: $A^{-1} \in \mathrm{GL}(n, K)$ und*

$$(A^{-1})^{-1} = A .$$

(c) *Für alle $A \in \mathrm{GL}(n, K)$ gilt: $A^T \in \mathrm{GL}(n, K)$ und*

$$(A^T)^{-1} = (A^{-1})^T .$$

Definition Sei V ein K-Vektorraum. Dann heißt

$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V) = \{ F : V \to V \mid F \text{ lineare Abbildung} \}$$

die **Menge der (K-)Endomorphismen von** V. Außerdem heißt

$$\mathrm{Aut}_K(V) := \{ F : V \to V \mid F \text{ Isomorphismus} \} \subseteq \mathrm{End}_K(V)$$

die **Menge der (K-)Automorphismen von** V.

Bemerkung 6.28. *Sei V ein endlich-dimensionaler K-Vektorraum über einem Körper K der Dimension $\dim_K V = n$ mit gegebener K-Basis $\mathcal{B} = (v_1, \ldots, v_n)$.*

(a) *Neben* $\mathrm{GL}(n, K) = \{A \in K^{n\times n} \mid A \text{ invertierbar}\}$ *spielen* $\mathrm{End}_K(V)$ *und* $Aut_K(V)$ *in der Theorie linearer Abbildungen eine wichtige Rolle. Man schreibt auch* $\mathrm{GL}(V)$ *anstelle von* $\mathrm{Aut}_K(V)$ *wegen der in* (c) *beschriebenen strukturellen Gleichheit mit* $\mathrm{GL}(n, K)$.

(b) *Mit den obigen Resultaten kann man einfach folgern, dass* $\mathrm{End}_K(V)$ *eine Ringstruktur besitzt. Genauer ist die Abbildung*

$$\begin{array}{rlcl} \Phi: & \mathrm{End}_K(V) & \to & K^{n\times n} \\ & F & \mapsto & \Phi(F) := M_{\mathcal{B}}^{\mathcal{B}}(F) \end{array}$$

ein so genannter Ringisomorphismus, d.h. eine bijektive Abbildung, welche zusätzlich die folgenden Bedingungen erfüllt:

(i) $\Phi(F+G) = \Phi(F) + \Phi(F)$ *für alle* $F, G \in \mathrm{End}_K(V)$.

(ii) $\Phi(F \circ G) = \Phi(F) \cdot \Phi(F)$ *für alle* $F, G \in \mathrm{End}_K(V)$.

(c) *Die Automorphismen von V über K entsprechen dann den Elementen aus* $\mathrm{GL}(n, K)$. *D.h. schränkt man die Abbildung Φ auf* $\mathrm{Aut}_K(V)$ *ein, so gilt:*

$$\begin{array}{rlcl} \Phi: & \mathrm{Aut}_K(V) & \to & \mathrm{GL}(n, K) \\ & F & \mapsto & \Phi(F) := M_{\mathcal{B}}^{\mathcal{B}}(F) \end{array}$$

ist ein Ringisomorphismus.

Anwendung: Invertierbare Matrix-Darstellung

Wir hatten gelernt, wie man eine Matrix invertiert. Hier ist eine wichtige Anwendung. Sei $A \in \mathrm{GL}(n, K)$. Dann ist $\operatorname{rang}(A) = n$ und das System $\mathcal{B} = (s_1, \ldots, s_n)$ der Spalten von A bildet eine Basis von $K^{n\times 1}$. Sei $\mathcal{C}$ die Standardbasis des $K^{n\times 1}$, so dass

$$I_{\mathcal{C}}(x) = x \qquad \text{für alle } x \in K^{n\times 1} \ .$$

Man verifiziert leicht, dass dann in dieser Situation

$$A = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_{K^{n\times 1}}) \ .$$

Die Matrix-Darstellung von $\mathrm{id}_{K^{n\times 1}}$ ist dann

$$I_{\mathcal{C}} \circ \mathrm{id}_{K^{n\times 1}} = F_A \circ I_{\mathcal{B}} \ .$$

Für alle $x \in K^{n\times 1}$ rechnen wir deshalb

$$(I_{\mathcal{C}} \circ \mathrm{id}_{K^{n\times 1}})(x) = I_{\mathcal{C}}(\mathrm{id}_{K^{n\times 1}}(x)) = I_{\mathcal{C}}(x) = x \ ,$$

und

$$(F_A \circ I_{\mathcal{B}})(x) = F_A(I_{\mathcal{B}}(x)) = A \cdot I_{\mathcal{B}}(x) \ .$$

Insgesamt haben wir also:

$$A \cdot I_{\mathcal{B}}(x) = x \qquad \text{für alle } x \in K^{n\times 1} \ .$$

Konsequenz: Sollten wir die Inverse von A berechnet haben, so können wir die Koordinaten von jedem $x \in K^{n\times 1}$ mit der Formel

$$I_{\mathcal{B}}(x) = A^{-1} x$$

berechnen.

6.5 Basiswechsel und Äquivalenz von Matrizen

In vielen Anwendungen ist es wichtig, Basen zu wechseln. Die Idee ist, zu einer gegebenen Abbildung Basen der Vektorräume zu finden, sodass die Transformationsmatrix besonders einfach ist.

Sei K ein Körper. Seien V, W endlich-dimensionale K-Vektorräume mit gegebenen Basen $\mathcal{B} = (v_1, \ldots, v_n)$ von V und $\mathcal{C} = (w_1, \ldots, w_m)$ von W. Außerdem sei $F : V \to W$ eine lineare Abbildung. Zu verschiedenen Basen gehören verschiedene Matrizen. Wechselt man die Basen $\mathcal{B}$ und $\mathcal{C}$, so ändern sich auch die die darstellenden Matrizen. Wir werden dies nun beschreiben. Als erstes Beispiel erinnern wir an Satz 2.35 und den Beweis zu Satz 6.8. Dort hatten wir bewiesen, dass die Dimensionsformel gilt:

$$\dim_K V = \dim_K \operatorname{Kern}(F) + \underbrace{\dim_K \operatorname{Bild}(F)}_{= r = \operatorname{rang}(F)} .$$

In dem Beweis wurde explizit eine Basis konstruiert. Diese verwenden wir für den folgenden Satz.

Satz 6.29. *Seien V, W zwei endlich-dimensionale K-Vektorräume mit $\dim_K V = n$ und $\dim_K W = m$. Ist $F : V \to W$ eine lineare Abbildung mit $\dim_K \operatorname{Bild}(F) = r$, so gibt es K-Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W, so dass*

$$M_{\mathcal{C}}^{\mathcal{B}}(F) = \begin{pmatrix} E_r & 0_{r,(n-r)} \\ 0_{(m-r),r} & 0_{(m-r),(n-r)} \end{pmatrix}$$

Dabei ist E_r die r-te Einheitsmatrix, und $0_{r,(n-r)}$, $0_{(m-r),r}$, $0_{(m-r),(n-r)}$ sind die Nullmatrizen mit den angegebenen Dimensionen.

Möchte man für $F \in \operatorname{End}_K(V)$ und $\mathcal{B} = \mathcal{C}$ eine einfachere darstellende Matrix von F bzgl. Basen $\mathcal{B}'$ und $\mathcal{C}'$ finden, so ist dieses Problem erheblich schwieriger. Das Problem werden wir in Kapitel 8 behandeln.

Wir leiten nun die allgemeine Formel für den Basiswechsel her und folgern daraus die Transformationsformel.

Satz 6.30. (Basiswechsel) *Seien* V, W *endlich-dimensionale* K*-Vektorräume der Dimension* $\dim_K V = n$ *und* $dim_K W = m$ *mit gegebenen* K*-Basen* $\mathcal{B}$ *von* V *und* $\mathcal{C}$ *von* W. *Außerdem sei* $F : V \to W$ *eine lineare Abbildung und sei* $M_{\mathcal{C}}^{\mathcal{B}}(F) \in K^{m \times n}$ *die* F *darstellende Matrix bzgl.* $\mathcal{B}$ *und* $\mathcal{C}$. *Seien nun weitere* K*-Basen* $\mathcal{B}'$ *von* V *und* $\mathcal{C}'$ *von* W *gegeben, und sei* $M_{\mathcal{C}'}^{\mathcal{B}'}(F) \in K^{m \times n}$ *die* F *darstellende Matrix bzgl. der Basen* $\mathcal{B}'$ *und* $\mathcal{C}'$. *Dann gilt:*

$$M_{\mathcal{C}'}^{\mathcal{B}'}(F) = M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_W) \cdot M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \ ,$$

wobei $M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_W) \in \mathrm{GL}(m, K)$ *und* $M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_W) \in \mathrm{GL}(n, K)$.

Man kann dies auch wie folgt schreiben:

$$M_{\mathcal{C}'}^{\mathcal{B}'}(F) = M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_W) \cdot M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V)^{-1} \ .$$

Setzt man

$$A := M_{\mathcal{C}}^{\mathcal{B}}(F) \quad , \quad A' := M_{\mathcal{C}'}^{\mathcal{B}'}(F) \quad , \quad Q := M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_W) \quad , \quad P := M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \ ,$$

so liest sich die Aussage als

$$A' = Q\,A\,P \ .$$

Die folgende Illustration wird in der Vorlesung gegeben:

$$\begin{tikzcd}
V \arrow[rrr, "F"] \arrow[ddd, "I_{\mathcal{B}'}"'] \arrow[dr, "\mathrm{id}_V"] & & & W \arrow[ddd, "I_{\mathcal{C}'}"] \\
& V \arrow[r, "F"] \arrow[d, "I_{\mathcal{B}}"] & W \arrow[ur, "\mathrm{id}_W"'] \arrow[d, "I_{\mathcal{C}}"] & \\
& K^{n\times 1} \arrow[r, "F_A"] & K^{m\times 1} \arrow[dr, "F_Q"] & \\
K^{n\times 1} \arrow[ur, "F_P"] \arrow[rrr, "F_{A'}"'] & & & K^{m\times 1}
\end{tikzcd}$$

Im Beweis verwendet man die folgenden Resultate und kombiniert diese mit Satz 6.23:

$$I_{\mathcal{B}} \circ \mathrm{id}_V = F_P \circ I_{\mathcal{B}'} \quad , \quad I_{\mathcal{C}'} \circ \mathrm{id}_W = F_Q \circ I_{\mathcal{C}'} \ ,$$

$$I_{\mathcal{C}} \circ F = F_A \circ I_{\mathcal{B}} \quad , \quad I_{\mathcal{C}'} \circ F = F_{A'} \circ I_{\mathcal{B}'} \ ,$$

$$F_P : K^{n\times 1} \to K^{n\times 1}, F_P(x) = P\,x \quad , \quad F_Q : K^{m\times 1} \to K^{m\times 1}, F_Q(x) = Q\,x \ ,$$

$$F_A : K^{n\times 1} \to K^{m\times 1}, F_A(x) = A\,x \quad , \quad F_{A'} : K^{n\times 1} \to K^{m\times 1}, F_{A'}(x) = A'\,x \ .$$

Definition Zwei Matrizen $A, B \in K^{m\times n}$ heißen **äquivalent**, wenn es invertierbare Matrizen $Q \in \mathrm{GL}(m, K)$ und $P \in \mathrm{GL}(n, K)$ gibt, so dass

$$B = Q\,A\,P \ .$$

Satz 6.31. *Sei K ein Körper.*

(a) *Zu einer beliebigen Matrix $A \in K^{m\times n}$ gibt es zwei Matrizen $Q \in \mathrm{GL}(m, K)$ und $P \in \mathrm{GL}(n, K)$, so dass A und die Matrix*

$$\begin{pmatrix} E_r & 0_{r,(n-r)} \\ 0_{(m-r),r} & 0_{(m-r),(n-r)} \end{pmatrix} \in K^{m\times n}$$

äquivalent sind. Hierbei ist $r = \mathrm{rang}(A)$.

(b) *Zwei Matrizen $A, B \in K^{m\times n}$ sind genau dann äquivalent, wenn gilt:*

$$\mathrm{rang}(A) = \mathrm{rang}(B) \ .$$

Definition Zwei quadratische Matrizen $A, B \in K^{n\times n}$ heißen **ähnlich**, wenn es eine invertierbare Matrix $P \in \mathrm{GL}(n, K)$ gibt, so dass

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

Satz 6.32. (Transformationsformel für Endomorphismen) *Sei V ein endlich-dimensionaler K-Vektorraum der Dimension $\dim_K V = n$ mit gegebener K-Basis $\mathcal{B}$. Ist nun $F \in \mathrm{End}_K(V)$ mit darstellender Matrix $M^{\mathcal{B}}_{\mathcal{B}}(F)$ bzgl. $\mathcal{B}$, dann gilt für jede weitere K-Basis $\mathcal{B}'$ von V mit darstellender Matrix $M^{\mathcal{B}'}_{\mathcal{B}'}(F)$ bzgl. $\mathcal{B}'$:*

$$M^{\mathcal{B}'}_{\mathcal{B}'}(F) = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V)\cdot M^{\mathcal{B}}_{\mathcal{B}}(F)\cdot M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) \ ,$$

wobei $M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) \in \mathrm{GL}(n, K)$. Setzt man

$$P := M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) \quad , \quad A := M^{\mathcal{B}}_{\mathcal{B}}(F) \quad , \quad A' := M^{\mathcal{B}'}_{\mathcal{B}'}(F) \ ,$$

so ist die Behauptung äquivalent zu

$$A' = P^{-1}\,A\,P \ .$$

6.6 Der duale Raum und lineare Funktionale

In späteren Vorlesungen (z.B. Funktionalanalysis) werden die Begriffe lineare Funktionale, Dualraum und Bidualraum eine Rolle spielen. Deshalbbwollen wir diese Begriffe hier kurz als Spezialfall erwähnen. Die Beweise ignorieren wir, da sie leichte Übungsaufgaben sind. Sie eignen sich jedoch ausgezeichnet zur Nachbereitung der bearbeiteten Themen.

Satz 6.33. *Sei K ein Körper und seien V, W endlich-dimensionale K-Vektorräume mit gegebenen K-Basen $\mathcal{B} = (v_1, \dots, v_n)$ von V und $\mathcal{C} = (w_1, \dots, w_m)$ von W. Das folgende System*

$$((F_{k\ell}\,;\, 1 \leq k \leq m\,,\, 1 \leq \ell \leq n)$$

der so genannten ***Elementarabbildungen*** *bildet eine K-Basis von $\mathrm{Hom}_K(V, W)$. Dabei ist $F_{k\ell} : V \to W$ definiert durch*

$$F_{k\ell}(v_i) := \begin{cases} w_k & , \text{für } i = \ell \\ 0 & , \text{für } i \neq \ell \end{cases} \qquad (1 \leq k \leq m\,, 1 \leq \ell \leq n)\,.$$

Insbesondere gilt also

$$\dim_K \mathrm{Hom}_K(V, W) = m\,n\,.$$

Der Spezialfall $W = K$ ist in dieser Situation sehr wichtig.

Definition Sei K ein Körper und sei V ein endlich-dimensionaler K-Vektorraum. Der Vektorraum $\mathrm{Hom}_K(V, K)$ heißt der zu V **duale Raum** bzw. der **Dualraum von** V. Wir schreiben

$$V^* := \mathrm{Hom}_K(V, K)\,.$$

Die Elemente $F : V \to K$ von $V^* = \mathrm{Hom}_K(V, K)$ heißen **lineare Funktionale**.

Korollar 6.34. *Sei K ein Körper und sei V ein endlich-dimensionaler K-Vektorraum mit gegebener Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V. Das System $(F_1, \dots, F_n)$ bildet eine K-Basis von $\mathrm{Hom}_K(V, W)$, wobei $F_\ell : V \to K$ definiert ist durch*

$$F_\ell(v_i) := \delta_{\ell i} = \begin{cases} 1 & , \text{für } i = \ell \\ 0 & , \text{für } i \neq \ell \end{cases} \qquad (1 \leq \ell \leq n)\,.$$

Insbesondere gilt also $\dim_K \mathrm{Hom}_K(V, K) = n$.

Definition Sei K ein Körper und sei V ein endlich-dimensionaler K-Vektorraum mit K-Basis $\mathcal{B} = (v_1, \ldots, v_n)$. Die in Korollar bzgl. $\mathcal{B}$ definierte K-Basis $(F_1, \ldots, F_n)$ des zu V dualen Vektorraums V^* heißt die zu $\mathcal{B} = (v_1, \ldots, v_n)$ **duale Basis**. Wir schreiben

$$(v_1^*, \ldots, v_n^*) := (F_1, \ldots, F_n) \ .$$

Für beliebiges $v \in V$ existieren $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$v = \sum_{j=1}^{n} \lambda_j v_j \in V \ . \quad \Rightarrow \quad F(v) = F\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j \underbrace{F(v_j)}_{=: a_j \cdot 1} = \sum_{j=1}^{n} \lambda_j a_j$$

Wir erhalten dann aus dem Korollar, dass

$$F_\ell(v) = \sum_{j=1}^{n} \lambda_j \underbrace{F_\ell(v_j)}_{=: \delta_{\ell j}} = \lambda_\ell \ . \quad \Rightarrow \quad F(v) = \sum_{j=1}^{n} F_\ell(v)\, a_j = \sum_{j=1}^{n} a_j F_\ell(v) \ .$$

Damit läßt sich nun die zu F **duale Abbildung** definieren (siehe später).

Satz 6.35. *Sei K ein Körper und sei V ein endlich-dimensionaler K-Vektorraum der Dimension n. Dann ist für $F \in \mathrm{Hom}_K(V, K)$ mit $F \neq 0$ immer*

$$\dim_K \mathrm{Kern}(F) = n - 1 \ .$$

7 Determinanten

Wir werden nun sehen, wie sich diese Resultate auf $n \times n$ Matrizen verallgemeinern. Dazu müssen wir Determinanten, Minoren, Adjungierte und Cramersche Regel einführen.

Was von der Schule nicht bekannt ist, ist die Tatsache, dass Determinanten ein wichtiges Beispiel von Multilinearformen darstellen. Determinanten werden in vielen Bereichen der Mathematik angewendet. Zum Beispiel beschreibt eine Determinante die Fläche oder das Volumen geometrischer Objekte, hilft beim Ermitteln von Lösungen von linearen Gleichungssystemen oder bei der Invertierbarkeit von quadratischen Matrizen.

7.1 Axiomatische Definition à la Weierstraß und Eigenschaften

Sei $A = (a_{ij}) \in K^{n\times n}$. Sind $z_1, \dots, z_n \in K^{1\times n}$ die Zeilen von A, so haben wir

$$A = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \in K^{n\times n} \quad , \quad \text{wobei } z_k := \begin{pmatrix} a_{k1} & \dots & a_{kn} \end{pmatrix} \in K^{1\times n} \text{ für } k \in \{1, \dots, n\} \quad .$$

Definition Sei K ein Körper und sei $n \in \mathbb{N}$. Eine Abbildung

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

heißt **Determinante**, wenn sie die folgenden Axiome erfüllt:

(D1) det ist in jeder Zeile **linear** (multilinear). D.h. für alle $k = 1, \dots, n$ gilt:

(D1a)

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_k + z_k' \\ \vdots \\ z_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ z_k \\ \vdots \\ z_n \end{pmatrix} + \det \begin{pmatrix} z_1 \\ \vdots \\ z_k' \\ \vdots \\ z_n \end{pmatrix}$$

für alle $z_1, z_2, \dots, z_k, \dots, z_n, z_k' \in K^{1\times n}$.

(D1b)

$$\det \begin{pmatrix} z_1 \\ \vdots \\ \lambda \cdot z_k \\ \vdots \\ z_n \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_k \\ \vdots \\ z_n \end{pmatrix}$$

für alle $z_1, z_2, \dots, z_k, \dots, z_n \in K^{1\times n}$ und $\lambda \in K$.

(D2) det ist **alternierend**: $\det(A) = 0$, falls A zwei gleiche Zeilen hat.

(D3) det ist **normiert**: $\det(E_n) = 1$.

Beispiel: Sei K ein beliebiger Körper. Wir betrachten die früher definierte Determinante für $n = 2$ als die Abbildung

$$\begin{aligned} \det : \quad & K^{2\times 2} && \to \quad K \\ A = & \begin{pmatrix} a & b \\ c & d \end{pmatrix} && \mapsto \quad \det(A) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := a\,d - b\,c \ . \end{aligned}$$

Wir verifizieren nun, dass diese Funktion tatsächlich eine Determinante ist.

(D1) det ist in jeder Zeile **linear**.

(D1a)

$$\begin{aligned} \det \begin{pmatrix} (a + a') & (b + b') \\ c & d \end{pmatrix} &= (a + a')\,d - (b + b')\,c \\ &= (a\,d - b\,c) + (a'\,d - b'\,c) \\ &= \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \det \begin{pmatrix} a' & b' \\ c & d \end{pmatrix} \ . \end{aligned}$$

(D1b) Für alle $\lambda \in K$:

$$\det \begin{pmatrix} \lambda\,a & \lambda\,b \\ c & d \end{pmatrix} = \lambda\,a\,d - c\,\lambda\,b = \lambda \det(A) = \det \begin{pmatrix} a & b \\ \lambda\,c & \lambda\,d \end{pmatrix}$$

(D2) det ist **alternierend**:

$$\det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = a\,b - b\,a \ .$$

(D3) det ist **normiert**:

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

Fragen: Wir werden Antworten zu den folgenden Fragen geben.

(a) Existiert eine Determinante? Wenn ja, wieviele?

(b) Wie kann ich die Determinante berechnen? Insbesondere, wie kann ich die Determinante einer Matrix effizient berechnen?

(c) Wie verhält sich die Determinante bei elementaren Zeilenumformungen?

(d) Welche speziellen Eigenschaften hat eine Determinante? Zum Beispiel, was ist $\det(A^T)$, wenn man $\det(A)$ kennt?

(e) Wie hängen der Rang einer Matrix und Determinante einer Matrix zusammen?

Im Hauptsatz für Determinanten werden wir sehen, dass genau eine Determinante existiert. Das werden wir anschließend verwenden. Wie üblich werden wir zuerst wichtige Eigenschaften aus den Axiomen ableiten und dann auch sehen, wie man eine Determinante einfach berechnen kann. Sei fortan K ein Körper und sei $n \in \mathbb{N}$.

Satz 7.1. *Sei* $A \in K^{n\times n}$ *und sei* $\det : K^{n\times n} \to K$ *eine Determinante. Dann gilt:*

(a) $\det(\lambda A) = \lambda^n \det(A)$ *für alle* $\lambda \in K$.

(b) *Hat* A *eine Nullzeile, so ist* $\det(A) = 0$.

Bemerkung 7.2. *Sei* $A \in K^{n\times n}$ *und sei* $\det : K^{n\times n} \to K$ *eine Abbildung. Dann ist die Eigenschaft* **(D1)** *offensichtlich zu der folgenden Bedingung äquivalent.*

(D1') *Für alle* $k = 1, \ldots, n$ *gilt:*

$$\det\begin{pmatrix} z_1 \\ \vdots \\ \lambda\cdot z_k + \lambda'\cdot z'_k \\ \vdots \\ z_n \end{pmatrix} = \lambda\cdot\det\begin{pmatrix} z_1 \\ \vdots \\ z_k \\ \vdots \\ z_n \end{pmatrix} + \lambda'\cdot\det\begin{pmatrix} z_1 \\ \vdots \\ z'_k \\ \vdots \\ z_n \end{pmatrix}$$

für alle $z_1, z_2, \ldots, z_k, \ldots, z_n, z'_k \in K^{1\times n}$ *und* $\lambda, \lambda' \in K$.

Noch allgemeiner gilt, wenn (D1) *erfüllt ist, für alle* $k = 1, \ldots, n$*:*

$$\det\begin{pmatrix} z_1 \\ \vdots \\ \sum_{i=1}^{m}\lambda_i\cdot v_i \\ \vdots \\ z_n \end{pmatrix} = \sum_{i=1}^{n}\lambda_i\cdot\det\begin{pmatrix} z_1 \\ \vdots \\ v_i \\ \vdots \\ z_n \end{pmatrix}$$

für alle $z_1, \ldots, z_{k-1}, z_{k+1}, \ldots, z_k, \ldots, z_n, v_1, \ldots, v_n \in K^{1\times n}$ *und* $\lambda_1, \ldots, \lambda_n \in K$.

Satz 7.3. (Elementare Zeilenumformungen und Determinanten)
Sei $A \in K^{n\times n}$, *sei* $\det : K^{n\times n} \to K$ *eine Determinante und seien* $k, \ell \in \{1, \ldots, n\}$.

(a) **Typ I:** *Entsteht* A_I *aus* A *durch Multiplikation der* k*-ten Zeile von* A *mit einem Skalar* $\lambda \in K^*$, *so gilt:*
$$\det(A_I) = \lambda\cdot\det(A)\ .$$

(b) **Typ II:** *Entsteht* A_{II} *aus* A *durch Addition des* μ*-fachen* ($\mu \in K$, $k \neq \ell$) *der* ℓ*-ten Zeile zur* k*-ten Zeile von* A, *so gilt:*
$$\det(A_{II}) = \det(A)\ .$$

(c) **Typ III:** *Entsteht* A_{III} *aus* A, *indem man die* k*-te Zeile und die* ℓ*-te Zeile von* A *vertauscht, so gilt:*
$$\det(A_{III}) = -\det(A)\ .$$

Korollar 7.4. *Sei* $\det : K^{n\times n} \to K$ *eine Determinante. Dann gilt für die Elementarmatrizen:*

(a) $\det(S_k(\lambda)) = \lambda$ *für alle* $\lambda \in K^*$ *und* $k \in \{1, \ldots, n\}$.

(b) $\det(Q_{k\ell}(\mu)) = 1$ *für alle* $\mu \in K$ *und* $k, \ell \in \{1, \ldots, n\}$, $k \neq \ell$.

(c) $\det(P_{k\ell}) = -1$ *für alle* $k, \ell \in \{1, \ldots, n\}$, $k \neq \ell$.

Wir erhalten aus Satz 7.3 sowie Korollar 7.4 mit den zugehörigen Elementarmatrizen:

(a) $\det(S_k(\lambda)A) = \det(A_I) = \lambda \cdot \det(A) = \det(S_k(\lambda)) \cdot \det(A)$.

(b) $\det(Q_{k\ell}(\mu)A) = \det(A_{II}) = \det(A) = \det(Q_{k\ell}(\mu)) \cdot \det(A)$.

(c) $\det(P_{k\ell}A) = \det(A_{III}) = -\det(A) = \det(P_{k\ell}) \cdot \det(A)$.

Satz 7.5. *Sei* $A = (a_{ij}) \in K^{n\times n}$ *und sei* $\det : K^{n\times n} \to K$ *eine Determinante.*

(a) *Ist* A *eine obere Dreiecksmatrix, also* $a_{ij} = 0$ *für alle* $i, j \in \{1, \ldots, n\}$ *mit* $i > j$, *so gilt:*

$$\det(A) = \prod_{i=1}^{n} a_{ii} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn} \ .$$

(b) *Geht* $B \in K^{n\times n}$ *aus* A *durch Anwendung einer endlichen Folge von elementaren Zeilenumformungen vom Typ II und Typ III hervor, so gilt:*

$$\det(A) = (-1)^p \cdot \det(B) \ ,$$

wobei p *gleich der Anzahl der elementaren Zeilenumformungen vom Typ III ist, d.h. gleich der Anzahl der Zeilenvertauschungen.*

Bemerkung 7.6. *Sei* $A = (a_{ij}) \in K^{n\times n}$ *und sei* $\det : K^{n\times n} \to K$ *eine Determinante. Um* $\det(A)$ *zu berechnen, überführt man* A *durch elementare Zeilenumformungen vom Typ II und Typ III in eine Matrix* $B = (b_{ij})$ *in Zeilenstufenform und merkt sich die Anzahl* p *der Zeilenvertauschungen. Dann berechnet sich die Determinante von* A *als*

$$\det(A) = (-1)^p \cdot \det(B) = \det(A) = (-1)^p \cdot \prod_{i=1}^{n} b_{ii}$$

Führt man aber auch elementaren Zeilenumformungen vom Typ I durch, so muss man noch **(D1b)** *bzw. Satz 7.3 beachten und die Determinante mit entsprechenden Vielfachen verändern. Dies ist dann der Fall, wenn man* A *in Gaußsche Normalform oder reduzierte Gaußsche Normalform überführt.*

Satz 7.7. *Sei* $A \in K^{n \times n}$ *und sei* $\det : K^{n \times n} \to K$ *eine Determinante. Dann sind die folgenden Aussagen äquivalent:*

(a) $\mathrm{rang}(A) = n$.

(b) $\det(A) \neq 0$.

(c) A *invertierbar.*

Der folgende Satz ist sehr wichtig, wie wir sehen werden. Im Beweis verwenden wir die Theorie der Elementarmatrizen. Damit ist klar, welch wichtiges Hilfsmittel die Elementarmatrizen auch für Beweise sind. Auch sollte sich jeder klarmachen, warum dieses Resultat wichtig ist. Die Determinante setzt das Produkt zweier Matrizen mit dem Produkt zweier Körperelemente in Verbindung.

Satz 7.8. (Multiplikativität der Determinante)
Seien $A, B \in K^{n \times n}$ *und sei* $\det : K^{n \times n} \to K$ *eine Determinante. Dann gilt:*

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

Korollar 7.9. *Sei* $\det : K^{n \times n} \to K$ *eine Determinante. Dann gilt:*

(a) *Ist* $A \in \mathrm{GL}(n, K)$*, so gilt:* $\det(A^{-1}) = (\det(A))^{-1}$.

(b) *Sind* $A, B \in \mathrm{GL}(n, K)$*, so gilt:* $AB \in \mathrm{GL}(n, K)$.

(c) *Sind* $A, B \in K^{n \times n}$*, so gilt:* $\det(AB) = \det(BA)$.

(d) *Sind* $A, B \in K^{n \times n}$ *ähnlich, so gilt:* $\det(A) = \det(B)$.

(e) *Sei* $\mathrm{SL}(n, K) := \{ A \in K^{n \times n} \mid \det(A) = 1 \}$. *Dann gilt:* $\mathrm{SL}(n, K)$ *ist eine multiplikative Untergruppe von* $\mathrm{GL}(n, K)$*, d.h.* $\mathrm{SL}(n, K) \subseteq \mathrm{GL}(n, K)$ *und* $\mathrm{SL}(n, K)$ *ist unter Matrizenmultiplikation wieder eine Gruppe.*

Satz 7.10. (Eindeutigkeitsatz für Determinanten)
Für jedes $n \in \mathbb{N}$ *gibt es höchstens eine Determinante* $\det : K^{n \times n} \to K$. *Dies bedeutet: Sind* $\det_1, \det_2 : K^{n \times n} \to K$ *zwei Determinanten, so gilt*

$$\det{}_1(A) = \det{}_2(A) \qquad \text{für alle } A \in K^{n \times n} \; .$$

Es bleibt noch die Frage der Existenz zu klären. Hier werden wir prinzipiell zwei Ansätze beschreiben (Leibniz-Formel, Laplace-Entwicklung). Um die Leibniz-Formel einzuführen, müssen wir einen kleinen Exkurs über symmetrische Gruppen eingehen. Zuerst liefern wir noch einen kleinen Einschub über den rationalen Funktionenkörper.

7.2 Permutationen

Definition Sei X eine beliebige Menge. Dann nennt man die Menge

$$\operatorname{Sym}(X) := \{ f : X \to X \mid f \text{ ist eine bijektive Abbildung} \}$$

die **symmetrische Gruppe auf** X. Man nennt $f \in \operatorname{Sym}(X)$ auch eine **Permutation** von X. Ist speziell X die n-elementige Menge $X = \{1, \ldots, n\}$ für $n \in \mathbb{N}$, so heißt

$$S_n := \operatorname{Sym}(X) = \big\{ \sigma : \{1, \ldots, n\} \to \{1, \ldots, n\} \mid \sigma \text{ ist eine Permutation} \big\}$$

die **symmetrische Gruppe von** n **Elementen.**

Bemerkung 7.11. *Für eine beliebige nichtleere Menge* X *ist* $(\operatorname{Sym}(X), \circ)$ *tatsächlich eine Gruppe, wobei* $\circ$ *die Verknüpfung von Abbildungen bezeichnet.*

Satz 7.12. *Ist* X *eine endliche Menge mit* $n \in \mathbb{N}$ *Elementen, so ist*

$$|\operatorname{Sym}(X)| = n!\ .$$

Insbesondere ist $(S_n, \circ)$ *für jedes* $n \in \mathbb{N}$ *eine Gruppe der Ordnung* $|S_n| = n!$.

Notation: Ist $\sigma \in S_n$, so schreiben wir σ mithilfe aller Bilder als

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n-1 & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n-1) & \sigma(n) \end{pmatrix} \in S_n\ .$$

Sind $\sigma, \tau \in S_n$, so schreiben wir auch $\sigma\tau$ für $\sigma \circ \tau$. Beim Rechnen geht man immer von rechts nach links vor. Wir sprechen auch vom Produkt von Permutationen.

Die Gruppe S_n ist im Allgemeinen nicht kommutativ für $n \geq 3$. Ein Beispiel für S_7 wird in der Vorlesung gegeben.

Für $\sigma \in S_n$ und $m \in \mathbb{N}$ bezeichnet σ^m die m**-malige Hintereinanderausführung** von σ. Dies bedeutet für $i \in \{1, \ldots, n\}$, dass

$$\sigma^m(i) = \underbrace{(\sigma \circ \ldots \circ \sigma)}_{m \text{ mal}}(i) = \sigma(\sigma(\ldots(\sigma(i))))\ .$$

Dazu setzen wir

$$\sigma^{-m} := (\sigma^{-1})^m \quad \text{und} \quad \sigma^0 := \mathrm{id}_X\ .$$

Definition Eine Permutation $\tau \in S_n$ heißt ein **Zykel der Länge** $m \in \mathbb{N}$ oder einfach ein m-**Zyklus**, wenn ein $j \in \{1, \dots, n\}$ existiert, so dass gilt:

(i) Die Werte $j, \tau(j), \tau^2(j), \dots, \tau^{m-1}(j)$ sind paarweise verschieden.

(ii) $\tau^m(j) = j$.

(iii) Für alle $x \notin \{j, \tau(j), \tau^2(j), \dots, \tau^{m-1}(j)\}$ haben wir $\tau(x) = x$.

Wir schreiben $\tau := (\, j\ \tau(j)\ \tau^2(j)\ \dots\ \tau^{m-1}(j)\,)$. Außerdem heißen zwei Zyklen

$$\tau_1 = (\, x_1\ x_2\ \dots\ x_m \,) \quad , \quad \tau_2 = (\, y_1\ y_2\ \dots\ y_r \,)$$

elementfremd oder **disjunkt**, wenn $\{x_1, \dots, x_m\} \cap \{y_1, \dots, y_r\} = \varnothing$.

Man kann leicht durch Fallunterscheidung verifizieren, dass **disjunkte Zykel kommutieren**. D.h. seien $\tau_1 = (\, x_1\ x_2\ \dots\ x_m \,)$, $\tau_2 = (\, y_1\ y_2\ \dots\ y_r \,) \in S_n$ Zykel der Länge m bzw. r, so gilt:

$$\{x_1, \dots, x_m\} \cap \{y_1, \dots, y_r\} = \varnothing \quad \Rightarrow \quad \tau_1 \circ \tau_2 = \tau_2 \circ \tau_1 \ .$$

Satz 7.13. *Für alle $n \in \mathbb{N}$ mit $n \geq 2$ lässt sich jedes $\sigma \in S_n$ als Produkt elementfremder Zyklen schreiben. Diese Darstellung ist bis auf die Reihenfolge eindeutig bestimmt.*

Definition Ein Zykel $\tau \in S_n$ der Länge 2 heißt **Transposition**. Es existieren dann Indizes $i, j \in \{1, \dots, n\}$ mit $\tau = (\, i\ j \,)$, d.h.

$$\tau(j) = i \quad , \quad \tau(i) = j \quad , \quad \tau(k) = k \quad \text{für alle } k \in \{1, \dots, n\} \setminus \{i, j\} \ .$$

τ vertauscht genau zwei Werte und hält sonst alle Werte fest.

Wir bemerken, dass für einen Zyklus $(i_1\, i_2 \dots i_k) \in S_n$ das inverse Element sofort gegeben ist, nämlich

$$(\, i_1\, i_2 \dots i_k \,)^{-1} = (\, i_k\, i_{k-1} \dots i_1 \,) \ .$$

Alle Transpositionen sind sogar selbst-invers, d.h.

$$(\, i\, j \,)^{-1} = (\, i\, j \,) \ .$$

Mittels obigem Satz und der folgenden Identität

$$(\, i_1\, i_2 \dots i_k \,) = (\, i_1\, i_2 \,) \circ (\, i_2\, i_3 \,) \circ \dots (\, i_{k-2}\, i_{k-1} \,) \circ (\, i_{k-1}\, i_k \,)$$

ergibt sich demnach:

Satz 7.14. *Jede Permutation lässt sich als endliches Produkt von Transpositionen schreiben. Dieses Produkt ist nicht eindeutig.*

Definition Sei $\sigma \in S_n$. Ein **Fehlstand** von σ ist ein Indexpaar (i, j) von Elementen $i, j \in \{1, \ldots, n\}$, wenn $i < j$ und $\sigma(i) > \sigma(j)$. Wir definieren dann das **Signum** bzw. **Vorzeichen** von σ als

$$\operatorname{sign}(\sigma) := (-1)^s ,$$

wobei s gleich der Anzahl der Fehlstände ist.

Lemma 7.15. *Für jede Permutation $\sigma \in S_n$ gilt:*

$$\operatorname{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} .$$

Satz 7.16. (Multiplikativität des Signums) *Für alle $\sigma, \tau \in S_n$ gilt:*

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau) .$$

Korollar 7.17. *Für alle $\sigma \in S_n$ gilt:*

$$\operatorname{sign}\left(\sigma^{-1}\right) = \operatorname{sign}(\sigma) .$$

Korollar 7.18. *Für alle $\sigma \in S_n$ und für alle Transpositionen $\tau \in S_n$ gilt:*

$$\operatorname{sign}(\sigma \circ \tau) = -\operatorname{sign}(\sigma) .$$

Korollar 7.19. *Ist $\sigma \in S_n$ ein Produkt von k Transpositionen, so gilt:*

$$\operatorname{sign}(\sigma) = (-1)^k .$$

Definition Für jedes $n \in \mathbb{N}$ definieren wir die **alternierende Gruppe** als

$$A_n := \{\sigma \in S_n \mid \operatorname{sign}(\sigma) = 1\} \subseteq S_n .$$

Dass A_n tatsächlich eine Gruppe ist, werden wir im folgenden Satz sehen. Es ist auch klar, dass

$$S_n \setminus A_n = \{\sigma \in S_n \mid \operatorname{sign}(\sigma) = -1\} \subseteq S_n .$$

Somit ist

$$S_n = A_n \ \dot{\cup} \ (S_n \setminus A_n) \ .$$

Für ein beliebiges (aber festes) $\pi \in S_n$ führen wir auch die folgenden Mengen ein:

$$\pi A_n = \{ \pi \circ \sigma \mid \sigma \in A_n \} = \{ \pi \circ \sigma \mid \operatorname{sign}(\sigma) = 1 \} \ ,$$

$$A_n \pi = \{ \sigma \circ \pi \mid \sigma \in A_n \} = \{ \sigma \circ \pi \mid \operatorname{sign}(\sigma) = 1 \} \ ,$$

Satz 7.20. (Satz über die alternierende Gruppe) *Für jedes $n \in \mathbb{N}$ ist $(A_n, \circ)$ eine Gruppe, wobei $\circ$ die Hintereinanderausführung von Permutationen bezeichnet.*

(a) *Für alle $\tau \in A_n$ gilt:* $\tau A_n = A_n \tau = A_n$.

(b) *Für alle $\tau \in S_n \setminus A_n$ gilt:* $\tau A_n = A_n \tau = S_n \setminus A_n$. *Damit haben wir*

$$S_n = A_n \ \dot{\cup} \ \tau A_n = A_n \ \dot{\cup} \ A_n \tau \ .$$

(c) *Für jedes $n \geq 2$ gilt:* $|A_n| = \frac{1}{2} n!$.

Aus dem vorangehenden Satz benötigen wir für die Theorie hauptsächlich folgenden Aussagenteil aus (b): Für alle $\tau \in S_n \setminus A_n$ gilt:

$$A_n \tau = S_n \setminus A_n \ .$$

Damit leiten wir unmittelbar ab, dass

$$S_n = A_n \ \dot{\cup} \ A_n \tau \ .$$

Insbesondere gelten diese Aussagen für eine Transposition τ .

7.3 Existenz der Determinante, Leibnizformel

Wir werden nun die Existenz einer Determinante beweisen und auch eine explizite Formel für die Determinante angeben. Diese ist für $n \geq 4$ zur Berechnung eher vernachlässigbar. Wir können ja Determinanten effizient mittels elementaren Zeilenumformungen berechnen. Zudem gibt es noch die Laplaceschen Entwicklungen. Relevant ist die Leibnizformel für Berechnungen im Falle $n = 2, 3$. Folgender Satz gehört bereits zur Leibnizformel.

Satz 7.21. *Sei $A \in K^{n\times n}$ und sei* det $: K^{n\times n} \to K$ *eine Determinante. Bezeichnen $z_1, \dots, z_n \in K^{1\times n}$ die Zeilen von A, dann gilt für alle $\sigma \in S_n$:*

$$\det \underbrace{\begin{pmatrix} z_{\sigma(1)} \\ \vdots \\ z_{\sigma(k)} \\ \vdots \\ z_{\sigma(\ell)} \\ \vdots \\ z_{\sigma(n)} \end{pmatrix}}_{=:A_\sigma} = \operatorname{sign}(\sigma) \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_k \\ \vdots \\ z_\ell \\ \vdots \\ z_n \end{pmatrix} = \operatorname{sign}(\sigma) \cdot \det(A)$$

Insbesondere gilt für $A = E_n = \begin{pmatrix} e_1 & \dots & e_k & \dots & e_\ell & \dots e_n \end{pmatrix} \in K^{n\times n}$:

$$\det \begin{pmatrix} e^T_{\sigma(1)} \\ \vdots \\ e^T_{\sigma(k)} \\ \vdots \\ e^T_{\sigma(\ell)} \\ \vdots \\ e^T_{\sigma(n)} \end{pmatrix} = \operatorname{sign}(\sigma) \ .$$

Satz 7.22. (Hauptsatz für Determinanten) *Sei K ein Körper und sei $n \in \mathbb{N}$. Dann existiert eine (und damit genau eine) Determinante* det $: K^{n\times n} \to K$. *Sie wird durch die **Leibnizformel** beschrieben: Für $A = (a_{ij}) \in K^{n\times n}$ ist*

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \ .$$

Satz 7.23. *Sei $A \in K^{n \times n}$. Dann gilt:*

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

Satz 7.24. (Blockmatrizen) *Sei $A \in K^{n \times n}$ eine Blockmatrix mit zwei Blöcken in oberer Dreiecksform, also von der Form*

$$A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix} \ ,$$

wobei $A_1 \in K^{n_1 \times n_1}$, $A_2 \in K^{n_2 \times n_2}$, $C \in K^{n_1 \times n_2}$ und $n_1 + n_2 = n$. Dann gilt:

$$\det(A) = \det(A_1) \cdot \det(A_2) \ .$$

Korollar 7.25. *Sei $A \in K^{n \times n}$ eine Blockmatrix mit ℓ Blöcken in oberer Dreiecksform, also von der Form*

$$A = \begin{pmatrix} A_1 & * & * & * \\ 0 & A_2 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & \dots & A_\ell \end{pmatrix} \ ,$$

wobei $A_i \in K^{n_i \times n_i}$ für $i \in \{1, \dots, \ell\}$ und $\sum\limits_{i=1}^{\ell} n_i = n$. Dann gilt:

$$\det(A) = \prod_{i=1}^{\ell} \det(A_i) \ .$$

7.4 Laplace-Entwicklung und Cramer'sche Regel

Hier besprechen wir die rekursive Entwicklung der Determinante und eine Anwendung auf lineare Gleichungssysteme. Die Laplace-Entwicklung beschreibt auch die Determinante einer Matrix, liefert also dasselbe Ergebnis wie die von der Leibnizformel beschriebene Determinante.

Sei K ein Körper und sei $n \in \mathbb{N}$ mit $n \geq 2$. Dies ist keine große Einschränkung, da der Fall $n = 1$ für uns uninteressant ist. Sei $A \in K^{n \times n}$ eine beliebige Matrix. Für alle $i, j \in \{1, \ldots, n\}$ definieren wir die folgenden Matrizen:

$$A_{ij} := \begin{pmatrix} a_{11} & \ldots & a_{1,j-1} & 0 & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots\ldots & a_{n,j-1} & 0 & a_{n,j+1} & \ldots & a_{nn} \end{pmatrix} \in K^{n \times n},$$

$$A'_{ij} := \begin{pmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ a_{n1} & \ldots\ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{nn} \end{pmatrix} \in K^{(n-1)\times(n-1)}.$$

Also entsteht A'_{ij} durch Wegstreichen der i-ten Zeile und j-ten Spalte von A.

Satz 7.26. (Laplace'scher Entwicklungssatz) *Es seien* $A = (a_{ij}) \in K^{n\times n}$*, wobei* K *ein Körper ist, und* $n \in \mathbb{N}$*. Dann gilt:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \qquad \textit{für alle } i = 1, \ldots, n \ .$$

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \qquad \textit{für alle } j = 1, \ldots, n \ .$$

Die erste Formel bezeichnet man als die ***Laplace-Entwicklung von*** $\det(A)$ ***nach der*** i***-ten Zeile****, und die zweite Formel als die* ***Laplace-Entwicklung von*** $\det(A)$ ***nach der*** j***-ten Spalte****.*

Man kann diesen Satz direkt beweisen, indem man die Axiome (D1)-(D3) verifiziert, und zwar für für alle $i, j \in \{1, \ldots, n\}$. Es genügt, die Aussage für $j = 1$ mit vollständiger Induktion nach n zu beweisen. Die Verallgemeinerung folgt mittels Spaltenvertauschungen und der Aussage über Determinante von Transponierten. Wir beweisen die Aussage in der Vorlesung auf einem anderen Weg. Genauer werden wir sehen, dass dieser Satz eine unmittelbare Folgerung aus einem tieferen Resultat über Adjungierte ist.

Lemma 7.27. *Sei* $A = (a_{ij}) \in K^{n\times n}$*. In Spaltenschreibweise bedeutet dies*

$$A = \begin{pmatrix} s_1 & \ldots & s_j & \ldots & s_n \end{pmatrix} \quad \textit{wobei} \quad s_j = \begin{pmatrix} a_{1j} & a_{2j} & \ldots & a_{nj} \end{pmatrix}^T \in K^{n\times 1} \ .$$

(a) *Für alle* $i, j \in \{1, \ldots, n\}$ *gilt:*

$$\det(A_{ij}) = (-1)^{i+j} \det(A'_{ij}) \ .$$

(b) *Bezeichnet*

$$E_n = \begin{pmatrix} e_1 & \ldots & e_i & \ldots & e_n \end{pmatrix} = (\delta_{ij}) \in K^{n\times n} \ ,$$

wobei die Spaltenvektoren die Einheitsvektoren sind, also

$$e_i^T = \begin{pmatrix} \delta_{1i} & \ldots & \delta_{i-1,i} & \delta_{ii} & \delta_{i+1,i} & \ldots & \delta_{ni} \end{pmatrix}^T = \begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{pmatrix}^T \ ,$$

so gilt für alle $i, j \in \{1, \ldots, n\}$*:*

$$\det(A_{ij}) = \det \begin{pmatrix} s_1 & \ldots & s_{j-1} & e_i & s_{j+1} & \ldots & s_n \end{pmatrix} \ .$$

Definition Sei $A = (a_{ij}) \in K^{n\times n}$.

- Wir definieren den **Minor von** A **zum Indexpaar** $(i\,,\,j)$ bzw. den **Minor zum Koeffizienten** a_{ij} als
$$M_{ij} := \det(A'_{ij}) \;.$$
- Wir definieren den **Cofaktor von** A **zum Indexpaar** $(i\,,\,j)$ bzw. den **Cofaktor zum Koeffizienten** a_{ij} als
$$c_{ij} := \det(A_{ij}) = (-1)^{i+j}\det(A'_{ij}) = (-1)^{i+j} M_{ij} \;.$$
- Die Matrix $C := (c_{ij}) = ((-1)^{i+j} M_{ij}) \in K^{n\times n}$, also
$$C = \begin{pmatrix} M_{11} & -M_{12} & \dots & (-1)^{1+j}M_{1j} & \dots & (-1)^{1+n}M_{1n} \\ -M_{21} & M_{22} & \dots & (-1)^{2+j}M_{2j} & \dots & (-1)^{2+n}M_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (-1)^{i+1}M_{i1} & (-1)^{i+2}M_{i2} & \dots & (-1)^{i+j}M_{ij} & \dots & (-1)^{1+n}M_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (-1)^{n+1}M_{n1} & (-1)^{n+2}M_{n2} & \dots & (-1)^{n+j}M_{nj} & \dots & M_{nn} \end{pmatrix} ,$$
heißt die **Matrix der Cofaktoren von** A.
- Wir definieren die **adjungierte Matrix von** A bzw. die **zu** A **gehörige Komplementärmatrix** als
$$\mathrm{adj}(A) := C^T = (c_{ji}) = \big((-1)^{i+j} M_{ji}\big) \in K^{n\times n} \;.$$

Satz 7.28. *Sei $A = (a_{ij}) \in K^{n\times n}$ und sei $E_n = (\delta_{ij}) \in K^{n\times n}$. Dann gilt:*
$$A\cdot \mathrm{adj}(A) = \mathrm{adj}(A)\cdot A = \det(A)\cdot E_n \;.$$
Genauer gilt für alle $i\,,\,j \in \{1\,,\,\dots\,,\,n\}$:
$$\sum_{k=1}^{n} c_{ki}\,a_{kj} = \delta_{ij}\cdot\det(A) = \sum_{k=1}^{n} a_{ik}\,c_{jk} \;.$$
Ist A invertierbar, so gilt insbesondere:
$$A^{-1} = (\det(A))^{-1}\cdot \mathrm{adj}(A) \;.$$

Damit ist Satz 7.26 ein Spezialfall dieses Satzes, denn für $i = j$ sind die Formeln aus Satz 7.28 genau die Formeln für die Laplace-Entwicklung. Per Definition der c_{ij} und Lemma 7.27 ist nämlich
$$c_{ij} = \det(A_{ij}) = (-1)^{i+j}\det(A'_{ij}) \;.$$

Wir fassen die Formeln im folgenden Korollar zusammen.

Korollar 7.29. (Laplace'scher Entwicklungssatz') *Sei K ein Körper und sei $n \in \mathbb{N}$. Sei $A = (a_{ij}) \in K^{n\times n}$. Ist $C := (c_{ij})$ die Matrix der Cofaktoren von A, so gilt:*

$$\det(A) = \sum_{j=1}^{n} a_{ij} c_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \qquad \text{für alle } i = 1, \dots, n \; .$$

$$\det(A) = \sum_{i=1}^{n} a_{ij} c_{ij} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A'_{ij}) \qquad \text{für alle } j = 1, \dots, n \; .$$

Bemerkung 7.30. *Die Adjungierte einer Matrix besitzt folgende Vorteile:*

(a) *Die Adjungierte einer Matrix ist im Allgemeinen nur von theoretischem Interesse. Für globale Aussagen über eine Matrix (z.B. mit Parametern) kann sie von Nutzen sein. Dies ist insbesondere der Fall, wenn symbolisches Rechnen erwünscht ist.*

(b) *Falls A nur ganzzahlige Einträge (oder Polynomeinträge) hat, so auch die Adjungierte. Also braucht man in solchen Fällen keine Brüche.*

Satz 7.31. (Cramer'sche Regel) *Sei $b \in K^{n\times 1}$ und sei*

$$A = \begin{pmatrix} s_1 & \dots & s_j & \dots & s_n \end{pmatrix} \in \mathrm{GL}(n, K) \; .$$

Da A invertierbar ist, wissen wir, dass das lineare Gleichungssystem $A x = b$ eine eindeutige Lösung hat, nämlich

$$x' := \begin{pmatrix} x_1 & \dots & x_i & \dots & x_n \end{pmatrix}^T = A^{-1} b \; .$$

Dann gilt für alle $i \in \{1, \dots, n\}$: Die i-te Komponente x_i von x' berechnet sich als

$$x_i = (\det(A))^{-1} \cdot \det\begin{pmatrix} s_1 & \dots & s_{i-1} & b & s_{i+1} & \dots & s_n \end{pmatrix} \; .$$

Einschub: Der Quotientenkörper von $K[\,t\,]$

Wir fügen hier die Konstruktion des Quotientenkörpers von $K[\,t\,]$ bzw. des rationalen Funktionenkörpers ein. Die allgemeinere Situation wird in der Algebra I behandelt, wo der Begriff des Quotientenkörper eines Integritätsringes konstruiert wird.

Sei K ein Körper. Dann haben die Ringe $\mathbb{Z}$ und $K[\,t\,]$ sehr viel gemeinsam. Beides sind zum Beispiel nullteilerfreie, kommutative Ringe. Wir wissen von der Schule, dass wir $\mathbb{Z}$ in die rationalen Zahlen einbetten können und dass $\mathbb{Q}$ ein Körper ist, der damit $\mathbb{Z}$ enthält. Elemente aus $\mathbb{Q}$ sind Brüche $\frac{a}{b}$, wobei $a, b \in \mathbb{Z}$ und $b \neq 0$. Die Darstellung ist eindeutig, wenn man gemeinsame Teiler von Zähler und Nenner 'wegkürzt'. Also ist $\frac{a}{b} = \frac{a \cdot c}{b \cdot c}$ für alle $c \in \mathbb{Z} \setminus \{\,0\,\}$. Wir konstruieren nun aus $K[\,t\,]$ einen Körper $K(\,t\,)$, indem wir ähnlich wie bei $\mathbb{Z}$ und $\mathbb{Q}$ vorgehen. Auf $K[\,t\,] \times (\,K[\,t\,] \setminus \{\,0\,\}\,)$ definieren wir eine Relation S wie folgt:

$$((\,f\,,\,g\,)\,,\,(\,f_1\,,\,g_1\,)) \in S \quad :\Leftrightarrow \quad f\,g_1 = g\,f_1 \ .$$

Man verifiziert, dass S eine Äquivalenzrelation ist, d.h. reflexiv, symmetrisch und transitiv ist. Wir schreiben $\frac{f}{g}$ für die Äquivalenzklasse zu $(\,f\,,\,g\,) \in K[\,t\,] \times (\,K[\,t\,] \setminus \{\,0\,\}\,)$. Damit enthält $\frac{f}{g}$ alle zu $(\,f\,,\,g\,)$ äquivalenten Elemente. Wir definieren $K(\,t\,)$ als die Menge aller Äquivalenzklassen bzgl. der Relation S, d.h.

$$K(\,t\,) := \left\{ \frac{f}{g} \;\middle|\; f\,,\,g \in K[t],\, g \neq 0 \right\} \ .$$

Wir führen auch zwei Operationen ein, die man von der Bruchrechnung kennt:

- Addition: $\dfrac{f}{g} + \dfrac{f_1}{g_1} := \dfrac{f\,g_1 + f_1\,g}{g\,g_1}$.
- Multiplikation: $\dfrac{f}{g} \cdot \dfrac{f_1}{g_1} := \dfrac{f\,f_1}{g\,g_1}$.

Dann ist $(\,K(\,t\,)\,,\,+\,,\,\cdot\,)$ ein Körper, genannt der **Körper der rationalen Funktionen** oder der **Quotientenkörper von** $K[\,t\,]$. Identifiziert man $f \in K[\,t\,]$ mit $\frac{f}{1} \in K(\,t\,)$, so ist damit $K[\,t\,] \subseteq K(\,t\,)$, und die Operationen von $K[\,t\,]$ stimmen mit denen von $K(\,t\,)$ überein. Außerdem lässt sich $\frac{f}{g}$ unter dieser Identifikation als $f \cdot g^{-1}$ schreiben.

8 Eigenwerttheorie

Die Eigenwerttheorie und die damit verbundenen Probleme spielen in den gesamten Naturwissenschaften ein Rolle. Wir werden die relevanten Fragestellungen für Matrizen über Körpern in der Linearen Algebra betrachten und die Theorie in der Algebra I vertiefen. Sei K ein Körper und sei $n \in \mathbb{N}$. Uns werden die folgenden, prinzipiellen Probleme interessieren:

(A) Eigenwertproblem: Gegeben sei eine Matrix $A \in K^{n\times n}$; gibt es einen Vektor $v \in K^{n\times 1} \setminus \{0\}$ und einen Skalar $\lambda \in K$, so dass

$$A\,v = \lambda\,v \ ?$$

(i) Ist $F_A : K^{n\times 1} \to K^{n\times 1}$ die Standardinterpretation von A, so ist das Problem aquivalent zur folgenden Frage: Gibt es einen Vektor $v \in K^{n\times 1} \setminus \{0\}$ und einen Skalar $\lambda \in K$, so dass

$$F_A(v) = \lambda\,v \ ?$$

(ii) Bestimmen Sie alle Skalare $\lambda \in K$, so dass $F_A(v) = A\,v = \lambda\,v$.

(iii) Sei $\lambda \in K$. Bestimmen Sie $\mathrm{NR}(\lambda E_n - A)$. D.h. bestimmen Sie alle $v \in K^{n\times 1}$ mit

$$F_A(v) = \lambda\,v \ .$$

(iv) Angenommen, $K^{n\times 1}$ besäße eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ aus verschiedenen Vektoren $v_1, \ldots, v_n \in K^{n\times 1}$ mit der Eigenschaft, dass es $\lambda_1, \ldots, \lambda_n \in K$ gibt, so dass

$$F_A(v_i) = A\,v_i = \lambda_i\,v_i \qquad \text{für alle } \ i = 1, \ldots, n \ .$$

Dann wäre F_A besonders einfach zu berechnen und die F_A darstellende Matrix bzgl. der Basis $\mathcal{B}$ von besonders einfacher Form.

Sei $v \in K^{n\times 1}$ beliebig. Dann existieren $\mu_1, \ldots, \mu_n \in K$, so dass

$$v = \sum_{i=1}^{n} \mu_i\,v_i \quad \Rightarrow \quad F_A(v) = \sum_{i=1}^{n} \mu_i\,F_A(v_i) = \sum_{i=1}^{n} \mu_i\,\lambda_i\,v_i \quad \text{und}$$

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{B}}(F_A) &= \Big(I_{\mathcal{B}}\,(F_A(v_1)) \ \ \ldots \ \ I_{\mathcal{B}}\,(F_A(v_n))\Big) \\ &= \Big(I_{\mathcal{B}}\,(\lambda_1 v_1) \ \ \ldots \ \ I_{\mathcal{B}}\,(\lambda_n v_n)\Big) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} . \end{aligned}$$

(B) Diagonalisierung: Gegeben sei eine Matrix $A \in K^{n\times n}$; ist A ähnlich zu einer Diagonalmatrix $D \in K^{n\times n}$? D.h. existieren eine Matrix $P \in \mathrm{GL}(n, K)$ sowie Skalare $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$P^{-1} A P = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} =: D \ ?$$

Falls ja, so hätten wir folgende Beobachtungen:

(i) Es existiert dann eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von $K^{n\times 1}$, so dass

$$M_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} .$$

(ii) Es gilt

$$P^{-1} A P = D \quad \Leftrightarrow \quad A = P D P^{-1} .$$

Folglich rechnen wir

$$A^2 = (P D P^{-1})(P D P^{-1}) = P D^2 P^{-1} .$$

Durch Iteration erhalten wir für alle $k \in N$:

$$A^k = P D^k P^{-1}$$

und

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

8.1 Eigenwerte und Eigenvektoren von Matrizen

Definition Sei $A \in K^{n\times n}$. Ein Skalar $\lambda \in K$ heißt ein **Eigenwert** von A, wenn es einen Vektor $v \in K^{n\times 1} \setminus \{0\}$ gibt, so dass

$$Av = \lambda v \ .$$

Jeder solche Vektor $v \in K^{n\times 1}\setminus\{0\}$ mit $Av = \lambda v$ heißt ein **Eigenvektor von A zum Eigenwert λ**.

Definition Sei $A \in K^{n\times n}$ und $\lambda \in K$. Dann heißt

$$\mathrm{Eig}(A, \lambda) := \left\{ x \in K^{n\times 1} \mid Ax = \lambda x \right\} \subseteq K^{n\times 1}$$

der **Eigenraum von A bzgl. λ**. Dessen Dimension

$$\dim_K \mathrm{Eig}(A, \lambda)$$

nennt man die **geometrische Vielfachheit von λ** bzgl. A.

Beobachtungen: Sei $A \in K^{n\times n}$ und $\lambda \in K$.
1. Fall: $\lambda \in K$ ist kein Eigenwert von A. Dann ist offensichtlich

$$\mathrm{Eig}(A, \lambda) = \{0\} \quad \text{und} \quad \dim_K \mathrm{Eig}(A, \lambda) = 0 \ .$$

2. Fall: $\lambda \in K$ ist ein Eigenwert von A. Dann existiert ein $v \in K^{n\times 1} \setminus \{0\}$, so dass

$$0 = \lambda v - Av = (\lambda E_n - A)v \quad \text{bzw.} \quad 0 = \lambda v - F_A(v) = (\lambda\, \mathrm{id}_{K^{n\times 1}} - F_A)(v) \ .$$

$$\Rightarrow \quad \mathrm{Eig}(A, \lambda) = \mathrm{Lös}(\lambda E_n - A, 0) = \mathrm{NR}(\lambda E_n - A) = \mathrm{Kern}(\lambda\, \mathrm{id}_{K^{n\times 1}} - F_A) \ .$$

Zudem ist nach der Dimensionsformel und der Definition der Nullität

$$1 \le \dim_K \mathrm{Eig}(A, \lambda) = \nu(\lambda E_n - A) = n - \mathrm{rang}(\lambda E_n - A) \ .$$

Dies ist äquivalent zu

$$\mathrm{rang}(\lambda E_n - A) \le n - 1 < n \ .$$

Also muss in diesem Fall die Matrix $\lambda E_n - A$ auf jeden Fall singulär sein.

Bemerkung 8.1. *Sei $A \in K^{n\times n}$ und sei $\lambda \in K$. Dann gilt:*

(a) $\mathrm{Eig}(A, \lambda)$ *ist tatsächlich ein Unterverktorraum von $K^{n\times 1}$.*

(b) *λ ist genau dann kein Eigenwert von A, wenn* $\mathrm{Eig}(A, \lambda) = \{0\}$.

λ ist genau dann ein Eigenwert von A, wenn $\dim_K \mathrm{Eig}(A, \lambda) \ge 1$.

(c) $\mathrm{Eig}(A, \lambda) = \mathrm{NR}(\lambda E_n - A) = \mathrm{Kern}(\lambda\, \mathrm{id}_{K^{n\times 1}} - F_A)$.

(d) *Nicht jede Matrix muss einen Eigenwert besitzen.*

Herleitung: Sei $A = (a_{ij}) \in K^{n\times n}$ und $\lambda \in K$ ein Eigenwert von A. Dann existiert ein $v \in K^{n\times 1} \setminus \{0\}$, so dass

$$F_A(v) = Av = \lambda v \ .$$

$$\Leftrightarrow \quad \dim_K \mathrm{Eig}(A, \lambda) \geq 1 \ .$$

$$\Leftrightarrow \quad \mathrm{rang}\,(\lambda E_n - A) < n \ .$$

$$\Leftrightarrow \quad \det(\lambda E_n - A) = 0 \ ,$$

nach Satz 7.7. Folglich spielt die Determinante von $\lambda E_n - A \in K^{n\times n}$ für alle $\lambda \in K$ eine besondere Rolle. Betrachtet man λ als eine Unbestimmte, so liefert das Polynom $\det(\lambda E_n - A)$ in λ alle Eigenwerte. Dass dies tatsächlich ein **Polynom** in λ ist, folgt aus der Leibnizformel. Ersetzen wir λ durch t, so erhalten wir das charakteristische Polynom von A.

Wir betrachten die folgende wichtige Matrix in der Unbestimmten t über K

$$t E_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{pmatrix} \in (K[t])^{n\times n} \ .$$

Festlegung:

- Da $K[t]$ kein Körper sondern nur ein nullteilerfreier, kommutativer Ring ist, sind die Ergebnisse über Determinanten nicht sofort anwendbar. Wir benötigen also eigentlich die Theorie der Determinanten von Matrizen über kommutativen Ringen.

- Man kann Determinanten von Matrizen über kommutativen Ringen vollkommen allgemein über die Leibnizformel definieren und erhält analoge Resultate.

- Noch eleganter ist die Betrachtungsweise über den Quotientenkörper $K(t)$ von $K[t]$. Dieser ist ein Körper, der alle Einträge von $t E_n - A$ enthält. Aus der Leibnizformel geht sogar hervor, dass bei der Determinanten von $t E_n - A$ keine "Nenner", d.h. Invertierungen, auftreten können.

Einschub: Polynome und Linearfaktoren

Sei K ein Körper und sei $K[t]$ der Polynomring in der Unbestimmten t über K. Dann wissen wir aus Proposition 3.3 und Proposition 3.4, dass Folgendes gilt:

- Für alle $f, g \in K[t]$ ist $\deg(fg) = \deg(f) + \deg(g)$ und $K[t]$ ist nullteilerfrei.
- Die Einheiten von $K[t]$ sind gerade die Einheiten von K, d.h. $(K[t])^* = K^*$.
- Ein normiertes Polynom vom Grad n ist von der Form $f = a_0 + \ldots + a_{n-1} t^{n-1} + t^n$.
- Für alle $\alpha \in K$ und $f = \sum_{j=0}^{n} a_j t^j \in K[t]$ ist
$$f(\alpha) = \sum_{j=0}^{n} a_j \alpha^j = a_0 + a_1 \alpha + \ldots + a_n \alpha^n \quad \in K \ .$$
- Gilt $f(\alpha) = 0$ für ein $\alpha \in K$, so heißt α eine **Nullstelle** von f.
- Division mit Rest in $K[t]$ ist auf eindeutige Art und Weise möglich. Für $f, g \in K[t]$ mit $g \neq 0$ gibt es eindeutige Polynome $q, r \in K[t]$, so dass
$$f = qg + r \qquad \text{und} \qquad \deg(r) < \deg(g) \ .$$
  Wir sagen g **teilt** f, geschrieben $g \mid f$, wenn in dieser Zerlegung $r = 0$ gilt.

Proposition 8.2. *Sei $f \in K[t]$ und sei $\alpha \in K$ eine Nullstelle von f. Dann existiert ein eindeutiges Polynom $q \in K[t]$ vom Grad* $\deg q = \deg(f) - 1$*, so dass*
$$f = q \cdot (t - \alpha) \ .$$

Korollar 8.3. *Jedes Polynom $f \in K[t]$ vom Grad* $\deg(f) = n \geq 1$ *besitzt höchstens n Nullstellen in K.*

Wir sagen, $f \in K[t]$ vom Grad $\deg(f) = n \geq 1$ **zerfällt über K in Linearfaktoren**, wenn $\beta_1, \ldots, \beta_n \in K$ existieren, welche nicht notwendigerweise verschieden sind, so dass
$$f = \prod_{i=1}^{n} (t - \beta_i) \ .$$
Sortiert man nach gleichen Nullstellen, so kann man in dem Fall schreiben
$$f = \prod_{i=1}^{m} (t - \lambda_i)^{r_i} \ ,$$
wobei $\lambda_1, \ldots, \lambda_m \in K$ paarweise verschieden sind und $r_i \in \mathbb{N}$. Es gilt natürlich, dass
$$\{\beta_1, \ldots, \beta_n\} = \{\lambda_1, \ldots, \lambda_m\} \ .$$
Dabei bezeichnet somit r_i die **Vielfachheit der Nullstelle** λ_i für alle $i = 1, \ldots, n$.

Für einen beliebigen Körper K weiß man im Allgemeinen jedoch nur, dass ein Polynom $f \in K[t]$ vom Grad $\deg(f) = n \geq 1$ wie folgt zerfällt: Sind $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Nullstellen von f, so existiert ein Polynom $q \in K[t]$, so dass

$$f = q \cdot \prod_{i=1}^{m} (t - \lambda_i)^{r_i}$$

und $q(\lambda) \neq 0$ für alle $\lambda \in K$, insbesondere also $q(\lambda_i) \neq 0$ für alle $i \in \{1, \ldots, m\}$.

Eine Ausnahme ist zum Beispiel der Fall $K = \mathbb{C}$. Das liegt an dem

Fundamentalsatz der Algebra: Jedes Polynom $f \in \mathbb{C}[t]$ vom Grad $\deg(f) \geq 1$ besitzt eine Nullstelle in $\mathbb{C}$. Insbesondere erhält man durch Iteration: Jedes Polynom $f \in \mathbb{C}[t]$ vom Grad $\deg(f) \geq 1$ zerfällt über $\mathbb{C}$ in Linearfaktoren.

Definition Sei $A \in K^{n \times n}$. Dann heißt

- $H_A := tE_n - A \in (K(t))^{n \times n}$ die **charakteristische Matrix von** A.
- $h_A := \det(tE_n - A) \in K(t)$ das **charakteristische Polynom von** A.

Satz 8.4. *Sei $A = (a_{ij}) \in K^{n \times n}$. Dann gilt:*

(a) *$\lambda \in K$ ist genau dann ein Eigenwert von A, wenn $h_A(\lambda) = 0$.*

(b) *$h_A \in K[t]$ ist ein normiertes Polynom vom Grad n. Es gilt:*

$$h_A = c_0 + \ldots + c_{n-1} t^{n-1} + t^n ,$$

*wobei wir mit der **Spur von** A, $\mathrm{Spur}(A) = \sum_{i=1}^{n} a_{ii}$, haben*

(i) $c_{n-1} = -\mathrm{Spur}(A)$.

(ii) $c_0 = (-1)^n \det(A)$.

Korollar 8.5. *Sei $A = (a_{ij}) \in K^{n \times n}$. Dann gilt:*

(a) *A hat höchstens n Eigenwerte.*

(b) *A ist genau dann invertierbar, wenn $\lambda = 0$ kein Eigenwert von A ist.*

(c) $h_{A^T} = h_A$.

(d) *Ist A eine obere oder untere Dreiecksmatrix, dann hat A die folgenden n Eigenwerte:*

$$\lambda_1 = a_{11} , \lambda_2 = a_{22}, \ldots , \lambda_n = a_{nn} .$$

Satz 8.6. *Sind $A, B \in K^{n \times n}$ ähnlich, so gilt:*

$$h_A = h_B .$$

8.2 Eigenwerte und Eigenvektoren von Endomorphismen

Sei V ein K-Vektorraum über einem Körper K. Wir betrachten nun

$$\mathrm{End}_K(V) = \{ F : V \to V \mid F \text{ lineare Abbildung} \}$$

Die Definition des vorherigen Abschnitts übertragen sich wie folgt auf Endomorphismen.

Definition Sei V ein K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Ein Skalar $\lambda \in K$ heißt ein **Eigenwert** von F, wenn es einen Vektor $v \in V \setminus \{0\}$, so dass

$$F(v) = \lambda v \ .$$

Jeder solche Vektor $v \in K^{n\times 1} \setminus \{0\}$ mit $F(v) = \lambda v$ heißt ein **Eigenvektor von** F **zum Eigenwert** λ.

Definition Sei V ein K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Für $\lambda \in K$ heißt

$$\mathrm{Eig}(F, \lambda) := \{ x \in V \mid F(x) = \lambda x \} \subseteq V$$

der **Eigenraum von** F **bzgl.** λ. Dessen Dimension $\dim_K \mathrm{Eig}(F, \lambda)$ nennt man auch die **geometrische Vielfachheit von** λ bzgl. F.

Bemerkung 8.7. *Sei V ein K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Dann gilt für alle $\lambda \in K$:*

(a) $\mathrm{Eig}(F, \lambda)$ *ist tatsächlich ein Untervektorraum von V.*

(b) *Ist F_A die Standardinterpretation von A, so gilt*

$$\mathrm{Eig}(A, \lambda) = \mathrm{Eig}(F_A, \lambda) \ .$$

Zudem ist λ genau dann ein Eigenwert von A, wenn λ ein Eigenwert von F_A ist.

(c) *λ ist genau dann kein Eigenwert von F, wenn* $\mathrm{Eig}(F, \lambda) = \{0\}$.
λ ist genau dann ein Eigenwert von F, wenn $\dim_K \mathrm{Eig}(F, \lambda) \geq 1$.

(d) $\mathrm{Eig}(F, \lambda) = \mathrm{Kern}(\lambda\, \mathrm{id}_V - F)$.

Satz 8.8. *Sei V ein n-dimensionaler K-Vektorraum über einem Körper K mit Basis $\mathcal{B} = (v_1, \dots, v_n)$ und sei $F \in \mathrm{End}_K(V)$. Es gilt: $\lambda \in K$ ist genau dann ein Eigenwert von F, wenn λ ein Eigenwert der F darstellenden Matrix*

$$A := \Psi(F) = M_{\mathcal{B}}^{\mathcal{B}}(F)$$

bzgl. $\mathcal{B}$ ist. Insbesondere ist dann

$$\mathrm{Eig}(F, \lambda) = \psi_{\mathcal{B}}(\mathrm{Eig}(A, \lambda)) = I_{\mathcal{B}}^{-1}(\mathrm{Eig}(A, \lambda)) \ .$$

Satz 8.9. *Sei V ein K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Sind $\lambda_1, \dots, \lambda_m$ paarweise verschiedene Eigenwerte von F und $v_1, \dots, v_m$ Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_m$, dann ist das System $(v_1, \dots, v_m)$ K-linear unabhängig.*

Korollar 8.10. *Sei $A \in K^{n\times n}$. Sind $\lambda_1, \dots, \lambda_m$ paarweise verschiedene Eigenwerte von A und $v_1, \dots, v_m$ zugehörige Eigenvektoren, dann ist das System $(v_1, \dots, v_m)$ linear unabhängig über K.*

Korollar 8.11. *Sei V ein n-dimensionaler K-Vektorraum über einem Körper K. Dann hat jeder Endomorphismus $F \in \mathrm{End}_K(V)$ höchstens n verschiedene Eigenwerte.*

Korollar 8.12. *Sei V ein n-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Besitzt F genau n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$ mit zugehörigen Eigenvektoren $v_1, \dots, v_n$, so ist $\mathcal{B} = (v_1, \dots, v_n)$ eine K-Basis von V und*

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} .$$

Definition Sei V ein n-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Das **charakteristische Polynom von** F wird definiert als

$$h_F := h_A = \det(t\,E_n - A) \in K[t] ,$$

wobei $A = M_{\mathcal{B}}^{\mathcal{B}}(F)$ die F darstellende Matrix bzgl. einer beliebigen Basis $\mathcal{B}$ von V ist. Dabei ist h_F unabhängig von der Wahl der Basis von V.

Bemerkung 8.13. *Sei V ein n-dimensionaler Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$.*

(a) *Für das charakteristische Polynom h_F gelten genau dieselben Eigenschaften wie für das charakteristische Polynom h_A von $A = M_{\mathcal{B}}^{\mathcal{B}}(F)$ in Satz 8.4.*

(b) *Die Nullstellen von h_F sind genau die Eigenwerte von F. Genauer gilt: $\lambda \in K$ ist genau dann ein Eigenwert von F, wenn λ ein Eigenwert von $A = M_{\mathcal{B}}^{\mathcal{B}}(F)$ ist. Dies ist genau dann der Fall, wenn $h_F(\lambda) = 0$.*

Definition Sei V ein n-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Für das charakteristische Polynom h_F und beliebiges $\lambda \in K$, existiert ein $q \in K[t]$ und ein $r \in \mathbb{N}_0$, so dass

$$h_F = (t - \lambda)^r \cdot q \ ,$$

wobei $q(\lambda) \neq 0$. Dann nennen wir

$$\mu(F, \lambda) := r$$

die **algebraische Vielfachheit von** λ.

Allgemein für einen beliebigen Körper K wissen wir, dass h_F als Polynom vom Grad n wie folgt zerfällt:

$$h_F = q \cdot \prod_{i=1}^{m} (t - \lambda_i)^{r_i} \ ,$$

wobei $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Eigenwerte von F, also die Nullstellen von h_F, sind. Dabei ist $q \in K[t]$ mit $q(\lambda_i) \neq 0$ für alle $i \in \{1, \ldots, m\}$. Zudem ist

$$\deg(h_F) = \deg(q) + \sum_{i=1}^{m} r_i \ .$$

In unserer Definition gilt $r_i = \mu(F, \lambda_i)$ und damit

$$h_F = q \cdot \prod_{i=1}^{m} (t - \lambda_i)^{\mu(F, \lambda_i)} \ .$$

Zudem setzen wir ein und erhalten

$$n = \deg(h_F) = \deg(q) + \sum_{i=1}^{m} \mu(F, \lambda_i) \ .$$

Schließlich, wenn $\lambda \in K$ kein Eigenwert von F ist, folgt automatisch, dass

$$\mu(F, \lambda_i) = 0 \ .$$

Wir kennen auch die **geometrische Vielfachheit** eines beliebigen $\lambda \in K$, nämlich

$$n_\lambda := \dim_K \mathrm{Eig}(F, \lambda) \ .$$

Satz 8.14. *Sei V ein n-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Für jeden Eigenwert λ von F ist die geometrische Vielfachheit kleiner oder gleich der algebraischen Vielfachheit von λ. D.h., es gilt:*

$$1 \leq \dim_K \mathrm{Eig}(F, \lambda) \leq \mu(F, \lambda) \ .$$

8.3 Diagonalisierung und Trigonalisierung

Wir betrachten in diesem Abschnitt Endomorphismen F eines endlich-dimensionalen K-Vektorraumes V über einem Körper K, für die Basen $\mathcal{B}$ von V existieren, so dass $M_{\mathcal{B}}^{\mathcal{B}}(F)$ Diagonalmatrizen oder obere Dreiecksmatrizen sind.

Definition Eine Matrix $D = (d_{ij}) \in K^{n\times n}$ heißt **Diagonalmatrix** oder **in Diagonalgestalt**, wenn $d_{ij} = 0$ für alle $i, j \in \{1, \ldots, n\}$ mit $i \neq j$. D.h. D ist von der folgenden Form

$$D = \begin{pmatrix} d_{11} & 0 & \ldots & 0 \\ 0 & d_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_{nn} \end{pmatrix} \in K^{n\times n} .$$

Definition Sei K ein Körper und sei $n \in \mathbb{N}$.

(a) Eine Matrix $A \in K^{n\times n}$ heißt **diagonalisierbar**, wenn A zu einer Diagonalmatrix ähnlich ist, d.h. wenn es eine invertierbare Matrix $P \in \mathrm{GL}(n, K)$ und eine Diagonalmatrix $D \in K^{n\times n}$ gibt, so dass

$$P^{-1} A P = D .$$

(b) Sei V ein n-dimensionaler K-Vektorraum. Ein Endomorphismus $F \in \mathrm{End}_K(V)$ heißt **diagonalisierbar**, wenn es eine Basis $\mathcal{B}$ von V gibt, so dass $M_{\mathcal{B}}^{\mathcal{B}}(F)$ eine Diagonalmatrix ist.

Satz 8.15. *Sei V ein endlich-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:*

(a) *F ist diagonalisierbar.*

(b) *V besitzt eine Basis aus Eigenvektoren von F.*

Korollar 8.16. *Sei V ein n-dimensionaler K-Vektorraum über einem Körper K. Besitzt $F \in \mathrm{End}_K(V)$ genau n verschiedene Eigenwerte, so ist F diagonalisierbar.*

Korollar 8.17. *Sei V ein n-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Ist F diagonalisierbar, so zerfällt das charakteristische Polynom h_F von F über K in Linearfaktoren.*

Ist F diagonalisierbar, so muss $F \in \mathrm{End}_K(V)$ nicht genau n verschiedene Eigenwerte besitzen. Der folgende Satz gibt ein Kriterium dazu an. Wir haben auch bereits Gegenbeispiele dafür gesehen. Zudem ist leider zu beachten, dass nicht jeder Endomorphismus diagonalisierbar ist.

Satz 8.18. *Sei V ein endlich-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:*

(a) *F ist diagonalisierbar.*

(b)

(i) *h_F zerfällt über K in Linearfaktoren, also*

$$h_F = \prod_{i=1}^{m} (t - \lambda_i)^{r_i} \qquad ,$$

wobei $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Eigenwerte von F sind.

(ii) *Die geometrische Vielfachheit jedes Eigenwertes stimmt mit seiner algebraischen Vielfachheit überein, also*

$$\dim_K \mathrm{Eig}(F, \lambda_i) = \mu(F, \lambda_i) \qquad \text{für alle } i = 1, \ldots, m \ .$$

Definition

(a) $A \in K^{n \times n}$ heißt **trigonalisierbar**, wenn A zu einer oberen Dreiecksmatrix ähnlich ist, d.h. wenn es eine Matrix $P \in \mathrm{GL}(n, K)$ gibt, so dass $P^{-1} A P$ eine obere Dreiecksmatrix ist.

(b) Sei V ein n-dimensionaler K-Vektorraum über einem Körper K. $F \in \mathrm{End}_K(V)$ heißt **trigonalisierbar**, wenn es eine Basis $\mathcal{B}$ von V gibt, so dass $M_{\mathcal{B}}^{\mathcal{B}}(F)$ eine obere Dreiecksmatrix ist.

Satz 8.19. *Sei V ein endlich-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:*

(a) *F ist trigonalisierbar.*

(b) *h_F zerfällt über K in Linearfaktoren.*

Wir können nun sofort einen Spezialfall des Satzes von Cayley-Hamilton, nämlich für $K = \mathbb{C}$ beweisen. Dazu definieren wir den Einsetzungshomomorphismus

$$\begin{array}{rrcl} \psi_F : & K[t] & \to & \mathrm{End}_K(V) \\ & p = \sum_{i=0}^{k} a_i t^i & \mapsto & \psi_F(p) := p(F) = \sum_{i=0}^{k} a_i F^i \ . \end{array}$$

Da $\mathrm{End}_K(V)$ ein Ring ist, ist $\psi_F(p)$ eine K-lineare Abbildung ist, wobei

$$\begin{array}{rrcl} p(F) : & V & \to & V \\ & v & \mapsto & p(F)(v) := \sum_{i=0}^{k} a_i F^i(v) \ . \end{array}$$

Satz 8.20. (Satz von Cayley-Hamilton über $\mathbb{C}$) Sei V ein endlich-dimensionaler K-Vektorraum über einem Körper K und sei $F \in \mathrm{End}_K(V)$. Dann gilt für das charakteristische Polynom h_F von F, dass $h_F(F)$ die Nullabbildung von V ist, d.h. dass

$$h_F(F) = 0 \ .$$

9 Vektorräume mit Skalarprodukt

In den bekannten Standardsituationen, wo der zugrundeliegende Körper $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $V = K^{n\times 1}$ ist, weiß man viel mehr als nur die Eigenschaften der Vektorräume. Man hat auch Abstände, Normen, Metriken, Längenmaße und andere geometrisch wichtige Hilfsmittel. Diese spielen zum Beispiel in der analytischen Geometrie eine Rolle und werden auch in vielen Bereichen der Analysis angewendet.

9.1 Euklidische Vektorräume

In diesem Abschnitt beschäftigen wir uns ausschließlich mit dem Fall, dass $K = \mathbb{R}$ und V ein reeller Vektorraum, also ein $\mathbb{R}$-Vektorraum, ist.

Definition Sei V ein $\mathbb{R}$-Vektorraum. Ein **reelles Skalarprodukt** in V ist eine Abbildung

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}\ ,$$

die den folgenden Axiomen genügt:

(B1) $\langle\ ,\ \rangle$ ist **linear** in der ersten Komponente:

(B1a) $\langle v + v', w\rangle = \langle v, w\rangle + \langle v', w\rangle$ für alle $v, v', w \in V$.

(B1b) $\langle \lambda v, w\rangle = \lambda\langle v, w\rangle$ für alle $v, w \in V$ und $\lambda \in \mathbb{R}$.

(S) $\langle\ ,\ \rangle$ ist **symmetrisch**: $\langle v, w\rangle = \langle w, v\rangle$ für alle $v, w \in V$.

(PD) $\langle\ ,\ \rangle$ ist **positiv definit**: $\langle v, v\rangle > 0$ für alle $v \in V \setminus \{0\}$.

Definition Wir nennen einen $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt $\langle\ ,\ \rangle$ einen **euklidischen Vektorraum**. Manchmal schreiben wir auch $(V, \langle\ ,\ \rangle)$.

Bemerkung 9.1. *Sei V ein $\mathbb{R}$-Vektorraum. Dann ist* **(B1)** *äquivalent zu*

(B1') *Für alle $v, v', w \in V$ und $\lambda, \lambda' \in \mathbb{R}$ gilt:*

$$\langle \lambda v + \lambda' v', w\rangle = \lambda\langle v, w\rangle + \lambda'\langle v', w\rangle\ .$$

Ist **(B1)** *erfüllt, so gilt für alle $v_1, \ldots, v_\ell, w \in V$ und $\lambda_1, \ldots, \lambda_\ell \in \mathbb{R}$:*

$$\left\langle \sum_{i=1}^{\ell} \lambda_i v_i, w \right\rangle = \sum_{i=1}^{\ell} \lambda_i \langle v_i, w\rangle\ .$$

Proposition 9.2. (Elementare Eigenschaften von reellen Skalarprodukten) Sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann gilt:

(a) Es gilt auch das Axiom

(B2) $\langle\ ,\ \rangle$ ist **linear in der zweiten Komponente**:

(B2a) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$ für alle $v, w, w' \in V$.

(B2b) $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$ für alle $v, w \in V$ und $\lambda \in \mathbb{R}$.

Dies ist wiederum äquivalent zu

(B2') $\langle v, \lambda w + \lambda' w' \rangle = \lambda \langle v, w \rangle + \lambda' \langle v, w' \rangle$ für alle $v, w, w' \in V$ und $\lambda, \lambda' \in \mathbb{R}$.

Man sagt in dem Fall auch, das Skalarprodukt sei **bilinear**.

(b) Für alle $v \in V$ gilt: $\langle v, v \rangle = 0 \quad \Leftrightarrow \quad v = 0$.

(c) Seien $w, w' \in V$. Falls $\langle v, w \rangle = \langle v, w' \rangle$ für alle $v \in V$, so gilt $w = w'$.

(d) Für alle $v_1, \ldots, v_\ell, w_1, \ldots, w_n \in V$ und $\lambda_1, \ldots, \lambda_\ell, \mu_1, \ldots, \mu_n \in \mathbb{R}$ gilt:

$$\left\langle \sum_{i=1}^{\ell} \lambda_i v_i, \sum_{j=1}^{n} \mu_j w_j \right\rangle = \sum_{i=1}^{\ell} \sum_{j=1}^{n} \lambda_i \mu_j \langle v_i, w_j \rangle .$$

Definition Sei V ein n-dimensionaler euklidischer Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$ und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Die Matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, definiert durch

$$a_{ij} := \langle v_i, v_j \rangle \qquad \text{für alle } i, j \in \{1, \ldots, n\} ,$$

heißt die **darstellende Matrix des Skalarproduktes** $\langle\ ,\ \rangle$ bezüglich der Basis $\mathcal{B}$. Wir schreiben auch $M_{\mathcal{B}}^{\mathcal{B}}(\langle\ ,\ \rangle)$.

Definition Es seien K ein Körper und $n \in \mathbb{N}$. Wir nennen eine Matrix $A \in K^{n \times n}$ eine **symmetrische Matrix**, wenn

$$A^T = A .$$

Ist $A = (a_{ij})$, so ist dies äquivalent zu

$$a_{ji} = a_{ij} \quad \text{für alle } i, j \in \{1, \ldots, n\} .$$

Satz 9.3. *Die darstellende Matrix eines reellen Skalarproduktes in einem n-dimensionalen euklidischen Vektorraum ist symmetrisch.*

In Satz 6.32 hatten wir gelernt, wie sich für einen Endomorphismus $F \in \mathrm{End}_K(V)$ in einem n-dimensionaler Vektorraum V über einem Körper K die zugehörigen Matrizen verändern. Seien $\mathcal{B}$ und $\mathcal{B}'$ zwei Basen von V mit F darstellenden Matrizen $A := M_{\mathcal{B}}^{\mathcal{B}}(F)$ bezüglich $\mathcal{B}$ und $A' = M_{\mathcal{B}'}^{\mathcal{B}'}(F)$ bzgl. $\mathcal{B}'$. Dann gilt:

$$M_{\mathcal{B}'}^{\mathcal{B}'}(F) = A' = P^{-1} A P = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}_V) M_{\mathcal{B}}^{\mathcal{B}}(F) M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) ,$$

wobei $P = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \in \mathrm{GL}(n, K)$.

Hierbei hängt P nicht von F, sondern nur von den Basen, ab. P beschreibt, wie sich die Koordinaten eines Vektors in V bzgl. der Basis $\mathcal{B}$ zu den Koordinaten bzgl. $\mathcal{B}'$ verändern.

Satz 9.4. (Transformationsformel für reelle Skalarprodukte) Sei V ein n-dimensionaler euklidischer Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Ferner seien $A := M_{\mathcal{B}}^{\mathcal{B}}(\langle\ ,\ \rangle)$ und $A' := M_{\mathcal{B}'}^{\mathcal{B}'}(\langle\ ,\ \rangle)$ die darstellenden Matrizen von $\langle\ ,\ \rangle$ bezüglich zweier Basen $\mathcal{B}$ und $\mathcal{B}'$. Dann gilt mit $P = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) \in \mathrm{GL}(n, \mathbb{R})$:

$$A' = P^T A P .$$

9.2 Unitäre Vektorräume

In diesem Abschnitt beschäftigen wir uns ausschließlich mit dem Fall, dass $K = \mathbb{C}$ und V ein komplexer Vektorraum, also ein $\mathbb{C}$-Vektorraum, ist. Sei

$$\mathbb{C} = \{ z = a + b\,i \mid a, b \in \mathbb{R} \}$$

die Menge der komplexen Zahlen, wobei $i^2 = -1$. Für $z = a + b\,i \in \mathbb{C}$ nennt man $\mathrm{Re}(z) = a$ den **Realteil** von z und $\mathrm{Im}(z) = b$ den **Imaginärteil** von z. Wir haben also $z = \mathrm{Re}(z) + \mathrm{Im}(z)\,i$. Man definiert die zu z **konjugiert komplexe Zahl** $\overline{z}$ als $\overline{z} := a - b\,i \in \mathbb{C}$. Ferner kennen wir für alle $z = a + b\,i, z_1, z_2 \in \mathbb{C}$ die Regeln:

- $\overline{\overline{z}} = z$.
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$.
- $\overline{z_1\, z_2} = \overline{z_1}\,\overline{z_2}$.
- $z\overline{z} = a^2 + b^2 \in \mathbb{R}_{\geq 0}$.
- $z + \overline{z} = 2\,\mathrm{Re}(z) = 2\,a$ und $z - \overline{z} = 2\,\mathrm{Im}(z)\,i = 2\,b\,i$.
- $z \in \mathbb{R} \Leftrightarrow \overline{z} = z$.

Außerdem definiert man den **Absolutbetrag** in $\mathbb{C}$,

$$|\ |\colon \mathbb{C} \to \mathbb{R}_{\geq 0}\ ,$$

wobei $|z| := \sqrt{a^2 + b^2}$. Schränkt man diesen Absolutbetrag auf $\mathbb{R}$ ein, so erhält man den üblichen Absolutbetrag auf $\mathbb{R}$. Für alle $z, z_1, z_2 \in \mathbb{C}$ gilt:

- $z\,\overline{z} = |z|^2$.
- $|z_1\, z_2| = |z_1|\,|z_2|$.
- $|z_1 + z_2| \leq |z_1| + |z_2|$.

Definition Sei $A := (a_{ij}) \in \mathbb{C}^{n\times n}$. Wir definieren die **konjugierte Transponierte** von A als

$$A^\star := (a^\star_{ij}) \in \mathbb{C}^{n\times n} \quad , \qquad \text{wobei } a^\star_{ij} := \overline{a_{ji}} \qquad \text{für alle } i, j \in \{1, \ldots, n\}\ .$$

Schreibweise: $A^\star = \overline{A}^T$, wobei $\overline{A}$ die Matrix beschreibt, welche durch Konjugation jedes Eintrages von A entsteht. Wir nennen eine Matrix $A \in \mathbb{C}^{n\times n}$ mit $A^\star = A$ **hermitesch**. Beachte: Für $A \in \mathbb{R}^{n\times n}$ ist $A^\star = A^T$. In diesem Fall sind die Begriffe hermitesch und symmetrisch identisch.

Definition Sei V ein $\mathbb{C}$-Vektorraum. Ein **(komplexes) Skalarprodukt** in V ist eine Abbildung

$$\langle\ ,\ \rangle : V \times V \to \mathbb{C}\ ,$$

die den folgenden Axiomen genügt:

(B1) $\langle\ ,\ \rangle$ ist **linear** in der ersten Komponente:

(B1a) $\langle v + v' , w \rangle = \langle v , w \rangle + \langle v' , w \rangle$ für alle $v , v' , w \in V$.

(B1b) $\langle \lambda v , w \rangle = \lambda \langle v , w \rangle$ für alle $v , w \in V$ und $\lambda \in \mathbb{C}$.

(S) $\langle\ ,\ \rangle$ ist **hermitesch**: $\overline{\langle v , w \rangle} = \langle w , v \rangle$ für alle $v , w \in V$.

(PD) $\langle\ ,\ \rangle$ ist **positiv definit**: $\langle v , v \rangle \in \mathbb{R}$ und $\langle v , v \rangle > 0$ für alle $v \in V \setminus \{0\}$.

Definition Wir nennen einen $\mathbb{C}$-Vektorraum V zusammen mit einem Skalarprodukt $\langle\ ,\ \rangle$ einen **unitären Vektorraum**. Manchmal schreiben wir auch $(V , \langle\ ,\ \rangle)$.

Beispiel: Hier sind einige Standard-Skalarprodukte.

(1) Für $V := \mathbb{C}^{n \times 1}$ ist das **Standard-Skalarprodukt** wie folgt gegeben:

$$\left\langle \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{=:x} , \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_{=:y} \right\rangle := \sum_{i=1}^{n} x_i \overline{y}_i = x^T \overline{y} = \overline{y}^T x\ .$$

Die Konjugation ist notwendig, da dies ohne die Konjugation der y_i kein Skalarprodukt ist.

(2) Sei $V := \mathbb{C}[t]_n$. Dann ist das Standard-Skalarprodukt in $\mathbb{C}[t]_n$ definiert als

$$\left\langle \sum_{i=0}^{n} a_i t^i , \sum_{i=0}^{n} b_i t^i \right\rangle := \sum_{i=0}^{n} a_i \overline{b}_i\ .$$

(3) Sei $V := K^{n \times n}$. Dann ist das Standard-Skalarprodukt in $K^{n \times n}$ definiert als

$$\langle A , B \rangle := \operatorname{Spur}\left(\overline{B}^T A\right) := \operatorname{Spur}\left(\overline{B}^T A\right) = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} \overline{b}_{ik}\ .$$

Bemerkung 9.5. *Sei V ein $\mathbb{C}$-Vektorraum. Dann ist* **(B1)** *äquivalent zu*

(B1') *Für alle $v, v', w \in V$ und $\lambda, \lambda' \in \mathbb{C}$ gilt:*

$$\langle \lambda v + \lambda' v', w \rangle = \lambda \langle v, w \rangle + \lambda' \langle v', w \rangle \; .$$

Ist **(B1)** *erfüllt, so gilt für alle $v_1, \ldots, v_\ell, w \in V$ und $\lambda_1, \ldots, \lambda_\ell \in \mathbb{C}$:*

$$\left\langle \sum_{i=1}^{\ell} \lambda_i v_i , w \right\rangle = \sum_{i=1}^{\ell} \lambda_i \langle v_i , w \rangle \; .$$

Proposition 9.6. (Elementare Eigenschaften von komplexen Skalarprodukten) Sei V ein unitärer Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$. Dann gilt:

(a) Es gilt auch das Axiom

(B2) $\langle \, , \, \rangle$ ist **konjugiert linear in der zweiten Komponente**:

(B2a) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$ für alle $v, w, w' \in V$.

(B2b*) $\langle v, \lambda w \rangle = \overline{\lambda} \langle v, w \rangle$ für alle $v, w \in V$ und $\lambda \in \mathbb{C}$.

Dies ist wiederum äquivalent zu

(B2*) $\quad \langle v, \lambda w + \lambda' w' \rangle = \overline{\lambda} \langle v, w \rangle + \overline{\lambda'} \langle v, w' \rangle$

für alle $v, w, w' \in V$ und $\lambda, \lambda' \in \mathbb{C}$. Man sagt in dem Fall auch, das Skalarprodukt sei **sesquilinear**.

(b) Für alle $v \in V$ gilt: $\langle v, v \rangle = 0 \quad \Leftrightarrow \quad v = 0$.

(c) Seien $w, w' \in V$. Falls $\langle v, w \rangle = \langle v, w' \rangle$ für alle $v \in V$, so gilt $w = w'$.

(d) Für alle $v_1, \ldots, v_\ell, w_1, \ldots, w_n \in V$ und $\lambda_1, \ldots, \lambda_\ell, \mu_1, \ldots, \mu_n \in \mathbb{C}$ gilt:

$$\left\langle \sum_{i=1}^{\ell} \lambda_i v_i , \sum_{j=1}^{n} \mu_j w_j \right\rangle = \sum_{i=1}^{\ell} \sum_{j=1}^{n} \lambda_i \overline{\mu_j} \langle v_i , w_j \rangle \; .$$

Definition Sei V ein n-dimensionaler unitärer Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$ und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Die Matrix $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, definiert durch

$$a_{ij} := \langle v_i , v_j \rangle \qquad \text{für alle } i, j \in \{1, \ldots, n\} \; ,$$

heißt die **darstellende Matrix des Skalarproduktes** $\langle \, , \, \rangle$ bezüglich der Basis $\mathcal{B}$. Wir schreiben auch $M_{\mathcal{B}}^{\mathcal{B}}(\langle \, , \, \rangle)$.

Satz 9.7. *Die darstellende Matrix eines komplexen Skalarproduktes in einem n-dimensionalen unitären Vektorraum ist hermitesch, d.h. es gilt in obiger Situation:*

$$A^T = \overline{A} \ .$$

Satz 9.8. (Transformationsformel für komplexe Skalarprodukte) *Sei V ein n-dimensionaler unitärer Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Ferner seien $A := M_{\mathcal{B}}^{\mathcal{B}}(\langle\ ,\ \rangle)$ und $A' := M_{\mathcal{B}'}^{\mathcal{B}'}(\langle\ ,\ \rangle)$ die darstellenden Matrizen von $\langle\ ,\ \rangle$ bezüglich zweier Basen $\mathcal{B}$ und $\mathcal{B}'$. Dann gilt mit $P = M_{\mathcal{B}}^{\mathcal{B}'}(\,\mathrm{id}_V\,) \in \mathrm{GL}(\,n\,,\mathbb{C}\,)$:*

$$A' = P^T A \overline{P} \ .$$

9.3 Norm und Metrik

In diesem Abschnitt beschäftigen wir uns ausschließlich mit dem Fall, dass $K = \mathbb{C}$ oder $K = \mathbb{R}$ und V ein K Vektorraum mit einem Skalarprodukt. Wir werden beide Fälle parallel betrachten und beweisen die Aussagen im Komplexen. Die Aussagen über $\mathbb{R}$ folgen auf ähnliche Art und Weise, indem man einfach die Konjugation weglässt. Sei also

$$K := \left\{ \begin{array}{c} \mathbb{R} \\ \mathbb{C} \end{array} \right\} \quad \text{und sei } V \text{ ein} \quad \left\{ \begin{array}{l} \text{euklidischer Vektorraum} \\ \text{unitärer Vektorraum} \end{array} \right\} .$$

Definition Wir nennen V einen **K-Vektorraum mit Skalarprodukt** $\langle\ ,\ \rangle$, wenn V entweder ein euklidischer $\mathbb{R}$-Vektorraum mit reellem Skalarprodukt $\langle\ ,\ \rangle$ oder ein unitärer $\mathbb{C}$-Vektorraum mit komplexem Skalarprodukt $\langle\ ,\ \rangle$ ist.

Satz 9.9. (Cauchy-Schwarzsche Ungleichung) Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann gilt für alle $v, w \in V$:

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle \langle w, w\rangle .$$

Gleichheit gilt genau dann, wenn das System (v, w) linear abhängig ist.

Definition Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann definieren wir die **Norm** oder **Länge** eines Vektors $v \in V$ als

$$\| v \| := \sqrt{\langle v, v\rangle} \quad \in \mathbb{R} .$$

Satz 9.10. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann ist die Abbildung*

$$\|\ \| : V \to \mathbb{R}$$

eine Norm in V. D.h. $\|\ \|$ erfüllt für alle $v, w \in V$ und $\lambda \in K$ die drei Bedingungen:

(a) $\| v \| \geq 0$ *und* $\| v \| = 0 \Leftrightarrow v = 0$.

(b) $\| \lambda v \| = |\lambda| \, \| v \|$.

(c) $\| v + w \| \leq \| v \| + \| w \|$. **(Dreiecksungleichung)**

Korollar 9.11. (Satz des Pythagoras) Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann gilt für alle $v, w \in V$:

$$\| v + w \|^2 = \| v \|^2 + \| w \|^2 + \langle v, w\rangle + \langle w, v\rangle .$$

Korollar 9.12. (Parallelogrammgleichung) Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann gilt für alle $v\,,\ w\ \in\ V$:

$$\|\, v+w\,\|^2\ +\ \|\, v-w\,\|^2\ =\ 2\left(\|\, v\,\|^2\ +\ \|\, w\,\|^2\right)\ .$$

Definition Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann definieren wir den **Abstand** zwischen zwei Vektoren $v\,,\ w\ \in\ V$ als

$$\mathrm{d}(v\,,\ w)\ :=\ \|\, v-w\,\|\ =\ \sqrt{\langle v-w\,,\ v-w\rangle}\ .$$

Satz 9.13. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann ist die Abbildung*

$$\mathrm{d}: V\times V\ \to\ \mathbb{R}$$

*eine **Metrik** in V. D.h.* d *erfüllt für alle $v\,,\ v'\,,\ w\ \in\ V$ die folgenden Bedingungen:*

(a) $\mathrm{d}(v\,,\ w)\ \geq\ 0$ *und* $\mathrm{d}(v\,,\ w)\ =\ 0 \Leftrightarrow v\ =\ w$.

(b) $\mathrm{d}(v\,,\ w)\ =\ \mathrm{d}(w\,,\ v)$.

(c) $\mathrm{d}(v\,,\ w)\ \leq\ \mathrm{d}(v\,,\ v')\ +\ \mathrm{d}(v'\,,\ w)$. **(Dreiecksungleichung)**

9.4 Orthonormalbasen und Gram-Schmidtsches Verfahren

Wir betrachten zentrale Aussagen über Orthogonalität in Vektorräumen mit Skalarprodukten. K und V seien wie im vorherigen Abschnitt.

Definition Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$.

(a) Zwei Vektoren $v\,, w \in V$ heißen **orthogonal**, wenn $\langle v\,, w \rangle = 0$.

Wir schreiben: $v \perp w$.

(b) Eine Familie $S = (v_i)_{i \in I}$ von Vektoren $v_i \in V$ heißt **orthogonal**, wenn je zwei verschiedene Elemente aus S orthogonal sind, d.h. wenn gilt:

$$\langle v_i\,, v_j \rangle = 0 \quad \text{für alle } i\,, j \in I \text{ mit } i \neq j \;.$$

(c) Ein Vektor $v \in V$ heißt **normiert** bzw. ein **Einheitsvektor**, wenn $\| v \| = 1$.

(d) Wir nennen eine Familie **orthonormiert** bzw. ein **Orthonormalsystem** von V, wenn sie orthogonal ist und alle Vektoren der Familie normiert sind.

Bemerkung 9.14. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$.*

(a) *Der Nullvektor ist zu allen Vektoren aus V orthogonal. Es gilt*

(i) $0 \perp v$ *für alle* $v \in V$.

(ii) *Falls* $v \perp w$ *für alle* $w \in V$, *so folgt* $v = 0$.

(b) *Die Familie $(v\,, w)$ von zwei Vektoren aus V ist genau dann orthonormiert, wenn*

$$v \perp w \qquad \textit{und} \qquad \| v \| = 1 = \| w \| \;.$$

(c) *Für alle $v \in V$ gilt:* $\| v \| = 1 \Leftrightarrow \langle v\,, v \rangle = 1$.

Proposition 9.15. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$. Dann gilt:*

(a) *Für alle $v \in V \setminus \{0\}$ ist $\frac{v}{\|v\|}$ normiert.*

(b) *Ist $S = (v_i)_{i \in I}$ eine orthogonale Familie aus V, so ist*

$$S' := (v_i')_{i \in I} \quad \text{mit} \quad v_i' := \frac{v_i}{\| v_i \|} \quad (i \in I)$$

eine orthonormierte Familie aus V. Insbesondere gilt für $I := \{1, \dots, \ell\}$: Ist $(v_1, \dots, v_\ell)$ eine orthogonale Familie von V, so ist

$$\left(\frac{v_1}{\| v_1 \|}, \dots, \frac{v_\ell}{\| v_\ell \|} \right)$$

ein Orthonormalsystem von V.

(c) *Ist $S = (v_i)_{i \in I}$ eine orthonormierte Familie aus V, dann ist*

$$\langle v_i , v_j \rangle = \delta_{ij} \qquad \text{für alle } i, j \in I \ ,$$

wobei δ_{ij} das Kroneckersymbol bezeichnet.

Definition Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$. Eine orthonormierte Familie S von Vektoren aus V, die zugleich eine Basis von V ist, heißt eine **orthonormierte Basis von V** bzw. eine **Orthonormalbasis von V**.

Satz 9.16. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$. Ferner sei $S = (v_1, \dots, v_\ell)$ eine orthogonale Familie aus V, welche den Nullvektor nicht enthält. Dann gilt für alle $v \in \mathrm{Spann}_K(v_1, \dots, v_\ell)$: Ist*

$$v = \sum_{i=1}^{\ell} \mu_i v_i \qquad \text{mit } \mu_1, \dots, \mu_\ell \in K \ ,$$

so berechnen sich die Koordinaten $\mu_1, \dots, \mu_\ell$ als

$$\mu_i = \frac{\langle v , v_i \rangle}{\| v_i \|^2} = \frac{\langle v , v_i \rangle}{\langle v_i , v_i \rangle} \qquad \text{für alle } i \in \{1, \dots, \ell\} \ .$$

Korollar 9.17. *Sei V ein n-dimensionaler K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Orthonormalbasis von V. Dann gilt für alle $v \in V$:*

$$v = \sum_{i=1}^{n} \langle v , v_i \rangle v_i \ .$$

Satz 9.18. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann ist jede orthogonale Familie S von Vektoren aus V, die den Nullvektor nicht enthält, linear unabhängig.*

Korollar 9.19. *Sei V ein n-dimensionaler K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann ist jede orthogonale Familie S von n Vektoren aus V, die den Nullvektor nicht enthält, eine Basis von V.*

Bemerkung 9.20. *Sei V ein n-dimensionaler K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$ und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Familie von Vektoren aus V. Dann sind die folgenden Aussagen äquivalent:*

(a) $\mathcal{B} = (v_1, \ldots, v_n)$ *ist eine Orthonormalbasis von V*

(b) *Es gilt:*

$$\langle v_i, v_j \rangle = \delta_{ij} \quad \text{für alle} \quad i, j \in \{1, \ldots, n\} \quad .$$

Mit anderen Worten:

$$M_{\mathcal{B}}^{\mathcal{B}}(\langle\ ,\ \rangle) = E_n \quad .$$

Satz 9.21. (Gram-Schmidtscher Orthogonalisierungssatz) Jeder endlich-dimensionale K-Vektorraum V mit Skalarprodukt $\langle\ ,\ \rangle$ besitzt eine Orthonormalbasis. Sei nämlich $(w_1, \ldots, w_n)$ eine beliebige Basis von V. Wir setzen

$$v_1 := w_1 \quad ,$$

und für alle $k = 1, \ldots, n-1$:

$$v_{k+1} := w_{k+1} - \sum_{j=1}^{k} \frac{\langle w_{k+1}, v_j \rangle}{\langle v_j, v_j \rangle} v_j = w_{k+1} - \sum_{j=1}^{k} \frac{\langle w_{k+1}, v_j \rangle}{\| v_j \|^2} v_j \quad .$$

Dann ist $(v_1, \ldots, v_n)$ eine orthogonale Familie, welche den Nullvektor nicht enthält. Setzen wir außerdem

$$v'_k := \frac{v_k}{\| v_k \|} \quad \text{für alle} \quad k \in \{1 \ldots, n\} \quad ,$$

dann ist $(v'_1, \ldots, v'_n)$ eine Orthonormalbasis von V.

9.5 Direkte Summe und Orthogonales Komplement

Die Aussagen dieses Abschnittes dienen zur Erklärung, was wirklich beim Orthogonalisierungsverfahren geschieht, und wir führen den Begriff der direkten Summe ein.

Definition Sei V ein K-Vektorraum über einem (beliebigen) Körper K. Sind U_1 und U_2 beliebige Untervektorräume von V mit

$$U_1 + U_2 = V \qquad \text{und} \qquad U_1 \cap U_2 = \{0\} \ ,$$

dann heißt V die **direkte Summe** von U_1 und U_2, in Zeichen,

$$V = U_1 \oplus U_2 \ .$$

Satz 9.22. *Sei V ein K-Vektorraum über einem (beliebigen) Körper K und seien U_1, U_2 Untervektorräume von V mit $V = U_1 + U_2$. Es sind die folgenden Aussagen äquivalent:*

(a) $V = U_1 \oplus U_2$.

(b) *Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung der Form*

$$v = u_1 + u_2 \quad (u_1 \in U_1, u_2 \in U_2) \ .$$

(c) *Aus einer Gleichung $u_1 + u_2 = 0$ mit $u_1 \in U_1$ und $u_2 \in U_2$ folgt $u_1 = u_2 = 0$.*

(d) *Für alle $u_1 \in U_1 \setminus \{0\}$ und $u_2 \in U_2 \setminus \{0\}$ ist das System (u_1, u_2) linear unabhängig über K.*

Als einfache Übungsaufgabe erhält man die folgende Dimensionsformel. Wir hatten bereits früher gesehen, dass in einem beliebigen K-Vektorraum V mit Untervektorräumen U_1, U_2 auch $U_1 + U_2$ und $U_1 \cap U_2$ Untervektorräume sind.

Satz 9.23. (Dimensionsformel für Untervektorräume) Sei V ein K-Vektorraum über einem (beliebigen) Körper K und seien U_1, U_2 Untervektorräume von V. Dann gilt:

$$\dim_K U_1 + \dim_K U_2 = \dim_K (U_1 + U_2) + \dim_K (U_1 \cap U_2) \ .$$

Korollar 9.24. *Sei V ein K-Vektorraum über einem (beliebigen) Körper K und seien U_1, U_2 Untervektorräume von V. Dann sind die folgenden Aussagen äquivalent:*

(a) $V = U_1 \oplus U_2$.

(b) $V = U_1 + U_2$ *und* $\dim_K U_1 + \dim_K U_2 = \dim_K V$.

Definition Sei V ein K-Vektorraum über einem (beliebigen) Körper K und sei U ein Untervektorraum von V. Ein Untervektorraum U' heißt ein **Komplement von U in V**, wenn

$$V = U \oplus U' \ .$$

Satz 9.25. *Jeder Untervektorraum U eines endlich-dimensionalen K-Vektorraumes über einem (beliebigen) Körper K besitzt ein Komplement in V. Insbesondere gilt für jedes Komplement U' von U:*

$$\dim_K U + \dim_K U' = \dim_K V \ .$$

Definition Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$, wobei entweder $K = \mathbb{C}$ oder $K = \mathbb{R}$. Ist U ein Untervektorraum von V, so definieren wir

$$U^{\perp} := \{ v \in V \mid \langle v, u \rangle = 0 \text{ für alle } u \in U \} = \{ v \in V \mid v \perp u \text{ für alle } u \in U \} \ .$$

$U^{\perp}$ heißt das **orthogonale Komplement von U in V**.

Zuerst muss man sich davon überzeugen, dass $U^{\perp}$ tatsächlich ein Komplement ist.

Proposition 9.26. (Elementare Eigenschaften des orthogonalen Komplements) Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$, wobei $K = \mathbb{C}$ oder $K = \mathbb{R}$, und sei U ein Untervektorraum von V. Dann gilt:

(a) $U^{\perp}$ ist ein Untervektorraum von V.

(b) $U \subseteq (U^{\perp})^{\perp}$.

(c) $\{0\}^{\perp} = V$.

(d) $V^{\perp} = \{0\}$.

Satz 9.27. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$, wobei $K = \mathbb{C}$ oder $K = \mathbb{R}$. Ist U ein endlich-dimensionaler Untervektorraum von V, dann gilt:*

(a) $V = U \oplus U^{\perp}$.

(b) $\dim_K (U^{\perp}) = \dim_K V - \dim_K U$.

(c) *Ist* $\dim_K V < \infty$, *so gilt* $U = (U^{\perp})^{\perp}$.

Einschub: Alternativer Beweis des Gram-Schmidtschen Satzes (OPTIONAL!)

Im Beweis dieses Satzes verwenden wir, dass U eine Orthonormalbasis besitzt, da U endlich-dimensional ist. D.h. wir verwendeten den Gram-Schmidtschen Orthogonalisierungssatz. Es gibt jedoch noch eine andere Variante der Vorgehensweise. Man könnte Satz 9.27 auch ohne eine orthonormierte Basis, sondern lediglich unter Verwendung einer Basis, beweisen. Dann wäre der Gram-Schmidtsche Orthogonalisierungssatz ein Korollar aus dem vorherigen Satz. Wir wären nun in der Lage, einen zweiten (sehr eleganten) Beweis des Gram-Schmidtscher Orthogonalisierungssatzes zu liefern. Eine Skizze des Beweises wird optional in der Vorlesung geliefert.

Korollar 9.28. (Gram-Schmidtscher Orthogonalisierungssatz') Sei $K = \mathbb{C}$ oder $K = \mathbb{R}$. Jeder endlich-dimensionale K-Vektorraum V mit Skalarprodukt $\langle\ ,\ \rangle$ besitzt eine Orthonormalbasis, also eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ mit

$$\langle v_i , v_j \rangle = \delta_{ij} \quad \text{für } i, j \in \{1, \ldots, n\} \ .$$

Bemerkung 9.29. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$, wobei $K = \mathbb{C}$ oder $K = \mathbb{R}$. Sei zudem U ein endlich-dimensionaler Untervektorraum von V und sei $\mathcal{B} = (v_1, \ldots, v_\ell)$ eine Orthonormalbasis von U. Dann existiert für alle $v \in V$ ein eindeutiger Vektor $w \in U$ mit $v - w \in U^\perp$. Diesen Vektor nennt man die **orthogonale Projektion** von v auf U, und es gilt:*

$$w = \mathrm{proj}_U(v) := \sum_{i=1}^{\ell} \langle v , v_i \rangle v_i \ .$$

Außerdem ist die folgende Abbildung linear:

$$\begin{array}{rccl} F: & V & \to & U \\ & v & \mapsto & F(v) := \sum_{i=1}^{\ell} \langle v , v_i \rangle v_i \end{array} .$$

Das Konzept der direkten Summe lässt sich wie folgt verallgemeinern.

Definition Sei V ein Vektorraum über einem (beliebigen) Körper K und seien $U_1, \ldots, U_m$ Untervektorräume von V. Wir nennen V die **direkte Summe** der U_i, wenn

(a) V ist gleich der formalen Summe der Untervektorräume $U_1, \ldots, U_m$, d.h.

$$V = \sum_{i=1}^{m} U_i = U_1 + \ldots + U_m = \{ u_1 + \ldots + u_m \mid u_i \in U_i \text{ für alle } i = 1, \ldots, m \} .$$

(b) Jeder Vektor $v \in V$ besitzt genau eine Darstellung der Form

$$v = \sum_{i=1}^{m} u_i \quad \text{wobei } u_i \in U_i \text{ für alle } i = 1, \ldots, m .$$

Schreibweise: $V = \bigoplus_{i=1}^{m} U_i = U_1 \oplus U_2 \oplus \ldots \oplus U_m$.

Satz 9.30. *Sei V ein Vektorraum über einem (beliebigen) Körper K und seien $U_1, \ldots, U_m$ Untervektorräume von V. Ist $V = \sum_{i=1}^{m} U_i$, so sind die folgenden Aussagen äquivalent:*

(a) $V = \bigoplus_{i=1}^{m} U_i$.

(b) *Aus einer Gleichung* $\sum_{i=1}^{m} u_i = 0$ *mit* $u_i \in U_i$ *folgt*

$$u_i = 0 \quad \text{für alle } i \in \{ 1, \ldots, m \} .$$

(c) *Für $k = 1, \ldots, m$ gilt:*

$$U_k \cap \sum_{\substack{i=1 \\ i \neq k}}^{m} U_i = \{ 0 \} .$$

Korollar 9.31. *Sei V ein n-dimensionaler Vektorraum über einem (beliebigen) Körper K und sei $F \in \mathrm{End}_K(V)$. Ferner seien $\lambda_1, \ldots, \lambda_m$ die sämtlichen Eigenwerte von F mit zugehörigen Eigenräumen $\mathrm{Eig}(F, \lambda_i)$. Bezeichnet*

$$V' = \sum_{i=1}^{m} \mathrm{Eig}(F, \lambda_i)$$

den von den zugehörigen Eigenräumen erzeugten Untervektorraum, so ist diese Summe bereits direkt, d.h. es gilt:

$$V' = \bigoplus_{i=1}^{m} \mathrm{Eig}(F, \lambda_i) .$$

Außerdem gilt: F ist genau dann diagonalisierbar, wenn

$$V = \bigoplus_{i=1}^{m} \mathrm{Eig}(F, \lambda_i) .$$

9.6 Orthogonale Matrizen und Orthogonale Endomorphismen

Wir haben die Existenz von Orthonormalbasen in endlich-dimensionalen Vektorräumen mit Skalarprodukt bewiesen. Im Folgenden betrachten wir die zugehörigen Darstellungsmatrizen, welche wir als orthogonal identifizieren werden. Das Ziel dieses Abschnitts ist es deshalb, orthogonale Matrizen und Endomorphismen zu klassifizieren. Zuerst untersuchen wir die so genannten orthogonalen Matrizen. In diesem Abschnitt sei wieder $K = \mathbb{C}$ oder $K = \mathbb{R}$, und es sei V ein K Vektorraum mit einem Skalarprodukt.

Definition Eine Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ heißt **orthogonal**, wenn gilt:

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

Eine Matrix $A \in \mathrm{GL}(n, \mathbb{C})$ heißt **unitär**, wenn gilt:

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

Definition Für $n \in \mathbb{N}$ definieren wir

- $O(n) := \{ A \in \mathrm{GL}(n, \mathbb{R}) \mid A \text{ orthogonal} \} = \left\{ A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{-1} = A^T \right\}$.
- $U(n) := \{ A \in \mathrm{GL}(n, \mathbb{C}) \mid A \text{ unitär} \} = \left\{ A \in \mathrm{GL}(n, \mathbb{C}) \mid A^{-1} = \overline{A}^T \right\}$.

Satz 9.32. *Für alle $n \in \mathbb{N}$ gilt:*

(a) $O(n)$ *ist eine Untergruppe von* $\mathrm{GL}(n, \mathbb{R})$, *genannt die* ***orthogonale Gruppe***.

(b) $U(n)$ *ist eine Untergruppe von* $\mathrm{GL}(n, \mathbb{C})$, *genannt die* ***unitäre Gruppe***.

(c) $|\det(A)| = 1$ *für alle* $A \in O(n) \cup U(n)$.

Satz 9.33. *Für eine Matrix $A \in \mathrm{GL}(n, \mathbb{R})$ sind die folgenden Aussagen äquivalent:*

(a) *A ist orthogonal.*

(b) $A^T A = E_n$.

(c) $A A^T = E_n$.

(d) *Die Spalten von A bilden eine Orthonormalbasis des $\mathbb{R}^{n\times 1}$ bezüglich des Standardskalarproduktes von $\mathbb{R}^{n\times 1}$.*

(e) *Die Zeilen von A bilden eine Orthonormalbasis des $\mathbb{R}^{1\times n}$ bezüglich des Standardskalarproduktes von $\mathbb{R}^{1\times n}$.*

(f) $A = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_V)$ *, wobei $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen eines n-dimensionalen euklidischen Vektorraumes sind.*

Satz 9.34. *Für eine Matrix $A \in \mathrm{GL}(n, \mathbb{C})$ sind die folgenden Aussagen äquivalent:*

(a) *A ist orthogonal.*

(b) $\overline{A}^T A = E_n$.

(c) $A \overline{A}^T = E_n$.

(d) *Die Spalten von A bilden eine Orthonormalbasis des $\mathbb{C}^{n\times 1}$ bezüglich des Standardskalarproduktes von $\mathbb{C}^{n\times 1}$.*

(e) *Die Zeilen von A bilden eine Orthonormalbasis des $\mathbb{C}^{1\times n}$ bezüglich des Standardskalarproduktes von $\mathbb{C}^{1\times n}$.*

(f) $A = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_V)$ *, wobei $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen eines n-dimensionalen unitären Vektorraumes sind.*

Mithilfe trigonometrischer Eigenschaften (Polarkoordinaten, Additionstheoreme) sind wir nun in der Lage, alle orthogonalen Matrizen zu klassifizieren.

Satz 9.35. *Sei $A \in O(2)$. Dann existiert ein $\alpha \in [0, 2\pi[$, so dass*

- *entweder*
$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad ;$$
in dem Fall ist $\det(A) = 1$ *und die Standardinterpretation von A beschreibt eine* ***Drehung*** *um α.*
- *oder*
$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \quad ;$$
in dem Fall ist $\det(A) = -1$ *und die Standardinterpretation von A beschreibt eine* ***Spiegelung*** *an einer Geraden.*

Umgekehrt ist jede Drehmatrix und jede Spiegelungsmatrix natürlich eine orthogonale Matrix. Insgesamt erhalten wir
$$O(2) = \left\{ \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \,\middle|\, 0 \le \alpha < 2\pi \right\} \cup \left\{ \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \,\middle|\, 0 \le \alpha < 2\pi \right\} .$$

Definition Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$. Wir nennen einen Endomorphismus $F \in \mathrm{End}_K(V)$ einen **orthogonalen Endomorphismus** für $K = \mathbb{R}$ bzw. einen **unitären Endomorphismus** für $K = \mathbb{C}$, wenn für alle $v, w \in V$ gilt
$$\langle F(v), F(w) \rangle = \langle v, w \rangle \; .$$
Einen orthogonalen bzw. unitären Isomorphismus nennt man eine **Isometrie**.

Proposition 9.36. *Sei V ein K-Vektorraum mit Skalarprodukt $\langle \, , \, \rangle$ und sei $F \in \mathrm{End}_K(V)$ orthogonal (bzw. unitär). Dann gilt:*

(a) *F ist injektiv. Ist $\dim_K V = n < \infty$, so ist F auch surjektiv, also eine Isometrie.*

(b) *Ist $v \perp w$, so ist auch $F(v) \perp F(w)$.*

(c) *Jeder Eigenwert von F hat den Betrag $|\lambda| = 1$.*

Satz 9.37. *Sei V ein endlich-dimensionaler K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Für $F \in \mathrm{End}_K(V)$ sind die folgenden Aussagen äquivalent:*

(a) $\langle F(v), F(w) \rangle = \langle v, w \rangle$ *für alle* $v, w \in V$.

(b) *Ist* $\mathcal{B} = (v_1, \ldots, v_n)$ *eine Orthonormalbasis von* V, *so ist auch* $(F(v_1), \ldots, F(v_n))$ *eine Orthonormalbasis von* V.

(c) $\| F(v) \| = \| v \|$ *für alle* $v \in V$. **(Längentreue)**

Es stellt sich nun die Frage, wie die darstellenden Matrizen von orthogonalen Matrizen aussehen können. Wir suchen natürlich wieder besonders einfache Darstellungsmatrizen. Zuerst erwähnen wir ein Resultat, welches sich aus der Begrifflichkeit ergibt.

Satz 9.38. *Sei V ein n-dimensionaler euklidischer Vektorraum mit gegebener Orthonormalbasis $\mathcal{B} = (v_1, \ldots, v_n)$. Dann gilt:*

$$F \in \mathrm{End}_{\mathbb{R}}(V) \text{ ist genau dann orthogonal, wenn } M_{\mathcal{B}}^{\mathcal{B}}(F) \text{ orthogonal ist.}$$

Satz 9.39. *Sei V ein n-dimensionaler unitärer Vektorraum mit gegebener Orthonormalbasis $\mathcal{B} = (v_1, \ldots, v_n)$. Dann gilt:*

$$F \in \mathrm{End}_{\mathbb{C}}(V) \text{ ist genau dann unitär, wenn } M_{\mathcal{B}}^{\mathcal{B}}(F) \text{ unitär ist.}$$

Satz 9.40. *Sei V ein 2-dimensionaler euklidischer Vektorraum und sei $F \in \mathrm{End}_{\mathbb{R}}(V)$ orthogonal. Dann existiert eine Orthonormalbasis $\mathcal{B} = (v_1, v_2)$ von V, so dass*

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{oder} \quad M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

wobei $\alpha \in [0, 2\pi[$.

Satz 9.41. (Darstellung unitärer Endomorphismen) Sei V ein n-dimensionaler unitärer Vektorraum. Für jeden unitären Endomorphismus $F \in \mathrm{End}_{\mathbb{C}}(V)$ gibt es eine Orthonormalbasis $\mathcal{B}$ von V, die aus Eigenvektoren von F besteht. Insbesondere ist F diagonalisierbar, und die Darstellungsmatrix ist von der Form

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} e^{i\alpha_1} & 0 & \dots & 0 \\ 0 & e^{i\alpha_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\alpha_n} \end{pmatrix},$$

wobei $0 \leq \alpha_j < 2\pi$ für alle $j \in \{1, \dots, n\}$.

Korollar 9.42. *Sei $A \in U(n)$. Dann existiert eine Matrix $P \in U(n)$, so dass*

$$\overline{P}^T A P = \begin{pmatrix} e^{i\alpha_1} & 0 & \dots & 0 \\ 0 & e^{i\alpha_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\alpha_n} \end{pmatrix},$$

wobei $0 \leq \alpha_j < 2\pi$ für alle $j \in \{1, \dots, n\}$.

Satz 9.43. (Darstellung orthogonaler Endomorphismen) Sei V ein n-dimensionaler euklidischer Vektorraum. Für jeden orthogonalen Endomorphismus $F \in \mathrm{End}_{\mathbb{R}}(V)$ gibt es eine Orthonormalbasis $\mathcal{B}$ von V, bzgl. der sich F wie folgt darstellen lässt:

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & 0 & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & 0 & & & \boxed{A_1} & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & \boxed{A_k} \end{pmatrix},$$

wobei es für jedes $j \in \{1, \ldots, k\}$ ein $\alpha_j \in [0, 2\pi[$ gibt, so dass

$$A_j = \begin{pmatrix} \cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix} .$$

Wir bemerken, dass man die Theorie noch etwas weiterführen kann, indem man selbst-adjungierte Endomorphismen einführt und für diese dann in Analogie zu den beiden zuletzt bewiesenen Sätzen die Resultate formuliert. Dabei muss man lediglich die Theorie selbst-adjungierter Endomorphismen in Zusammenhang mit symmetrischen bzw. hermiteschen Matrizen stellen und die Beweise etwas modifizieren. Damit erhält man den Spektralsatz und die Hauptachsentransformation. Beides ist sehr gut in der üblichen Literatur erklärt.

Viele weiterführende Aussagen über reelle symmetrische Matrizen, Spektralsatz, normale Endomorphismen, selbst-adjungierte Endomorphismen, Hauptachsentransformation werden als Übungsaufgaben behandelt und können einfach verifiziert werden.