Lineare Algebra

Andreas Kollross

7. Februar 2023

Inhaltsverzeichnis

Kapitel 1

Logik, Mengenlehre, Relationen

Eine *Aussage* ist ein sprachliches Gebilde, von dem es sinnvoll ist, zu fragen, ob es wahr oder falsch ist. Beispiele dafür sind:

> "Es regnet."
>
> "Die Erde ist eine Scheibe."

Dabei geht es nur darum, ob das sprachliche Gebilde von seiner Bauart her geeignet ist, einen Wahrheitswert zu haben, also wahr oder falsch zu sein. Keine Aussagen sind etwa: "Guten Appetit!" oder "Wie spät ist es?".

Aussagen sind in der Mathematik allgegenwärtig, jeder mathematische Satz ist eine Aussage und es gehört zu den täglichen Aufgaben jeder Mathematikerin und jedes Mathematikers, zu beurteilen, ob Aussagen wahr oder falsch sind. Dabei interessieren uns natürlich vor allem mathematische Aussagen wie

> "13 ist eine gerade Zahl."
>
> "99991 ist eine Primzahl."

"Es gibt unendlich viele Primzahlen."

Jeder dieser Sätze ist eine Aussage, denn wir können, zumindest prinzipiell, jedem einen Wahrheitswert, nämlich w (=wahr) oder f (=falsch) zuordnen. Dabei kommt es nicht darauf an, ob wir tatsächlich in der Lage sind, richtig über den Wahrheitswert zu entscheiden. Es gibt sehr viele mathematische Aussagen, von denen (noch) niemand weiß, ob sie wahr oder falsch sind.

1.1 Aussagenlogik

In der Aussagenlogik geht es darum, einzelne Aussagen miteinander zu verknüpfen. Dabei betrachten wir die einzelnen Aussagen zunächst nicht mehr im Detail, sondern wir sehen sie im Folgenden nur als Bausteine, als sogenannte *Elementaraussagen*, an, aus denen wir größere Aussagen, sogenannte *aussagenlogische Formeln*, aufbauen wollen. Die einzelnen Aussagen, die an einer solchen Formel beteiligt sind, bezeichnen wir in diesem Abschnitt deshalb nur noch mit einem Buchstaben wie p, q, etc. Die Verknüpfungen heißen *logische Junktoren*.

Die einfachste Möglichkeit, eine solche aussagenlogische Formel zu bauen, ist die *Negation*, also die Verneinung einer Aussage. Die Verneinung der Aussage "Es regnet." ist natürlich die Aussage "Es regnet nicht." Die Negation einer Aussage p bezeichnen wir mit $\neg p$. Die Aussage $\neg p$ (gesprochen "nicht p") ist definiert als die Aussage, die dann wahr ist, wenn p falsch ist und die dann falsch ist, wenn p wahr ist. Das kann man in Form einer Tabelle so ausdrücken:

p	$\neg p$
w	f
f	w

Eine Tabelle, die den Wahrheitswert einer aussagenlogischen Formel in Abhängigkeit der Wahrheitswerte der beteiligten Einzelaussagen angibt, nennt man *Wahrheitstafel* oder *Wahrheitstabelle*. Aus zwei Einzelaussagen p und q kann man eine

Reihe weiterer aussagenlogischer Formeln aufbauen. Die wichtigsten Möglichkeiten dafür wollen wir im Folgenden kennenlernen.

Konjunktion: Aus zwei Aussagen p und q kann man die Aussage $p \wedge q$ (gesprochen "p und q") bilden. Diese zusammengesetzte Aussage ist genau dann wahr, wenn beide Aussagen p und q wahr sind.

Disjunktion: Dies ist die Verknüpfung zweier Aussagen mit "oder", als Formel: $p \vee q$ (gesprochen "p oder q"). Die Aussage $p \vee q$ ist wahr, wenn p oder q (oder beide) wahr sind, sie ist genau dann falsch, wenn sowohl p als auch q falsch sind.

Die Konjunktion und Disjunktion können wir in Form von Wahrheitstabellen wie folgt darstellen.

p	q	$p \wedge q$
w	w	w
f	w	f
w	f	f
f	f	f

p	q	$p \vee q$
w	w	w
f	w	w
w	f	w
f	f	f

Äquivalenz: "Äquivalent" bedeutet "gleichwertig". Werden zwei Aussagen p und q mit der Äquivalenzverknüpfung zu einer Gesamtaussage zusammengebaut, so ist die Gesamtaussage genau dann wahr, wenn die beiden einzelnen Aussagen p und q den gleichen Wahrheitswert haben. Die so definierte Aussage $p \Leftrightarrow q$ (gesprochen "p ist zu q äquivalent." oder auch "p genau dann, wenn q.") ist also genau dann wahr, wenn p und q beide wahr oder beide falsch sind.

Implikation: Hier werden zwei Aussagen p und q zu der Gesamtaussage $p \Rightarrow q$ (gesprochen "wenn p, dann q." oder "p impliziert q.") zusammengefügt. Diese Gesamtaussage soll den Sachverhalt ausdrücken, dass wann immer p wahr ist, auch q wahr sein muss. Man kann sie durch $(\neg p) \vee q$ ausdrücken.

Auch die Äquivalenz und Implikation wollen wir durch Wahrheitstabellen darstellen.

p	q	$p \Leftrightarrow q$
w	w	w
f	w	f
w	f	f
f	f	w

p	q	$p \Rightarrow q$
w	w	w
f	w	w
w	f	f
f	f	w

Die Definition der Implikation mag überraschen. Wir wollen dafür ein Beispiel geben. Ist p die Aussage "Es regnet." und q die Aussage "Die Straße ist nass.", dann bedeutet die zusammengesetzte Aussage $p \Rightarrow q$: "Wenn es regnet, ist die Straße nass." Wenn nun p wahr ist, wenn es also regnet, dann muss auch q wahr, also die Straße nass sein, damit die Gesamtaussage wahr ist. Ist aber p falsch, es regnet also nicht, dann kann q wahr oder falsch sein, die Straße kann nass oder trocken sein, die Gesamtaussage ist in beiden Fällen wahr.

Eine aussagenlogische Formel, die unabhängig vom Wahrheitswert ihrer Teilformeln immer wahr ist, heißt *allgemeingültig* oder *Tautologie*. Ein Beispiel für eine Tautologie ist $p \vee (\neg p)$. Zwei aussagenlogische Formeln heißen *äquivalent*, wenn sie bei gleicher Belegung ihrer Elementaraussagen mit Wahrheitswerten stets die gleichen Wahrheitswerte liefern. (Dies ist wiederum genau dann der Fall, wenn die beiden Formeln mit der Äquivalenz verknüpft eine allgemeingültige Formel ergeben.)

Dazu wollen wir ein Beispiel betrachten. Wir wollen mit Hilfe einer Wahrheitstabelle beweisen, dass die beiden aussagenlogischen Formeln $\neg(p \wedge q)$ und $(\neg p) \vee (\neg q)$ äquivalent sind. Da beide Formeln nur die zwei Elementaraussagen p und q enthalten, genügt es, 4 Fälle durchzugehen. Das Ergebnis notieren wir in der folgenden Wahrheitstafel.

p	q	$\neg(p \wedge q)$	$(\neg p) \vee (\neg q)$	$(\neg(p \wedge q)) \Leftrightarrow ((\neg p) \vee (\neg q))$
w	w	f	f	w
f	w	w	w	w
w	f	w	w	w
f	f	w	w	w

Allgemein kann man immer mit Hilfe einer Wahrheitstabelle nachprüfen, ob eine aussagenlogische Formel allgemeingültig ist oder ob zwei Formeln äquivalent sind. Bei n beteiligten Elementaraussagen muss man dazu 2^n Fälle überprüfen. So kann man auch die Allgemeingültigkeit der folgenden aussagenlogischen Formeln beweisen.

Satz 1.1.1. *Die folgenden aussagenlogischen Formeln sind allgemeingültig.*

(i) Doppelte Verneinung: $\neg(\neg p) \Leftrightarrow p$.

(ii) De Morgansche Regeln:

  (a) $\neg(p \vee q) \Leftrightarrow (\neg p) \wedge (\neg q)$,

  (b) $\neg(p \wedge q) \Leftrightarrow (\neg p) \vee (\neg q)$.

(iii) Distributivgesetze für "und" und "oder":

  (a) $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$,

  (b) $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$.

1.2 Quantoren und Prädikatenlogik

Ein Prädikat ist ein Ausdruck, der die Form einer Aussage hat, aber Variablen enthält, zum Beispiel

"m ist eine gerade Zahl."

Eine Aussage wird daraus erst, wenn wir angeben, für welche m die Aussage gelten soll. Eine Möglichkeit dazu sind Quantoren. Sei M eine Menge, z.B. ein Zahlbereich, und sei $p(m)$ für jedes Element aus M eine Aussage. Dann können wir mit dem *Allquantor* die Aussage

"Für alle m aus M ist die Aussage $p(m)$ wahr."

bilden, mit Formelsymbolen ausgedrückt:

$$\forall m \in M\colon p(m).$$

Wenn zum Beispiel $\mathbb{N}$ den Zahlbereich der natürlichen Zahlen $1, 2, 3, \ldots$ bezeichnet und $p(m)$ die Aussage "m ist eine gerade Zahl.", wird daraus:

"Für alle m aus $\mathbb{N}$ ist m eine gerade Zahl."

Besser formuliert: "Jede natürliche Zahl ist gerade." Diese Aussage ist selbstverständlich falsch.

Eine anderer Quantor ist der *Existenzquantor*. Mit ihm können wir eine Aussage der Form

"Es gibt ein m aus M, so dass die Aussage $p(m)$ wahr ist."

bilden, in Formelsymbolen geschrieben:

$$\exists m \in M\colon p(m).$$

Ersetzt man im obigen Beispiel den Allquantor durch den Existenzquantor, erhält man die Aussage:

"Es gibt ein m aus $\mathbb{N}$, so dass m eine gerade Zahl ist."

Besser formuliert: "Es gibt eine gerade natürliche Zahl." Diese Aussage ist selbstverständlich wahr.

Eine Warnung: Existenz und Allquantoren darf man im allgemeinen nicht vertauschen. So ist etwa die Aussage

$$\forall n \in \mathbb{N} \;\; \exists m \in \mathbb{N} \colon m > n$$

wahr, sie bedeutet: “Zu jeder natürlichen Zahl gibt es eine, die größer ist.” Die Aussage

$$\exists m \in \mathbb{N} \;\; \forall n \in \mathbb{N} \colon m > n$$

aber ist falsch, denn sie bedeutet: “Es gibt eine natürliche Zahl, die größer ist als alle natürlichen Zahlen.”

Häufig wird auch der *Quantor der eindeutigen Existenz* verwendet, dabei wird der Existenzquantor mit einem Ausrufungszeichen versehen. Die Formel

$$\exists^! m \in M : p(m).$$

bedeutet

“Es gibt *genau* ein m aus M, so dass die Aussage $p(m)$ wahr ist.”

Beispielsweise ist die Aussage

$$\exists^! m \in \mathbb{N} \colon m + 1 = 3.$$

wahr, denn es gibt genau eine natürliche Zahl m, so dass $m + 1$ gleich 3 ist (nämlich 2).

Eine *Menge* ist eine wohldefinierte Gesamtheit von Objekten, den *Elementen* der Menge. Beispiele sind

Die Menge der Jahreszeiten $J = \{$Frühling, Sommer, Herbst, Winter$\}$.

Die Menge der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Die Menge der ganzen Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

Die Menge der rationalen Zahlen $\mathbb{Q} = \{\frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{N}\}$.

Mit

$$x \in M$$

bezeichnen wir die Aussage, die genau dann wahr ist, wenn das Objekt x ein Element der Menge M ist, mit $x \notin M$ die Verneinung davon. So sind z.B. die Aussagen $1 \in \mathbb{N}$, $-12 \in \mathbb{Z}$, $\frac{1}{2} \in \mathbb{Q}$ wahr, die Aussagen $-1 \in \mathbb{N}$, $\frac{1}{2} \in \mathbb{Z}$, $\sqrt{2} \in \mathbb{Q}$ falsch.

Mengen können endlich oder unendlich viele Elemente besitzen. Es gibt verschiedene Möglichkeiten, Mengen zu definieren: Durch *Aufzählung*, z.B. $M = \{1, 2, 3\}$ oder $G := \{2, 4, 6, \ldots\}$. Durch *Aussonderung*, z.B. $H := \{n \in \mathbb{N} \mid n \text{ ist gerade}\}$ aus einer bereits definierten Menge. Dabei bezeichnet $\{x \in M \mid p(m)\}$ die Menge, die aus allen Elementen von M besteht, für die die Aussage $p(m)$ gilt. Bei einer Menge kommt es nicht darauf an, wie sie beschrieben wird, sondern welche Elemente sie enthält. So sind die beiden eben definierten Mengen G und H gleich.

Die *leere Menge* ist die Menge, die kein Element enthält. Man schreibt $\emptyset$ oder $\{\}$ für die leere Menge. Ein Element kann in einer Menge nicht mehrfach auftreten, also bezeichnen $\{-5, -5\}$ und $\{-5\}$ dieselbe Menge. Mengen können andere Mengen als Elemente enthalten, so ist etwa $\{1, \{2, 3\}\}$ eine Menge, die die beiden Elemente 1 und $\{2, 3\}$ enthält. Sie ist *nicht* gleich der Menge $\{1, 2, 3\}$.

Gilt für zwei Mengen A und B, dass jedes Element von A auch ein Element von B ist, so sagt, man A ist eine *Teilmenge* von B, in Zeichen: $A \subseteq B$. Ist A eine Teilmenge von B und ist A nicht gleich B, so sagt man, A ist eine *echte Teilmenge* von B.

Für eine Menge M definieren wir die *Potenzmenge* von M

$$\mathrm{Pot}(M) := \{A \mid A \subseteq M\}$$

als die Menge aller Teilmengen von M. Zum Beispiel gilt

$$\mathrm{Pot}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Die Potenzmenge einer endlichen Menge M hat 2^n Elemente, wenn n die Anzahl der Elemente von M bezeichnet.

Mengen können Mengen enthalten, aber es kann keine Menge aller Menge geben, denn wäre $M = \{X \mid X \text{ Menge}\}$ selbst eine Menge, dann würde sich M selbst als Element enthalten, $M \in M$. Andererseits könnte man von M die Teilmenge aller Mengen, die sich nicht selbst als Element enthalten,

$$R = \{X \in M \mid X \notin X\},$$

definieren. Nun folgt aber aus $R \in R$, dass $R \notin R$ gilt, und umgekehrt. Dies ist ein Widerspruch, die sogenannte *Russellsche Antinomie*. Also kann es die Menge R, und somit auch die Menge M, nicht geben, d.h. die Gesamtheit aller Mengen ist selbst keine wohldefinierte Menge.

Sind mehrere Mengen gegeben, so kann man aus diesen neue konstruieren. Wir geben dazu einige Möglichkeiten an.

Schnittmenge: Sind A und B zwei Mengen, so definiert man die Schnittmenge $A \cap B$ durch

$$A \cap B := \{x \in A \mid x \in B\},$$

d.h. als die Menge aller Objekte, die sowohl in A als auch in B als Elemente enthalten sind. Ist I eine nichtleere Menge von Mengen A_i, so bezeichnet

$$\bigcap_{i \in I} A_i$$

die Menge aller Objekte, die in jeder der Mengen A_i enthalten sind.

Vereinigung: Sind A und B zwei Mengen, so definiert man die Vereinigung $A \cup B$ durch

$$A \cup B := \{x \mid x \in A \vee x \in B\},$$

d.h. als die Menge, die ein Objekt genau dann enthält, wenn es in A oder in B (oder in beiden) als Element enthalten ist. Ist I eine Menge von Mengen A_i, so bezeichnet

$$\bigcup_{i \in I} A_i$$

die Menge aller Objekte, die in mindestens einer der Mengen A_i enthalten sind.

Komplement: Sind A und B zwei Mengen, so definiert man das Komplement $A \setminus B$ von A in B, durch

$$A \setminus B := \{x \in A \mid x \notin B\}$$

als die Menge, die ein Objekt genau dann enthält, wenn es in A, aber nicht in B als Element enthalten ist. Dabei ist es nicht notwendig, dass B eine Teilmenge von A ist, z.B. gelten $\mathbb{Z} \setminus \mathbb{N} = \{\ldots, -2, -1, 0\}$ und $\mathbb{N} \setminus \mathbb{Z} = \emptyset$.

Kartesisches Produkt: Sind A und B zwei Mengen, so bezeichnet $A \times B$ die Menge aller geordneten Paare (a, b), wobei $a \in A$ und $b \in B$ gelten:

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

Z.B. gilt:

$$\{0, 1\} \times \{3, 6, 9\} = \{(0, 3), (0, 6), (0, 9), (1, 3), (1, 6), (1, 9)\}.$$

Analog definiert man das kartesische Produkt von n Mengen als

$$A_1 \times A_2 \times \ldots \times A_n := \{(a_1, a_2, \ldots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\}.$$

Man bezeichnet z.B. mit $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ die Menge aller geordneten Tripel von reellen Zahlen und mit $\mathbb{R}^n$ die Menge aller *n-Tupel* von reellen Zahlen.

1.3 Relationen

Seien A und B zwei Mengen. Unter einer *Relation* über A und B versteht man eine Teilmenge des kartesischen Produkts $A \times B$. Sei etwa $A = \{1, 2\}$, $B = \{1, 2, 3\}$, dann ist

$$A \times B = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3)\}.$$

Die Relation $\geq$, d.h. “größer gleich”, ist gegeben durch die Teilmenge

$$\{(1,1),(2,1),(2,2)\}.$$

Mit einer Relation über A und B beschreibt man eine Beziehung, die zwischen einem Element von A und einem von B bestehen kann. Ist das Paar (a,b) in der Relation enthalten, sagt man, a und b *stehen in Relation.*

Ein anderes Beispiel ist

$$\{(x,y) \in \mathbb{R}^2 \mid y = x^2\}.$$

Die Elemente von $\mathbb{R}^2$, die diese Relation erfüllen, liegen auf der Standardparabel. Hier gilt, dass es zu jedem $x \in \mathbb{R}$ genau ein $y \in \mathbb{R}$ gibt, nämlich x^2, das mit x in Relation steht. Relationen mit dieser Eigenschaft nennt man *Abbildungen*.

1.4 Äquivalenzrelationen

Dieser Abschnitt wird erst später benötigt, beim ersten Lesen kann er übersprungen werden.

Definition 1.4.1. Sei M eine Menge und R eine Relation auf $M \times M$. Dass zwei Elemente a, b von M in Relation stehen, drücken wir durch die Schreibweise $a \sim b$ aus. Die Relation heißt *Äquivalenzrelation*, wenn sie die drei folgenden Eigenschaften besitzt.

(Ä1) *Reflexivität:* $\forall a \in M \colon a \sim a$.

(Ä2) *Symmetrie:* $\forall a, b \in M \colon a \sim b \Rightarrow b \sim a$.

(Ä3) *Transitivität:* $\forall a, b, c \in M \colon (a \sim b \wedge b \sim c) \Rightarrow a \sim c$.

Ein elementares Beispiel für eine Äquivalenzrelation ist die Gleichheitsrelation auf einer beliebigen Menge, wobei jedes Element nur mit sich selbst in Relation steht. Für eine Äquivalenzrelation $\sim$ ist die *Äquivalenzklasse* eines Elements $x \in M$ durch

$$[x] := \{y \in M \mid y \sim x\}$$

definiert.

Satz 1.4.2. *Ist M eine Menge und $\sim$ eine Äquivalenzrelation auf M, so gilt, dass die Menge der Äquivalenzklassen*

$$\{[x] \mid x \in M\}$$

eine disjunkte Zerlegung *oder* Partition *von M bildet, d.h. es gilt*

$$M = \bigcup_{x \in M} [x]$$

und je zwei Äquivalenzklassen stimmen entweder überein oder haben leere Schnittmenge.

Umgekehrt, ist $\{M_j \mid j \in J\}$ eine disjunkte Zerlegung einer Menge M, d.h. es gilt $M_j \cap M_k = \emptyset$ falls $j \neq k$ und $M = \bigcup_{j \in J} M_j$, so ist durch

$$x \sim y \Leftrightarrow \exists j \in J \colon x, y \in M_j$$

eine Äquivalenzrelation auf M definiert.

Beweis. Wir zeigen zunächst, dass zwei Äquivalenzklassen, die eine nichtleere Schnittmenge haben, übereinstimmen: Gilt $z \in [x] \cap [y]$, so folgt für ein $w \in [x]$ zunächst $w \sim x$ und $z \sim x$, was wegen der Symmetrie zu $x \sim z$ äquivalent ist. Aus der Transitivität folgt $w \sim z$. Nun folgt mit $z \sim y$ und Transitivität, dass w mit y in Relation steht. Wir haben $[x] \subseteq [y]$ gezeigt. Analog zeigt man $[y] \subseteq [x]$.

Aus der Reflexivität folgt, dass $x \in [x]$ für alle $x \in M$ gilt, d.h. jedes Element ist in einer Äquivalenzklasse enthalten. Dies zeigt, dass die Vereinigung aller Äquivalenzklassen ganz M ergibt.

Für den zweiten Teil der Behauptung müssen wir überprüfen, dass für die angegebene Relation die drei Bedingungen (Ä1), (Ä2), (Ä3) gelten. Die Reflexivität folgt aus $M = \bigcup_{j \in J} M_j$, denn da jedes x in einem M_j enthalten ist, gilt $x \sim x$. Die Symmetrie gilt offensichtlich. Um die Transitivität zu zeigen, benutzen wir die Disjunktheit der Zerlegung. Falls nämlich $x \sim y$ und $y \sim z$ gelten, so folgt zunächst, dass es ein $j \in J$ gibt, so dass $x, y \in M_j$ und dass es ein $k \in J$ gibt, so dass $y, z \in M_k$. Es gilt also $y \in M_j \cap M_k$, was $j = k$ und somit $x \sim z$ zeigt. $\square$

Auf $M = \mathbb{Z}$ können wir für $p \in \mathbb{N}$ die folgende Relation definieren.

$$n \sim m \text{ genau dann wenn } n - m \text{ durch } p \text{ teilbar ist.}$$

Die Äquivalenzklassen zu dieser Äquivalenzrelation werden *Restklassen* genannt, die Äquivalenzklasse von n besteht aus den ganzen Zahlen, die bei Division durch p den gleichen Rest wie bei n ergeben. Für $p = 2$ gibt es zwei Restklassen, die geraden Zahlen und die ungeraden Zahlen.

1.5 Abbildungen

Seien A und B Mengen und sei $F \subseteq A \times B$ eine Relation. Die Relation F heißt *Abbildung* oder *Funktion* von A nach B, wenn gilt

$$\forall x \in A \;\; \exists^! y \in B \colon (x, y) \in F.$$

Eine Abbildung kann man sich auch als eine *Zuordnungsvorschrift* denken, die jedem Element a von A ein eindeutig bestimmtes Element von B zuordnet. Daher schreiben wir $F \colon A \to B$ und bezeichnen das eindeutig bestimmte Element aus B, das mit $x \in A$ in Relation steht, mit $F(x)$. Das Element $F(x)$ nennen wir auch den *Wert* von F bei x oder das *Bild* des Elements x. Gilt $y = F(x)$, so sagen wir auch, x wird auf y *abgebildet* und schreiben dafür $x \mapsto y$. Ein Element $x \in A$ mit $F(x) = y$ bezeichnet man auch als ein *Urbild* von y. Die Menge A heißt *Definitionsbereich*, die Menge B *Bildbereich* oder *Wertebereich* von F.

Für eine Abbildung $F\colon A \to B$ heißt

$$\mathrm{Bild}(F) := \{y \in B \mid \exists x \in A\colon F(x) = y\}$$

das *Bild* der Abbildung F. Gilt $\mathrm{Bild}(F) = B$, stimmt also das Bild mit dem Bildbereich überein, dann nennt man die Abbildung F *surjektiv*. Eine Abbildung $F\colon A \to B$ ist genau dann surjektiv, wenn jedes $y \in B$ mindestens ein Urbild besitzt.

Für eine Teilmenge $M \subseteq A$ nennt man

$$F(M) := \{y \in B \mid \exists x \in M\colon F(x) = y\}$$

das *Bild* der Teilmenge M. Es gilt $\mathrm{Bild}(F) = F(A)$.

Für eine Teilmenge $N \subseteq B$ nennt man

$$F^{-1}(N) := \{x \in A \mid F(x) \in N\}$$

das *Urbild* der Teilmenge N. Es gilt $F^{-1}(B) = A$. Gilt für jede einelementige Teilmenge $\{y\}$ von B, dass das Urbild $F^{-1}(y)$ aus höchstens einem Element besteht, so heißt die Abbildung *injektiv*.

Beispiel 1.5.1. Die Abbildung $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist weder injektiv noch surjektiv, denn jede positive Zahl y aus $\mathbb{R}$ hat zwei Urbilder (nämlich $\pm\sqrt{y}$); für negative Zahlen ist die Menge der Urbilder leer.

Eine Abbildung heißt *bijektiv*, wenn sie sowohl injektiv als auch surjektiv ist. Genau dann, wenn eine Abbildung $F\colon A \to B$ bijektiv ist, gibt es dazu eine *Umkehrabbildung*, d.h. eine Abbildung $G\colon B \to A$, so dass $G(F(a)) = a$ für alle $a \in A$ und $F(G(b)) = b$ für alle $b \in B$ gilt, wie man sich leicht überlegt. Wenn eine Umkehrabbildung existiert, ist sie eindeutig bestimmt.

Beispielsweise ist die Abbildung $F\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x + 1$ bijektiv und $G\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x - 1$ die zugehörige Umkehrabbildung.

Bei einer endlichen Menge nennt man die Anzahl der Elemente die *Mächtigkeit* der Menge. Allgemeiner nennt man zwei Mengen *gleichmächtig*, wenn es eine bijektive Abbildung zwischen ihnen gibt. Gibt es eine surjektive Abbildung $M \to N$ zwischen zwei Mengen, so sagt man, dass die Mächtigkeit von M größer gleich der Mächtigkeit von N ist.

Ist X eine Menge so bezeichnen wir mit id_X die *identische Abbildung*

$$\mathrm{id}_X\colon X \to X, \quad x \mapsto x,$$

die jedes Element von X wieder auf sich selbst abbildet. Sind $F\colon A \to B$ und $G\colon B \to C$ zwei Abbildungen, so ist durch

$$G \circ F\colon A \to C, \quad a \mapsto G(F(a)),$$

die *Verkettung* von F mit G definiert.

Für eine Abbildung $F\colon A \to B$ gilt

$$F \circ \mathrm{id}_A = F \quad \text{und} \quad \mathrm{id}_B \circ F = F.$$

Sind zwei Abbildungen gegeben, die eine Menge in sich abbilden, so erhält man im Allgemeinen verschiedene Abbildungen, je nachdem, in welcher Reihenfolge man verkettet: Seien die beiden Funktionen $f, g\colon \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = x^2$, $g(x) = x + 1$. Dann gilt $(f \circ g)(x) = x^2 + 2x + 1$, aber $(g \circ f)(x) = x^2 + 1$.

Die Verkettung von Abbildungen ist jedoch *assoziativ*: Sind $F\colon A \to B$, $G\colon B \to C$ und $H\colon C \to D$ drei Abbildungen, dann gilt:

$$(H \circ G) \circ F = H \circ (G \circ F).$$

In der Tat, wendet man die Abbildungen auf den beiden Seiten des Gleichheitszeichens auf ein Element $a \in A$ an, so erhält man für die linke Seite

$$((H \circ G) \circ F)(a) = (H \circ G)(F(a)) = H(G(F(a)))$$

und für die rechte Seite

$$(H \circ (G \circ F))(a) = H((G \circ F)(a)) = H(G(F(a))).$$

Ist $f\colon M \to N$ eine Abbildung und $A \subseteq M$ eine Teilmenge, dann definiert

$$f|_A\colon A \to N, \quad f|_A(a) := f(a)$$

eine Abbildung von A nach N, diese Abbildung wird die *Einschränkung* der Abbildung f auf A genannt. Es gilt, dass die Einschränkung einer injektiven Abbildung stets injektiv ist.

Kapitel 2

Gruppen, Ringe, Körper

2.1 Gruppen

In diesem Kapitel werden wir ein paar der wichtigsten *algebraischen Strukturen* kennenlernen. Grob gesprochen handelt es sich dabei um Rechen- oder Zahlbereiche, also Mengen, auf denen Rechenoperationen wie z.B. Addition und Multiplikation erklärt sind. Solche Rechenoperationen kann man wie folgt einführen: Sei M eine Menge. Eine *Verknüpfung* oder *binäre Operation* auf M ist eine Abbildung $M \times M \to M$. Eine Verknüpfung weist also einem Paar von Elementen aus M wieder ein Element aus M zu, so wie die *Addition* von natürlichen Zahlen einem Paar (m, n) von natürlichen Zahlen die Summe $m + n$ zuweist.

Definition 2.1.1. Sei G eine Menge mit einer Verknüpfung $*\colon G \times G \to G$, $(a, b) \mapsto a * b$. Die Menge G, zusammen mit der Verknüpfung $*$, heißt *Gruppe*, wenn die folgenden drei Eigenschaften gelten:

(G1) Die Verknüpfung ist *assoziativ*, d.h. es gilt $(a * b) * c = a * (b * c)$ für alle $a, b, c \in G$.

(G2) Es gibt ein *neutrales Element*, d.h. es gibt ein $e \in G$, so dass für alle $a \in G$

gilt $a * e = e * a = a$.

(G3) Es gibt zu jedem Element ein *inverses Element*, d.h. für alle $a \in G$ gibt es ein $b \in G$, so dass gilt $a * b = b * a = e$.

Hat die Verknüpfung $*$ zusätzlich die Eigenschaft der *Kommutativität*

(G4) Für alle $a, b \in G$ gilt: $a * b = b * a$,

so wird die Gruppe G eine *abelsche* (oder *kommutative*) *Gruppe* genannt.

Die Eigenschaften (G1), (G2) und (G3) werden auch *Gruppenaxiome* genannt. Die Verknüpfung $*$ wird oft auch einfach als *Multiplikation* der Gruppe bezeichnet, das Verknüpfungssymbol lässt man oft weg, d.h. man schreibt ab für $a*b$. Assoziativität einer Verknüpfung bedeutet, dass man bei der Verknüpfung von drei Elementen die Klammerung weglassen kann, denn $a(bc) = (ab)c$ und man schreibt dafür einfach abc.

Beispiele für Gruppen sind: Die Menge der ganzen Zahlen $\mathbb{Z}$ mit der Addition als Verknüpfung; hier hat die Null die Rolle eines neutralen Elementes. Die Menge $\mathbb{Q} \setminus \{0\}$ der rationalen Zahlen, die nicht gleich Null sind, mit der Multiplikation als Verknüpfung. Die Menge $\mathbb{R}^n$ der n-Tupel reeller Zahlen mit der *komponentenweisen Addition*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

als Verknüpfung. Ein neutrales Element ist das n-Tupel $(0, 0, \ldots, 0)$, das nur Nullen enthält. Diese drei Gruppen sind abelsche Gruppen.

Um den Umgang mit den Gruppenaxiomen zu demonstrieren, zeigen wir, dass es in jeder Gruppe nur ein Element e geben kann, das die Eigenschaften eines neutralen Elements hat: Sei $f \in G$ ein Element mit der Eigenschaft, das $f * a = a$ für alle $a \in G$ gilt. Dann folgt $e = f * e = f$.

Außerdem gilt für jedes $a \in G$, dass das inverse Element zu a eindeutig bestimmt ist. Deshalb darf man a^{-1} für das eindeutig bestimmte inverse Element zu a schreiben. Dies zeigt man wie folgt: Seien $b, c \in G$ mit $b * a = e$ und $a * c = e$. Daraus schließen wir

$$b = b * e = b * (a * c) = (b * a) * c = e * c = c.$$

Ist M eine Menge, so können wir Gesamtheit aller bijektiven Abbildungen

$$\mathrm{Sym}(M) := \{f\colon M \to M \mid f \text{ ist bijektiv}\}$$

betrachten. Diese ist eine Gruppe mit der Verkettung von Abbildungen als Verknüpfung.

Beweis. Wir bemerken zuerst, dass die Verkettung von zwei bijektiven Abbildungen wieder bijektiv ist; damit ist die Verknüpfung auf $\mathrm{Sym}(M)$ wohldefiniert. Die Gruppeneigenschaft (G1) gilt, da die Verkettung von Abbildungen assoziativ ist. (G2) gilt, da id_M die Eigenschaften eines neutralen Elementes hat. (G3) gilt, da die Umkehrabbildung einer bijektiven Abbildung ein inverses Element zu dieser Abbildung ist, denn es gilt $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_M$. □

Wie bereits gezeigt, ist die Verkettung von Abbildungen im Allgemeinen nicht kommutativ. Die Gruppe $\mathrm{Sym}(M)$ heißt *symmetrische Gruppe* über M, ihre Elemente heißen *Permutationen* von M. Im Spezialfall $M = \{1, 2, \dots, n\}$ setzt man $\mathrm{Sym}(n) := \mathrm{Sym}(M)$. Die Gruppe $\mathrm{Sym}(n)$ hat $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n$ Elemente.

Eine Permutation $\mathrm{Sym}(n)$ kann man definieren, in dem man die Bilder der Zahlen von 1 bis n einzeln angibt, z.B. ist durch

$$1 \mapsto 3, \quad 2 \mapsto 1, \quad 3 \mapsto 2$$

ein Element von $\mathrm{Sym}(3)$ festgelegt. Oft wird eine Permutation auch in Form einer Tabelle angegeben, dabei stehen in der ersten Zeile die Zahlen von 1 bis n und in der zweite Zeile darunter jeweils die Bilder. Die obige Permutation würde also als

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

notiert werden.

Gruppen treten bei der Beschreibung von Symmetrien auf. Beispielsweise kann man bei einem gleichseitigen Dreieck in der Ebene alle Drehungen um den Mittelpunkt betrachten, die das Dreieck wieder mit sich zur Deckung bringen. Dies sind die Drehungen um $0°$, $120°$ und $240°$. Diese Drehungen bilden, aufgefasst, als Abbildungen der Ebene, eine Gruppe mit der Verkettung von Abbildungen als Verknüpfung. Ähnlich kann man bei einem Würfel im dreidimensionalen Raum die Menge aller Drehungen betrachten, die den Würfel in sich überführen, die sogenannte *Würfelgruppe*. Man überlegt sich leicht, das diese Gruppe 24 Elemente hat.

Definition 2.1.2. Sei G eine Gruppe. Eine nichtleere Teilmenge $H \subseteq G$ wird als Untergruppe von G bezeichnet, falls die folgenden beiden Eigenschaften gelten:

(UG1) Die Teilmenge H ist abgeschlossen unter der Gruppenverknüpfung, d.h. es gilt $a * b \in H$ für alle $a, b \in H$.

(UG2) Die Teilmenge H ist abgeschlossen unter der Inversenbildung, d.h. es gilt $a^{-1} \in H$ für alle $a \in H$.

Es folgt sofort aus diesen beiden Eigenschaften, dass jede Untergruppe von G das neutrale Element von G enthält, denn eine nichtleere Teilmenge $H \subseteq G$ enthält mindestens ein Element $a \in H$ und nach (UG2) somit auch a^{-1}. Dann folgt aber mit (UG1), dass $a * a^{-1} = e$ in H liegt.

Daraus ist ersichtlich, dass H mit der Verknüpfung $*$, eingeschränkt auf Paare von Elementen in H, selbst eine Gruppe ist, mit demselben neutralen Element wie G.

Ein Beispiel für eine Untergruppe ist die Teilmenge

$$\{\ldots, -4, -2, 0, 2, 4, \ldots\}$$

der geraden Zahlen in der Gruppe der ganzen Zahlen $\mathbb{Z}$ mit der Addition als Verknüpfung. Die ungeraden Zahlen

$$\{\ldots, -3, -1, 1, 3, \ldots\}$$

hingegen sind keine unter der Addition abgeschlossene Teilmenge von $\mathbb{Z}$, denn die Summe zweier ungerader Zahlen ist gerade.

Ein weiteres Beispiel für eine Untergruppe ist die Teilmenge von $\mathrm{Sym}(3)$, die aus den beiden Permutationen

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

besteht. Das linke erste Element ist die identische Abbildung auf $\{1, 2, 3\}$. Um also die Abgeschlossenheit bezüglich der Verkettung von Abbildungen einzusehen, genügt es zu festzustellen, dass das zweite Element mit sich selbst verknüpft, wieder die identische Abbildung auf $\{1, 2, 3\}$ ergibt, was auch zeigt, dass dieses Element sein eigenes Inverses ist.

Für jede Gruppe ist $\{e\}$ eine Untergruppe, die *triviale Untergruppe*.

Definition 2.1.3. Seien G und H zwei Gruppen. Eine Abbildung $f\colon G \to H$ heißt *Gruppenhomomorphismus*, wenn $f(ab) = f(a)f(b)$ für alle $a, b \in G$ gilt.

Definition 2.1.4. Sei $f\colon G \to H$ ein Gruppenhomomorphismus und sei $e \in H$ das neutrale Element. Dann heißt

$$\ker(f) := f^{-1}(\{e\}) = \{g \in G \mid f(g) = e\}$$

der *Kern* von f.

Beispiel 2.1.5. Sei $G = \mathbb{R}^2$ mit der komponentenweisen Addition als Gruppenverknüpfung und $(0, 0)$ als neutralem Element. Dann ist $f\colon \mathbb{R}^2 \to \mathbb{R}$, definiert durch $f(x_1, x_2) := x_1 - x_2$ ein Gruppenhomomorphismus. Denn es gilt

$$\begin{aligned} f((x_1, x_2) + (y_1, y_2)) &= f(x_1 + y_1, x_2 + y_2) = \\ &= x_1 + y_1 - x_2 - y_2 = f(x_1, x_2) + f(y_1, y_2). \end{aligned}$$

Das Urbild des neutralen Elements in $\mathbb{R}$ ist die Diagonale

$$\{(x, x) \mid x \in \mathbb{R}\}.$$

Sie ist eine Untergruppe von $\mathbb{R}^2$ mit der komponentenweisen Addition, wie man leicht nachprüft.

Satz 2.1.6. *Sei $f\colon G \to H$ ein Gruppenhomomorphismus. Dann bildet f das neutrale Element von G auf das neutrale Element von H ab. Außerdem sind* $\ker(f)$ *und* $\mathrm{Bild}(f)$ *Untergruppen von G bzw. H.*

Beweis. Sei $e \in G$ das neutrale Element. Dann folgt $f(e) = f(ee) = f(e)f(e)$, was zeigt, das $f(e)$ das (eindeutig bestimmte) neutrale Element von H ist. Dies zeigt auch, dass das neutrale Element von G in $\ker(f)$ liegt.

Seien $a, b \in \ker(f)$. Dann gilt $f(ab) = f(a)f(b) = ee = e$, also liegt ab im Kern von f. Da ein Gruppenhomomorphismus das neutrale Element von G auf das von H abbildet, folgt $e\, f(a^{-1}) = f(a)f(a^{-1}) = f(aa^{-1}) = f(e) = e$, was zeigt, das $a^{-1} \in \ker(f)$ gilt.

Seien nun $a, b \in \mathrm{Bild}(f)$. Dann gibt es $g, k \in G$, so dass $a = f(g)$ und $b = f(k)$. Damit folgt $ab = f(g)f(k) = f(gk) \in \mathrm{Bild}(f)$. Außerdem folgt $af(g^{-1}) = f(g^{-1})a = f(e) = e$, was zeigt, dass $f(g^{-1}) \in \mathrm{Bild}(f)$ das zu a inverse Element ist. $\square$

2.2 Ringe

Definition 2.2.1. Sei R eine Menge auf der zwei Verknüpfungen $+$ und $\cdot$, genannt *Addition* und *Multiplikation*, gegeben sind. Wir sagen, R ist ein *Ring*, wenn folgende Eigenschaften gelten.

R ist eine abelsche Gruppe bezüglich der Addition:

(R1) Die Addition ist assoziativ: $\forall a, b, c \in R\colon (a + b) + c = a + (b + c)$.

(R2) Es gibt ein neutrales Element 0 bezüglich der Addition:
$\forall a \in R\colon a + 0 = a$.

(R3) Zu jedem Element in R gibt es ein inverses Element bezüglich der Addition: $\forall a \in R \;\; \exists b \in R\colon a + b = 0$.

(R4) Die Addition ist kommutativ: $\forall a, b \in R\colon a + b = b + a$.

Die Multiplikation ist assoziativ:

(R5) $\forall a, b, c \in R\colon (a \cdot b) \cdot c = a \cdot (b \cdot c)$.

Es gelten die Distributivgesetze:

(R6) $\forall a, b, c \in R\colon a \cdot (b + c) = a \cdot b + a \cdot c$.

(R7) $\forall a, b, c \in R\colon (a + b) \cdot c = a \cdot c + b \cdot c$.

Gibt es zusätzlich noch ein neutrales Element bezüglich der Multiplikation, nennt man den Ring einen *Ring mit Eins* oder *unitären Ring*. Ist zusätzlich die Multiplikation kommutativ, sagt man der Ring ist *kommutativ*.

Als ein Beispiel für das Arbeiten mit den Axiomen eines Rings beweisen wir:

Satz 2.2.2. *In jedem Ring R und für alle $x \in R$ gilt:* $0 \cdot x = x \cdot 0 = 0$.

Beweis. Wir zeigen nur $0 \cdot x = 0$, der andere Fall geht analog. Dazu wenden wir das Distributivgesetz (R7) an:

$$0 \cdot x + 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x.$$

Addieren wir nun zu beiden Seiten der Gleichung $-(0 \cdot x)$, so erhalten wir $0 \cdot x = 0$. $\square$

Man beachte, dass es in einem Ring im Allgemeinen keine inversen Elemente bezüglich der Multiplikation gibt. Ein Beispiel für einen kommutativen Ring mit Eins sind die ganzen Zahlen mit der üblichen Addition und Multiplikation. Ein weiteres Beispiel für einen kommutativen Ring mit Eins ist der Ring der Polynome mit reellen Koeffizienten in einer (formalen) Variablen, d.h. von Ausdrücken der Form

$$a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0,$$

wobei die Koeffizienten $a_0, \ldots, a_n$ reelle Zahlen sind, die wie gewohnt addiert und multipliziert werden. Das größte n, so dass der Koeffizient a_n ungleich Null ist, heißt der *Grad* des Polynoms. Das Nullpolynome hat den Grad $-\infty$.

2.3 Matrizen

Eine $n \times m$-*Matrix* mit Einträgen in einem Ring R ist ein rechteckiges Schema

$$\begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix},$$

von Elementen $a_{ij} \in R$ die in einer Tabelle mit n Zeilen und m Spalten angeordnet sind. Man sagt, i ist der Zeilenindex und j der Spaltenindex des Eintrags a_{ij}. Die Menge der $n \times m$-Matrizen mit Einträgen in R wird mit $M_{n,m}(R)$ bezeichnet. Die obige Matrix wird auch mit (a_{ij}) bezeichnet.

Sind A und B zwei Matrizen mit Einträgen aus dem selben Ring R und stimmt die Spaltenanzahl n von A mit der Zeilenanzahl von B überein, so kann man das *Matrizenprodukt* AB als die Matrix definieren, deren Eintrag in der i-ten Zeile und j-ten Spalte durch

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

gegeben ist.

Beispiel 2.3.1.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ -2 & 3 \end{pmatrix}$$

Das Matrizenprodukt AB ist

$$AB = \begin{pmatrix} 1 \cdot 1 - 2 \cdot 1 - 3 \cdot 2 & 1 \cdot 0 + 2 \cdot 1 + 3 \cdot 3 \\ 4 \cdot 1 - 5 \cdot 1 - 6 \cdot 2 & 4 \cdot 0 + 5 \cdot 1 + 6 \cdot 3 \end{pmatrix} = \begin{pmatrix} -7 & 11 \\ -13 & 23 \end{pmatrix}.$$

Wie man leicht nachprüft, gilt hier $AB \neq BA$, die Matrixmultiplikation ist nicht kommutativ.

Für zwei Matrizen $A, B \in M_{n,m}(R)$ kann man die Summe $A + B$ durch die komponentenweise Addition erklären, d.h. der Eintrag in der i-ten Zeile und j-ten Spalte von $A + B$ ist

$$a_{ij} + b_{ij}.$$

Die Menge $M_{n,m}$ der $n \times m$-Matrizen mit Einträgen in R wird mit dieser Addition zu einer abelschen Gruppe mit der *Nullmatrix*

$$\begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

als neutralem Element.

Im Spezialfall von *quadratischen Matrizen* $n = m$, schreibt man für $M_{n,n}(R)$ kürzer $M_n(R)$. Man kann je zwei quadratische Matrizen in $M_n(R)$ sowohl mit der Addition von Matrizen als auch mit Hilfe der Matrixmultiplikation verknüpfen und die Menge $M_n(R)$ wird damit zu einem Ring mit Eins. (Die dazu benötigten Eigenschaften der Matrizenmultiplikation, nämlich Assoziativität und Distributivität rechnet man anhand der Definition nach, siehe Hilfssatz 2.3.2.) Das Einselement ist dabei gegeben durch die $n \times n$-*Einheitsmatrix*

$$E := E_n := \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix},$$

d.h. der Matrix, bei der alle Einträge auf der *Diagonalen* (dort, wo Zeilen- und Spaltenindex übereinstimmen) gleich 1 sind und alle anderen Einträge gleich 0.

Hier benutzen wir die Vereinbarung, dass Matrixeinträge, die gleich Null sind, leer gelassen werden können.

Die Einträge der Einheitsmatrix kann man auch durch das sogenannte *Kronecker-Symbol*

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

beschreiben, es gilt $E_n = (\delta_{ij})$.

Wir haben also gesehen, dass die Menge $M_n(R)$ der $n \times n$-Matrizen mit Einträgen in einem Ring R selbst einen Ring bildet. Der Matrizenring besitzt *Nullteiler* falls $n \geq 2$, d.h. es gibt Elemente $a, b \in M_n(R)$, so dass $ab = 0$, obwohl weder a noch b gleich Null sind, denn es gilt z.B.:

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

Hilfssatz 2.3.2. *Die Multiplikation von Matrizen ist assoziativ, d.h. es gilt*

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

für je drei Matrizen $A \in M_{m,n}(R)$, $B \in M_{n,p}(R)$ und $C \in M_{p,q}(R)$. Außerdem gelten die Distributivgesetze $A(B+C) = AB+AC$ und $(A+B)C = AC+BC$.

Beweis. Durch direktes Nachrechnen. Seien $A = (a_{ij}) \in M_{m,n}(R)$, $B = (b_{ij} \in M_{np}(R)$, $C = (c_{ij}) \in M_{pq}(R)$. Dann ist der ij-te Eintrag von AB gegeben durch $\sum_{k=1}^{n} a_{in} b_{nj}$. Damit erhalten wir für den ij-ten Eintrag von $(AB)C$ den Ausdruck

$$\sum_{\ell=1}^{p} \sum_{k=1}^{n} (a_{ik} b_{k\ell}) c_{\ell j}$$

Andererseits ist der ij-te Eintrag von BC gegeben durch $\sum_{\ell=1}^{p} b_{i\ell} c_{\ell j}$ und wir erhalten wir den ij-ten Eintrag von $A(BC)$

$$\sum_{k=1}^{n} \sum_{\ell=1}^{p} a_{ik} (b_{k\ell} c_{\ell j})$$

Damit folgt nun das Assoziativgesetz für die Matrizenmultiplikation aus dem Assoziativität der Multiplikation und der Kommutativität der Addition im Ring R. Die beiden Distributivgesetze rechnet man auf ähnliche Weise nach. $\square$

2.4 Körper

Definition 2.4.1. Ein kommutativer Ring mit Eins heißt *Körper*, wenn jedes Element außer das neutrale Element 0 der Addition ein multiplikatives Inverses besitzt.

Wir fassen sämtliche Axiome, die in einem Körper gelten, noch einmal zusammen: Ein *Körper* ist eine Menge K, auf der zwei Verknüpfungen $+$ und $\cdot$, genannt *Addition* und *Multiplikation*, gegeben sind, für die folgende Eigenschaften gelten:

(K1) Die Addition ist assoziativ:
$\forall a, b, c \in K\colon (a + b) + c = a + (b + c)$.

(K2) Es gibt ein neutrales Element 0 bezüglich der Addition:
$\forall a \in K\colon a + 0 = a$.

(K3) Zu jedem Element in K gibt es ein inverses Element bezüglich der Addition:
$\forall a \in K \;\; \exists b \in K\colon a + b = 0$.

(K4) Die Addition ist kommutativ:
$\forall a, b \in K\colon a + b = b + a$.

(K5) Die Multiplikation ist assoziativ:
$\forall a, b, c \in K\colon (a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(K6) Es gibt ein neutrales Element 1 bezüglich der Multiplikation:
$\forall a \in K\colon a \cdot 1 = a$.

(K7) Zu jedem Element in $K \setminus \{0\}$ gibt es ein inverses Element bezüglich der Multiplikation:
$\forall a \in K \;\; \exists b \in K\colon a \cdot b = 1$.

(K8) Die Multiplikation ist kommutativ:
$\forall a, b \in K\colon a \cdot b = b \cdot a.$

(K9) Es gilt das Distributivgesetz:
$\forall a, b, c \in K\colon a \cdot (b + c) = a \cdot b + a \cdot c.$

Ein Körper K ist somit eine abelsche Gruppe bezüglich der Addition und $K \setminus \{0\}$ ist eine abelsche Gruppe bezüglich der Multiplikation.

Ein Körper ist ein Zahlenbereich, in dem wir addieren, subtrahieren, multiplizieren und dividieren können, so wie wir dies beim Rechnen mit rationalen oder reellen Zahlen gewohnt sind.

Beispiele 2.4.2. (i) Die rationalen Zahlen, also die Menge der Brüche

$$\mathbb{Q} = \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z}, n \in \mathbb{N} \setminus \{0\} \right\},$$

mit der üblichen Addition und Multiplikation, bilden einen Körper.

(ii) Der Körper $\mathbb{R}$ der reellen Zahlen, dargestellt durch unendliche Dezimalbrüche.

(iii) Die *komplexen Zahlen* $\mathbb{C}$ sind definiert als die Menge der Paare (a, b) von reellen Zahlen mit komponentenweiser Addition und der Multiplikation

$$(a, b) \cdot (c, d) := (ac - bd, ad + bc)$$

Führt man die Abkürzung i für das Element $(0, 1)$ ein, so kann man jede komplexe Zahl (a, b) in der Form $a+ib$ schreiben und mit komplexen Zahlen wie von reellen Zahlen gewohnt rechnen, wenn man zusätzlich $i^2 = -1$ beachtet.

Man definiert $\Re((a, b)) = a$ als den *Realteil* und $\Im((a, b)) = b$ als den *Imaginärteil* einer komplexen Zahl (a, b). Für jede komplexe Zahl $z = (a, b)$ definiert man die zu z konjugiert komplexe Zahl

$$\bar{z} = (a, -b) = \Re(z) - i\,\Im(z).$$

Die komplexen Zahlen bilden einen Körper. Das multiplikative Inverse eines Elements $z \neq 0$ ist gegeben durch $z^{-1} = \frac{\bar{z}}{z\bar{z}}$. Die komplexen Zahlen enthalten die reellen Zahlen als *Teilkörper*.

(iv) Der Körper $\mathbb{Z}_2$ mit zwei Elementen. Er ist definiert durch die folgende Additions- und Multiplikationstabelle.

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad\qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Man überlegt sich leicht, dass aus den Körperaxiomen bereits folgt, dass die Addition und die Multiplikation in einem Körper mit zwei Elementen wie in den beiden Tabellen beschrieben sein muss. Mit anderen Worten: Je zwei Körper mit zwei Elementen sind isomorph.

Satz 2.4.3. *In einem Körper gibt es keine Nullteiler: Sei K ein Körper und seien $a, b \in K$. Falls $ab = 0$, dann gilt $a = 0$ oder $b = 0$.*

Beweis. Angenommen, es gilt $ab = 0$, aber $a \neq 0$. Dann gibt es ein multiplikatives Inverses a^{-1} von a und wir erhalten

$$0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$$

Also folgt $b = 0$. □

Der Ring der reellen Polynome in einer Variablen ist ein kommutativer Ring mit Eins. In ihm gibt es keine Nullteiler, aber er ist dennoch kein Körper. Beides folgt aus der Tatsache, dass sich beim Multiplizieren zweier Polynome sich die Grade der Polynome addieren.

Kapitel 3

Vektorräume

3.1 Vektorräume

Definition 3.1.1. Sei K ein Körper. Ein *Vektorraum* über dem Körper K, auch K-Vektorraum genannt, ist eine Menge V, zusammen mit einer Addition $+$, so dass V bezüglich $+$ eine abelsche Gruppe bildet, zusammen mit einer weiteren Abbildung

$$K \times V \to V, \qquad (\lambda, v) \mapsto \lambda \cdot v,$$

genannt *Skalarmultiplikation*, für die folgende Eigenschaften gelten.

(V1) $\forall \lambda, \mu \in K \;\; \forall v \in V \colon (\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$.

(V2) $\forall \lambda, \mu \in K \;\; \forall v \in V \colon (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$.

(V3) $\forall \lambda \in K \;\; \forall v, w \in V \colon \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

(V4) $\forall v \in V \colon 1 \cdot v = v$.

Die Elemente der Menge V werden als *Vektoren* bezeichnet. Die Elemente des Körpers K werden auch *Skalare* genannt.

Das wichtigste Beispiel für einen K-Vektorraum ist die Menge K^n der n-Tupel von Zahlen aus K mit der kompentenweisen Addition und der wie folgt definierten Skalarmultiplikation.

$$\lambda \cdot (x_1, \ldots, x_n) := (\lambda x_1, \ldots, \lambda x_n).$$

Zum Beispiel kann man sich den $\mathbb{R}^3$ als den dreidimensionalen Anschauungsraum und den $\mathbb{R}^2$ als die Zeichenebene, beide durch kartesische Koordinaten parametrisiert, vorstellen. Dabei entspricht der komponentenweise Addition die anschauliche Vektoraddition durch Hintereinandersetzen von Vektorpfeilen. Der Skalarmultiplikation eines Vektors v mit einer reellen Zahl entspricht dabei der zentrischen Streckung das Vektors v um den Faktor λ.

Ein weiteres Beispiel ist die Menge der reellen Funktionen. Wir definieren

$$V := \{f : \mathbb{R} \to \mathbb{R}\}.$$

als die Menge aller Abbildungen $f : \mathbb{R} \to \mathbb{R}$. Zwei solche Funktionen lassen sich addieren, indem man sie wie folgt überlagert:

$$(f + g)(t) := f(t) + g(t),$$

d.h. indem man also punktweise die Summe ihre Funktionswerte bei einer reellen Zahl t zum Wert der zu definierenden Funktion $f + g$ bei t erklärt. Die Skalarmultiplikation erklärt man ebenfalls punktweise:

$$(\lambda \cdot f)(t) := \lambda f(t).$$

Dieses Beispiel lässt sich abwandeln, indem man etwa V durch die Menge aller stetigen Funktionen ersetzt. Sind nämlich f und g zwei stetige Funktionen und λ eine reelle Zahl, so sind $f + g$ und λf wieder stetige Funktionen. Auch die Menge der differenzierbaren Funktionen $\mathbb{R} \to \mathbb{R}$ bildet einen $\mathbb{R}$-Vektorraum.

Als ein Beispiel für das Arbeiten mit den Vektorraumaxiomen zeigen wir, dass in jedem Vektorraum

$$0 \cdot v = 0$$

gilt, wobei hier die auf der linken Seite des Gleichheitszeichens stehende Null die Null im Körper K bezeichnet, während die auf der rechten Seite des Gleichheitszeichens stehende Null das Nullelement im Vektorraum meint. Wir formen um:

$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v,$$

was zeigt, dass $0 \cdot v$ das neutrale Element in der abelschen Gruppe V ist. Weitere Aussagen, die in jedem Vektorraum gelten, sind

$$(-\lambda)v = -(\lambda v) = \lambda(-v) \quad \text{für alle} \quad \lambda \in K, v \in V$$

und

$$\lambda \cdot 0 = 0,$$

wobei hier $0 \in V$ gemeint ist. Es ist eine gute Übung, diese Aussagen zu beweisen.

3.2 Untervektorräume

Definition 3.2.1. Sei V ein K-Vektorraum. Eine nichtleere Teilmenge $U \subseteq V$ heißt *Untervektorraum* von V, wenn gilt

(UV1) Für alle $v, w \in U$ ist auch die Summe $v + w$ wieder ein Element von U.

(UV2) Für alle Vektoren $u \in U$ und alle Körperelemente $\lambda \in K$ is auch λu wieder ein Element von U.

Ein Unterraum ist also eine Teilmenge eines Vektorraums, die nichtleer ist, sowie abgeschlossen unter Addition und unter Multiplikation mit beliebigen Skalaren.

Es lässt sich zeigen, dass eine Teilmenge eines Vektorraums genau dann selbst wieder ein Vektorraum mit der auf die Teilmenge eingeschränkten Vektoraddition und Skalarmultiplikation ist, wenn sie ein Untervektorraum ist.

Ein Beispiel für einen Untervektorraum im Anschauungsraum $V := \mathbb{R}^3$ ist etwa die xy-Ebene

$$U := \{(x, y, 0) \mid x, y \in \mathbb{R}\}.$$

Wie man sich sofort überlegt, ist diese Teilmenge des $\mathbb{R}^3$ nicht leer, abgeschlossen unter Addition und Multiplikation mit beliebigen Skalaren.

In der Ebene $\mathbb{R}^2$ ist jede Gerade, die durch den Ursprung verläuft, ein Untervektorraum.

Ein weiteres Beispiel ist die Teilmenge der stetigen Funktionen der Menge aller Abbildungen $f : \mathbb{R} \to \mathbb{R}$.

In jedem Vektorraum V sind $\{0\}$ und V selbst Untervektorräume von V.

Da ein Untervektorraum $U \subseteq V$ nach Definition nicht leer ist, enthält er mindestens ein Element v von V. Dann folgt aus $0 \cdot v = 0$, dass nach(UV2) das Nullelement von V in U enthalten ist. Jeder Untervektorraum enthält also die Null. Mit anderen Worten, $\{0\} \subseteq V$ ist der kleinste Untervektorraum in jedem Vektorraum V ist.

Satz 3.2.2. *Sei V ein Vektorraum und sei $\{U_j \mid j \in J\}$ eine Menge von Untervektorräumen von V. Dann ist die Schnittmenge*

$$\bigcap_{j \in J} U_j$$

wieder ein Untervektorraum von V.

Beweis. Da jeder Untervektorraum U_j von V das Nullelement von V enthält, ist 0 auch in der Schnittmenge $\bigcap_{j \in J} U_j$ enthalten, diese ist somit nicht leer. Seien nun $v, w \in \bigcap_{j \in J} U_j$. Dann gilt für jedes $j \in J$, dass sowohl v als auch w in U_j enthalten sind. Da $U_j \subseteq V$ ein Untervektorraum ist, gilt $v + w \in U_j$ für jedes $j \in J$. Dies zeigt, dass $v + w$ in der Schnittmenge aller U_j enthalten ist. Sei nun $\lambda \in K$. Wegen der Untervektorraumeigenschaft der U_j liegt λv in jedem U_j und somit in der Schnittmenge aller U_j. $\square$

Die Vereinigung von Untervektorräumen ist im Allgemeinen kein Untervektorraum. Als Gegenbeispiel betrachten wir die Ebene $\mathbb{R}^2$. Die x-Achse und die y-Achse sind Untervektorräume, also liegen die beiden Vektoren $(1, 0)$ und $(0, 1)$ in der Vereinigung der beiden Achsen. Aber ihre Summe $(1, 0) + (0, 1) = (1, 1)$ liegt außerhalb des Achsenkreuzes. Somit ist die Vereinigung der x-Achse und der y-Achse im $\mathbb{R}^2$ kein Untervektorraum von $\mathbb{R}^2$.

3.3 Spann und Linearkombinationen

Definition 3.3.1. Sei V ein Vektorraum und $M \subseteq V$ eine Teilmenge. Der *Spann* oder *Aufspann* von M ist definiert als die Schnittmenge aller Untervektorräume, die M enthalten.

Der Spann einer Teilmenge $M \subseteq V$ ist also der kleinste Untervektorraum von V, der die Teilmenge M enthält. Es gilt stets $M \subseteq \mathrm{span}(M)$. Der Spann der leeren Menge ist $\{0\}$.

Sei $V = \mathbb{R}^2$ und sei M die Vereinigung der x-Achse mit der y-Achse. Dann enthält M jeden Vektor $(x, 0)$ mit $x \in \mathbb{R}$ und jeden Vektor $(0, y)$ mit $y \in \mathbb{R}$. Somit enthält $\mathrm{span}(M)$ als Untervektorraum von $\mathbb{R}^2$ auch jedes Paar (x, y) und es folgt $\mathrm{span}(M) = \mathbb{R}^2$.

Definition 3.3.2. Sei V ein Vektorraum und seien $v_1, \ldots, v_n$ endlich viele Vektoren in V. Dann heißt jedes Element

$$\lambda_1 v_1 + \ldots + \lambda_n v_n$$

mit $\lambda_1, \ldots, \lambda_n \in K$ eine *Linearkombination* der Vektoren $v_1, \ldots, v_n$. Die Zahlen $\lambda_1, \ldots, \lambda_n$ nennt man die *Koeffizienten* der Linearkombination. Sind alle Koeffizienten gleich Null, so sagt man, es handelt sich um die *triviale Linearkombination* der $v_1, \ldots, v_n$.

Eine Linearkombination über null Elemente ist eine leere Summe, sie stellt den Nullvektor dar.

Hilfssatz 3.3.3. *Sei V ein Vektorraum und $M \subseteq V$ eine Teilmenge. Dann ist die Menge aller Linearkombinationen*

$$\{\lambda_1 v_1 + \ldots + \lambda_n v_n \,|\, n \in \mathbb{N} \cup \{0\}, v_1, \ldots, v_n \in M, \lambda_1, \ldots, \lambda_n \in K\},$$

gebildet über endliche Teilmengen von M, ein Untervektorraum von V.

Beweis. Die Menge ist nicht leer, denn sie enthält $0 \in V$, dargestellt als Linearkombination von null Elementen. Sind

$$\lambda_1 v_1 + \ldots + \lambda_n v_n \quad \text{und} \quad \mu_1 w_1 + \ldots + \mu_m w_m$$

zwei Elemente aus der Menge, so enthält die Menge auch deren Summe, die Linearkombination

$$\lambda_1 v_1 + \ldots + \lambda_n v_n + \mu_1 w_1 + \ldots + \mu_m w_m.$$

Ebenso ist

$$\lambda \cdot (\lambda_1 v_1 + \ldots + \lambda_n v_n) = (\lambda\lambda_1) v_1 + \ldots + (\lambda\lambda_n) v_n$$

wieder enthalten. □

Satz 3.3.4. *Sei V ein Vektorraum und sei M eine Teilmenge. Dann besteht die Teilmenge* $\mathrm{span}(M) \subseteq V$ *aus allen Linearkombinationen $\lambda_1 v_1 + \ldots + \lambda_n v_n$ mit $n \in \mathbb{N} \cup \{0\}$, $v_1, \ldots, v_n \in M$, $\lambda_1, \ldots, \lambda_n \in K$.*

Beweis. Die Menge aller Linearkombination, die über endliche Teilmengen von Elementen in M gebildet werden, enthält M und ist nach dem Hilfssatz 3.3.3 ein Untervektorraum von V. Außerdem sind diese Linearkombinationen in jedem Untervektorraum von V enthalten, der M enthält, somit auch in $\mathrm{span}(M)$. Es gilt aber, dass $\mathrm{span}(M)$ der kleinste solche Untervektorraum ist. □

Aus dem Satz folgt sofort, dass für jeden Vektor $v \in V$ gilt:

$$\mathrm{span}(\{v\}) = \{\lambda v \mid \lambda \in k\}.$$

Für den Aufspann eines Vektors v schreibt man statt $\mathrm{span}\{v\}$ oft auch Kv.

3.4 Summen von Untervektorräumen

Definition 3.4.1. Seien $U_1, U_2, \ldots U_k$ Untervektorräume von V. Dann heißt der Untervektorraum

$$U_1 + U_2 + \ldots + U_k := \mathrm{span}(U_1 \cup U_2 \cup \ldots \cup U_k)$$

die *Summe* der Untervektorräume $U_1, U_2, \ldots, U_k$.

Die Summe der Untervektorräume $U_1, U_2, \ldots U_k$ besteht aus allen Vektoren v in V, die sich darstellen lassen als

$$v = u_1 + u_2 + \ldots + u_n, \quad u_1 \in U_1, u_2 \in U_2, \ldots, u_n \in U_n,$$

wie man sich sofort überlegt.

Beispiel 3.4.2. Betrachten wir den $\mathbb{R}^3$ und die drei Untervektorräume

$$U_1 := \mathrm{span}((1,0,0)), \quad U_2 := \mathrm{span}((0,1,0)), \quad U_3 := \mathrm{span}((1,1,0)).$$

Dann ist

$$U_1 + U_2 + U_3 = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0\}.$$

Definition 3.4.3. Seien $U_1, U_2, \ldots U_k$ Untervektorräume von V. Dann heißt ein Untervektorraum U von V die *direkte Summe* von $U_1, U_2, \ldots, U_k$ wenn sich jeder Vektor auf *genau eine* Weise als Summe

$$u_1 + u_2 + \ldots + u_n, \quad u_1 \in U_1, u_2 \in U_2, \ldots, u_n \in U_n,$$

darstellen lässt. Ist U die direkte Summe der Untervektorräume $U_1, U_2, \ldots, U_k$, so schreibt man $U = U_1 \oplus U_2 \oplus \ldots \oplus U_k$.

Beispiel 3.4.4. Bei dem Beispiel 3.4.2 handelt es sich um eine Summe von Untervektorräumen, die *keine* direkte Summe ist. Denn z.B. der Vektor $(1, 1)$ lässt als $(1, 0) + (0, 1) + (0, 0)$ oder auch als $(0, 0) + (0, 0) + (1, 1)$ darstellen. Dagegen ist $\mathbb{R}^2$ die direkte Summe der beiden Untervektorräume, die durch die x-Achse bzw. die y-Achse gegeben sind.

Satz 3.4.5. *Seien U_1, U_2 Untervektorräume von V. Genau dann ist die Summe $U = U_1 + U_2$ direkt, wenn U die beiden folgenden Eigenschaften hat:*

(i) $U = U_1 + U_2$*,*

(ii) $U_1 \cap U_2 = \{0\}$.

Beweis. Angenommen, die Summe $U = U_1 + U_2$ ist direkt und $u \in U_1 \cap U_2$. Dann gilt insbesondere $u \in U_1$ und $u \in U_2$, woraus auch $-u \in U_2$ folgt. Dann hat der Nullvektor in U die beiden Darstellungen $0 + 0$ und $u - u$, wobei jeweils der erste Summand in U_1 ist und der zweite in U_2. Wegen der eindeutigen Darstellbarkeit von 0 folgt $u = 0$.

Umgekehrt, sei nun $U = U_1 + U_2$ und gelte $U_1 \cap U_2 = \{0\}$. Angenommen für $u \in U$ gelte $u = u_1 + u_2 = u_1' + u_2'$ mit $u_1, u_1' \in U_1$ und $u_2, u_2' \in U_2$. Dann folgt $u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2 = \{0\}$. Also gelten $u_1 = u_1'$ und $u_2 = u_2'$. $\square$

Satz 3.4.5 gilt *nicht* für 3 oder mehr Untervektorräume, wie das Beispiel 3.4.2 zeigt: Dort gilt $U_1 \cap U_2 \cap U_3 = \{0\}$, aber nicht $U_1 + U_2 + U_3 = U_1 \oplus U_2 \oplus U_3$.

3.5 Lineare Unabhängigkeit und Basen

Definition 3.5.1. Sei V ein Vektorraum. Seien $v_1, \ldots, v_n$ Vektoren von V. Die Vektoren $v_1, \ldots, v_n$ heißen *linear abhängig*, wenn sich der Nullvektor als eine nichttriviale Linearkombination der Vektoren $v_1, \ldots v_n$ darstellen lässt, d.h. wenn es eine Linearkombination $\lambda_1 v_1, \ldots, \lambda_n v_n = 0$ gibt, wobei mindestens ein Koeffizient λ_j ungleich Null ist. Andernfalls heißen die Vektoren $v_1, \ldots, v_n$ *linear unabhängig*.

Allgemeiner, sei V ein Vektorraum und M eine Teilmenge von V. Die Teilmenge M heißt *linear unabhängig*, falls für jede endliche Teilmenge $\{v_1, \ldots, v_n\}$ von M die Vektoren $v_1, \ldots, v_n$ linear unabhängig sind. Andernfalls heißt die Teilmenge von V *linear abhängig*.

Beispiel 3.5.2. *Sei $V = \mathbb{R}^2$ und seien $v_1 = (1,0)$, $v_2 = (0,1)$, $v_3 = (1,1)$. Dann ist sind Vektoren v_1, v_2, v_3 linear abhängig, denn die Linearkombination $v_1 + v_2 - v_3$ ergibt den Nullvektor. Die Vektoren v_1, v_2 sind jedoch linear unabhängig, denn falls eine Linearkombination $\lambda_1 v_1 + \lambda_2 v_2 = (\lambda_1, \lambda_2)$ den Nullvektor ergibt, so folgt $\lambda_1 = \lambda_2 = 0$, d.h. es handelt sich um die triviale Linearkombination.*

Aus der Definition der linearen Unabhängigkeit folgt sofort, dass eine einelementige Teilmenge $\{v\}$ eines Vektorraum genau dann linear unabhängig ist, wenn $v \neq 0$ gilt, denn in einem Vektorraum gilt $\lambda \cdot v = 0$ für $\lambda \in K$ und $v \in V$ genau wenn $\lambda = 0$ oder $v = 0$. Ebenfalls sofort aus der Definition folgt, dass jede Teilmenge einer linear unabhängigen Menge von Vektoren eines Vektorraums wieder linear unabhängig ist. Insbesondere ist die leere Menge linear unabhängig.

Satz 3.5.3. *Seien $v_1, \ldots, v_n$ Vektoren eines Vektorraums. Genau dann wenn die Vektoren $v_1, \ldots, v_n$ linear unabhängig sind, lässt sich jedes Element aus ihrem Aufspann $\operatorname{span}\{v_1, \ldots, v_n\}$ als Linearkombination von $v_1, \ldots, v_n$ mit eindeutig bestimmten Koeffizienten darstellen.*

Beweis. Sei $v \in \operatorname{span}\{v_1, \ldots, v_n\}$, d.h. es gibt Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Angenommen, die Darstellung von v ist nicht eindeutig, d.h. es gibt $\mu_1, \ldots, \mu_n \in K$ so, dass $v = \mu_1 v_1 + \ldots + \mu_n v_n$, wobei es ein $j \in \{1, \ldots, n\}$ gibt, so dass

$\lambda_j \neq \mu_j$. Dann ist

$$0 = (\lambda_1 - \mu_1)v_1 + \ldots + (\lambda_n - \mu_n)v_n$$

Eine Linearkombination, bei der nicht alle Koeffizienten Null sind.

Umgekehrt, sei $0 = \alpha_1 v_1 + \ldots + \alpha_n v_n$ eine Linearkombination, bei der nicht alle Koeffizienten gleich Null sind. Dann ist

$$v = (\lambda_1 + \alpha_1)v_1 + \ldots + (\lambda_n + \alpha_n)v_n.$$

eine weitere Darstellung von v. $\square$

Bemerkung 3.5.4. Dass sich jeder Vektor aus dem Spann von $v_1, \ldots, v_n$ als Linearkombination dieser Vektoren mit eindeutig bestimmten Koeffizienten schreiben lässt, kann man in Formeln auch so ausdrücken:

$$\operatorname{span}\{v_1, \ldots, v_n\} = Kv_1 \oplus \ldots \oplus Kv_n.$$

Hilfssatz 3.5.5. *Sei B eine linear unabhängige Teilmenge eines Vektorraums V und sei $v \in V$ ein Vektor. In diesem Fall ist die Teilmenge*

$$B \cup \{v\}$$

genau dann linear unabhängig, wenn v nicht in $\operatorname{span}(B)$ *liegt.*

Beweis. Angenommen, $v \in \operatorname{span}(B)$. Dann gibt es $v_1, \ldots, v_k \in B$ und Koeffizienten $\lambda_1, \ldots, \lambda_k$, so dass $v = \lambda_1 v_1 + \ldots + \lambda_k v_k$. Also ist

$$(-1)v + \lambda_1 v_1 + \ldots + \lambda_k v_k$$

eine nichttriviale Linearkombination, die Null ergibt.

Umgekehrt, sei $B \cup \{v\}$ linear abhängig, d.h. es gibt eine nichttriviale Linearkombination

$$\lambda_0 v + \lambda_1 v_1 \ldots + \lambda_k v_k = 0.$$

mit Vektoren $v_1, \ldots, v_k \in B$. Dann ist der Koeffizient λ_0 verschieden von Null, denn sonst wären $v_1, \ldots, v_k \in B$ linear abhängig. Also gilt

$$v = -\frac{\lambda_1}{\lambda_0}v_1 - \ldots - \frac{\lambda_k}{\lambda_0}v_k.$$ $\square$

Vereinbarung: Ab jetzt schreiben wir die Elemente von K^n stets als *Spaltenvektoren*, d.h. als $n \times 1$-Matrizen, also

$$K^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \; x_1, \dots, x_n \in K \right\}.$$

Beispiel 3.5.6. Die Menge der folgenden Vektoren wird als *Standardbasis* oder *kanonische Basis* von K^n bezeichnet:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Diese bilden in der Tat eine Basis, denn jeder Vektor in K^n lässt sich als Linearkombination

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 e_1 + \dots + \lambda_n e_n$$

mit eindeutig bestimmten Koeffizienten darstellen.

Definition 3.5.7. Sei V ein Vektorraum. Eine Teilmenge $E \subseteq V$ heißt *Erzeugendensystem* von V, wenn $\operatorname{span}(E) = V$ gilt. Eine Teilmenge $B \subseteq V$ heißt *Basis* von V, wenn sie ein linear unabhängiges Erzeugendensystem von V ist.

Satz 3.5.8. *Für eine Teilmenge B eines Vektorraums V sind äquivalent:*

(i) B ist eine Basis von V.

(ii) B ist ein minimales Erzeugendensystem von V.

(iii) B ist eine maximal linear unabhängige Teilmenge von V.

Erläuterung: Dass B ein *minimales Erzeugendensystem* von V ist, bedeutet: Es gilt $V = \operatorname{span}(B)$ und für alle $v \in B$ ist $B \setminus \{v\}$ kein Erzeugendensystem von

V. Dass B eine *maximal linear unabhängige Teilmenge* von V ist, bedeutet: B ist linear unabhängig ist und für alle $v \in V \setminus B$ ist $B \cup \{v\}$ nicht linear unabhängig.

Beweis von Satz 3.5.8. Wir zeigen zunächst, dass (i) sowohl (ii) als auch (iii) impliziert. Angenommen, B ist eine Basis. Dann ist B eine linear unabhängige Teilmenge von V und ein Erzeugendensystem von V. Angenommen, B wäre kein minimales Erzeugendensystem von V, d.h. es gibt ein Element $v \in B$, so dass $B \setminus \{v\}$ ein Erzeugendensystem ist. Dann gilt $v \in \operatorname{span}(B \setminus \{v\})$, was nach Hilfssatz 3.5.5 im Widerspruch dazu steht, dass $B = (B \setminus \{v\}) \cup \{v\}$ linear unabhängig ist. Angenommen, B wäre keine maximal linear unabhängige Menge, d.h. es gibt ein Element $v \in V \setminus B$, so dass $B \cup \{v\}$ linear unabhängig ist. Nach Hilfssatz 3.5.5 liegt dann v nicht im Spann von B, im Widerspruch dazu, dass B ein Erzeugendensystem von V ist.

Gelte nun (ii), sei also B ein minimales Erzeugendensystem von V. Wir zeigen, dass B linear unabhängig ist. Angenommen, B ist linear abhängig, d.h. es gibt eine nichttriviale Linearkombination

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

mit $v_1, \ldots, v_n \in B$. Da ein Koeffizient, sagen wir λ_j, ungleich Null ist, lässt sich v_j als Linearkombination der Vektoren $v_1, \ldots v_{j-1}, v_{j+1}, \ldots, v_n$ ausdrücken, was zeigt, dass $B \setminus \{v_j\}$ bereits ein Erzeugendensystem von V ist. Also war B kein minimales Erzeugendensystem.

Gelte (iii), sei also B maximal linear unabhängig. Sei $v \in V$ ein beliebiger Vektor. Dann ist $B \cup \{v\}$ eine linear abhängige Teilmenge von V und es gibt darin eine nichttriviale Linearkombination, die Null ergibt. In dieser Linearkombination kommt v mit einem Koeffizienten ungleich Null vor, denn sonst wäre B bereits linear abhängig gewesen. Es gibt also $\lambda_0, \lambda_1, \ldots, \lambda_n \in K$, so dass

$$\lambda_0 v + \lambda_1 v_1 + \ldots + \lambda_n v_n = 0,$$

und $\lambda_0 \neq 0$. Dann folgt

$$v = -\frac{\lambda_1}{\lambda_0} v_1 - \ldots - \frac{\lambda_n}{\lambda_0} v_n.$$

Also gilt $v \in \operatorname{span}(B)$ und wir haben gezeigt, dass B ein Erzeugendensystem von V ist. $\square$

Satz 3.5.9. *Jeder Vektorraum besitzt eine Basis.*

Beweis. Wir führen den Beweis nur für den Fall, dass ein Vektorraum V *endlich erzeugt* ist, d.h. für den Fall, in dem es eine endliche Teilmenge $E \subseteq V$ gibt, so dass $\operatorname{span}(E) = V$ ist. Wir zeigen per Induktion, dass jedes endliche Erzeugendensystem E ein minimales Erzeugendensystem enthält. Denn entweder ein Erzeugendensystem E ist bereits minimal, oder es gibt einen Vektor $v \in E$, so dass $E \setminus \{v\}$ immer noch ein Erzeugendensystem ist. $\square$

Bemerkung 3.5.10. Für den Fall, dass V nicht endlich erzeugt ist, benötigt man zum Beweis das sogenannte *Zornsche Lemma* oder auch das dazu äquivalente *Auswahlaxiom*. Beides sind starke Axiome aus der Mengenlehre. Wir wollen darauf hier nicht eingehen, sondern stattdessen einfach Satz 3.5.9, der zu diesen beiden Axiomen sogar äquivalent ist, als richtig annehmen.

Satz 3.5.11 (Basisauswahlsatz)**.** *Sei E ein endliches Erzeugendensystem von einem Vektorraum V. Dann enthält E eine Basis von V.*

Beweis. Dies wurde im Beweis von Satz 3.5.9 gezeigt. $\square$

3.6 Dimension

Satz 3.6.1 (Basisergänzungssatz)**.** *Sei V ein Vektorraum und sei die endliche Menge von Vektoren $E = \{w_1, \ldots, w_m\}$ eine Erzeugendensystem von V. Sei B eine linear unabhängige Teilmenge von V. Dann gibt es Vektoren $v_1, \ldots, v_k \in E$, so dass $B \cup \{v_1, \ldots, v_k\}$ eine Basis von V ist.*

Beweis. Wir führen den Beweis durch Induktion. Falls B bereits ein Erzeugendensystem von V ist, ist es bereits eine Basis und es ist nichts mehr zu zeigen. Wir zeigen, dass es andernfalls einen Vektor $w \in E \setminus \mathrm{span}(B)$ gibt, so dass $B \cup \{w\}$ immer noch linear unabhängig ist.

Falls B kein Erzeugendensystem von V ist, gibt es einen Vektor $v \in V$, so dass $v \notin \mathrm{span}(B)$. Da E ein Erzeugendensystem ist, gibt es Koeffizienten $\mu_1, \ldots, \mu_m \in K$, so dass $\mu_1 w_1 + \ldots + \mu_m w_m = v$. Nun gibt es mindestens einen Vektor w_j, so dass w_j nicht in $\mathrm{span}(B)$ enthalten ist, denn sonst wäre auch v in $\mathrm{span}(B)$ enthalten. Da $w := w_j$ nicht im Spann on B liegt, folgt aus Hilfssatz 3.5.5, dass $B \cup \{w\}$ linear unabhängig ist.

Dasselbe Argument können wir nun auf die linear unabhängige Menge $B \cup \{w\}$ anstelle von B noch einmal anwenden und induktiv fortfahren. Spätestens nach dem m-ten Schritt (wenn man nämlich alle Vektoren aus E zu B hinzugefügt hat) ist die linear unabhängige Menge B ein Erzeugendensystem von V, also eine Basis, geworden. □

Satz 3.6.2 (Austauschlemma)**.** *Seien $A = \{v_1, \ldots, v_n\}$ und $B = \{w_1, \ldots, w_m\}$ Basen des Vektorraums V. Sei $i \in \{1, \ldots, n\}$. Dann gibt es einen Index $j \in \{1, \ldots, n\}$, so dass auch*

$$\{v_1, \ldots, v_{i-1}, w_j, v_{i+1}, \ldots, v_n\}$$

eine Basis von V ist.

Beweis. Setze $A' = \{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$. Dann gibt es einen Vektor $w_j \in B$, so dass $w_j \notin \mathrm{span}(A')$, denn wenn alle Elemente der Basis B im Spann von A lägen, wäre A' schon ein Erzeugendensystem von V, im Widerspruch dazu, dass die Basis A ein minimales Erzeugendensystem von V ist. Daraus folgt mit Hilfssatz 3.5.5, dass die Menge $A' \cup \{w_j\}$ linear unabhängig ist.

Es bleibt zu zeigen, dass $A' \cup \{w_j\}$ auch ein Erzeugendensystem von V ist. Wir können $w_j \neq v_i$ annehmen, denn sonst ist nichts zu zeigen. Jedenfalls wissen wir, dass

$$A \cup \{w_j\} = \{v_1, \ldots, v_n, w_j\}$$

ein Erzeugendensystem von V ist. Nun können wir den Basisergänzungssatz anwenden, der uns sagt, dass wir die linear unabhängige Menge $A' \cup \{w_j\}$ mit Vektoren aus dem Erzeugendensystem $A \cup \{w_j\}$ zu einer Basis ergänzen können. Da alle Vektoren aus $A \cup \{w_j\}$ bis auf v_i bereits in $A' \cup \{w_j\}$ enthalten sind, gibt es dazu nur zwei Möglichkeiten: Entweder $A' \cup \{w_j\}$ ist selbst schon eine Basis oder es wird zu $A \cup \{w_j\}$ ergänzt. Aber A war bereits maximal linear unabhängig, also verbleibt nur die Möglichkeit, dass

$$A' \cup \{w_j\} = \{v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_n\}$$

eine Basis ist. $\square$

Hilfssatz 3.6.3. *Je zwei endliche Basen eines Vektorraums bestehen aus gleich vielen Elementen.*

Beweis. Angenommen, $A = \{v_1, \dots, v_n\}$ und $B = \{w_1, \dots, w_m\}$ sind zwei endliche Basen von V mit n bzw. m Elementen. Dann können wir nach dem Austauschlemma Schritt für Schritt jeden Vektor in A durch einen in B ersetzen. Damit bauen wir A zu einer Basis von V um, die aus n Vektoren besteht, von denen jeder ein Element von B ist. Falls $n < m$ gilt, ist diese Basis eine echte Teilmenge von B, im Widerspruch dazu, dass B ein minimales Erzeugendensystem von V ist. Dies zeigt $n \geq m$. Vertauscht man die Rollen von A und B, folgt $m \geq n$. $\square$

Hat ein Vektorraum ein endliche Basis, dann nennen wir die Anzahl der Elemente die *Dimension* des Vektorraums. Nach dem vorangehenden Hilfssatz ist diese Zahl wohldefiniert.

Satz 3.6.4. *Sei V ein Vektorraum, der eine Basis aus n Elementen besitzt. Dann gilt:*

(i) Jede Teilmenge von V mit $n+1$ oder mehr Elementen ist linear abhängig.

(ii) Jedes Erzeugendensystem von V enthält mindestens n Elemente.

(iii) Jedes Basis von V besteht aus n Elementen.

Beweis. Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V.

(i) Sei A eine linear unabhängige Teilmenge von V mit $n+1$ Elementen. Dann lässt sich A nach dem Basisergänzungssatz mit endlich vielen Elementen aus B zu einer endlichen Basis mit mehr als n Elementen ergänzen. Widerspruch zu Hilfssatz 3.6.3.

(ii) Sei E ein Erzeugendensystem von V mit m Elementen. Nach dem Basisauswahlsatz lässt sich aus E eine Basis aus E auswählen; diese hat m oder weniger Elemente. Aus (i) folgt dann, dass $n \leq m$ gilt.

(iii) Folgt aus direkt aus Hilfssatz 3.6.3 bzw. aus (i) und (ii).

$\square$

Eine Folgerung aus Satz 3.6.4 (i) ist, dass Untervektorräume von endlich-dimensionalen Vektorräumen selbst wieder endlich-dimensional sind, allgemeiner gilt folgendes.

Hilfssatz 3.6.5. *Sei V ein endlich-dimensionaler Vektorraum der Dimension n. Dann hat jeder* echte Untervektorraum *von V, d.h. jeder Untervektorraum $U \subseteq V$ mit $U \neq V$ eine echt kleiner Dimension als V, d.h. es gilt* $\dim(U) \leq \dim(V) - 1$. *Außerdem gilt: Jede linear unabhängige Teilmenge von V mit n Elementen ist eine Basis von V.*

Beweis. Sei $U \subseteq V$ ein Untervektorraum der Dimension n mit $U \neq V$ und sei $b_1, \dots, b_n$ eine Basis von U. Dann gibt es ein Element $v \in V \setminus U$. Nach Hilfssatz 3.5.5 sind dann $b_1, \dots, b_n, v$ linear unabhängig, Widerspruch zu Satz 3.6.4 (i). Um den zweiten Teil der Aussage zu beweisen, nehmen wir an, dass $v_1, \dots, v_n$ linear unabhängige Vektoren in V sind. Dann bilden diese eine Basis von $W := \operatorname{span}\{v_1, \dots, v_n\}$. Dann gilt $\dim(W) = n$ und nach dem ersten Teil folgt $V = W$. $\square$

Definition 3.6.6. Besitzt ein Vektorraum eine Basis mit endlich vielen Elementen, so nennen wir die Anzahl der Elemente die *Dimension* $\dim(V)$ des Vektorraums. Besitzt der Vektorraum keine endliche Basis, so setzen wir $\dim(V) = \infty$.

Der Begriff Dimension ist wohldefiniert, denn gibt es eine endliche Basis, so haben alle Basen eines Vektorraums nach Teil (iii) des vorangehenden Satzes gleich viele Elemente. Wir bemerken, dass $\dim(K^n) = n$ gilt, denn die Standardbasis $e_1, \ldots, e_n$ von K^n hat n Elemente.

Beispiel 3.6.7. Als Beispiel für einen unendlich-dimensionalen Vektorraum betrachten wir die Menge der reellen Polynome

$$t \mapsto a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0,$$

aufgefasst als Funktionen $\mathbb{R} \to \mathbb{R}$. Diese Menge bildet einen Vektorraum über $\mathbb{R}$. Es gilt, dass die *Monome* $\{ t \mapsto t^k \mid k \in \mathbb{N} \cup \{0\} \}$ eine linear unabhängige Teilmenge bilden, denn die einzige Linearkombination von Monomen, die die Nullfunktion ergibt, ist die triviale. Daraus folgt mit Satz 3.6.4, dass es in diesem Vektorraum keine endliche Basis geben kann. Es folgt, dass auch der $\mathbb{R}$-Vektorraum der Funktionen $\mathbb{R} \to \mathbb{R}$ unendlich-dimensional ist.

Satz 3.6.8 (Dimensionsformel für Untervektorräume)**.** *Sei V ein Vektorraum und seien U_1, U_2 endlich-dimensionale Untervektorräume von V. Dann gilt*

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

Beweis. Als Untervektorraum des endlich-dimensionalen Vektorraums U_1 ist $U_1 \cap U_2$ ebenfalls endlich-dimensional. Es existiert eine Basis $b_1, \ldots, b_s$ von $U_1 \cap U_2$, die wir nach dem Basisergänzungssatz zu einer Basis $b_1, \ldots, b_m$ mit $m \geq s$ von U_1 ergänzen können. Andererseits können wir die Basis $b_1, \ldots, b_s$ auch zu einer Basis $b_1, \ldots, b_2, v_1, \ldots, v_k$ von U_2 ergänzen. Um den Beweis abzuschließen, zeigen wir, dass

$$b_1, \ldots, b_m, v_1, \ldots, v_k$$

eine Basis von $U_1 + U_2$ ist. Offensichtlich bilden diese Vektoren ein Erzeugendensystem von $U_1 + U_2$. Es bleibt zu zeigen, dass sie linear unabhängig sind. Angenommen, es gibt eine Linearkombination, die Null ergibt:

$$0 = \lambda_1 b_1 + \ldots + \lambda_m b_m + \lambda_{m+1} v_1 + \ldots + \lambda_{m+k} v_k.$$

Dann folgt

$$\lambda_1 b_1 + \ldots + \lambda_m b_m = -(\lambda_{m+1} v_1 + \ldots + \lambda_{m+k} v_k) \in U_1 \cap U_2.$$

Aber die Vektoren $v_1, \ldots, v_k$ liegen außerhalb von $U_1 \cap U_2$, was $\lambda_{m+1} = \ldots = \lambda_{m+k} = 0$ zeigt. Damit folgt auch $\lambda_1 = \ldots = \lambda_m = 0$. Es gilt also

$$\begin{aligned} \dim(U_1 + U_2) &= m + k = m + (s + k) - s = \\ &= \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2). \end{aligned}$$

□

Kapitel 4

Lineare Abbildungen

4.1 Lineare Abbildungen

Definition 4.1.1. Seien V und W zwei Vektorräume über demselben Körper K. Eine Abbildung $f\colon V \to W$ heißt *linear*, wenn sie folgende Eigenschaften hat.

(L1) $\forall v_1, v_2 \in V\colon f(v_1 + v_2) = f(v_1) + f(v_2)$.

(L2) $\forall v \in V, \lambda \in K\colon f(\lambda v) = \lambda f(v)$.

Bemerkung 4.1.2. Man kann die beiden Eigenschaften wie folgt zu einer zusammenfassen:

(L) $\forall v_1, v_2 \in V, \lambda_1 \lambda_2 \in K\colon f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2)$.

Für jede lineare Abbildung f gilt $f(0) = 0$, denn $f(0) = f(0 \cdot v) = 0 \cdot f(v) = 0$.

Beispiele 4.1.3. (i) Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x, y) = 2x - y$. Dann ist f eine lineare Abbildung, denn es gilt:

$f((x_1, y_1) + (x_2, y_2)) = f(x_1 + x_2, y_1 + y_2) = 2x_1 + 2x_2 - y_1 - y_2 = f(x_1, y_1) + f(x_2, y_2)$ und $f(\lambda x, \lambda y) = 2\lambda x - \lambda y = \lambda f(x, y)$.

(ii) Sei $A \in M_{m,n}(K)$ eine $m \times n$-Matrix mit Einträgen in dem Körper K. Dann ist $f\colon K^n \to K^m$ eine lineare Abbildung, wobei die Elemente in K^n als $n \times 1$-Matrizen, d.h. als *Spaltenvektoren*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

aufgefasst werden. Die Additivität (L1) folgt aus dem Distributivgesetz für die Matrixmultiplikation: $A(v+w) = Av + Aw$. Die Verträglichkeit mit der Skalarmultiplikation (L2) $A(\lambda v) = \lambda A(v)$ folgt direkt aus der Definition der Matrixmultiplikation.

(iii) Sei V der Vektorraum der differenzierbaren reellen Funktionen $f\colon \mathbb{R} \to \mathbb{R}$. Die Differentiation $f \mapsto f'$ ist eine lineare Abbildung, denn die Ableitung von Funktionen verhält sich additiv: $(f+g)' = f' + g'$ und es gilt $(\lambda \cdot f)' = \lambda \cdot (f')$.

(iv) Die Verkettung von linearen Abbildungen ist wieder linear. Seien nämlich $f\colon V \to W$ und $g\colon W \to X$ lineare Abbildungen sind und $v_1, v_2 \in V$, $\lambda_1, \lambda_2 \in K$. Dann folgt

$$g(f(\lambda_1 v_1 + \lambda_2 v_2)) = g(\lambda_1 f(v_1) + \lambda_2 f(v_2)) = \lambda_1 g(f(v_1)) + \lambda_2 g(f(v_2)).$$

Definition 4.1.4. Seien V und W Vektorräume über demselbe Körper K und sei $f\colon V \to W$ eine lineare Abbildung. Dann sind auch die folgenden Sprechweisen üblich.

(i) Man sagt, f ist ein *Vektorraumhomomorphismus* oder *linearer Homomorphismus.*

(ii) Falls f injektiv ist, sagt man, f ist ein *Vektorraummonomorphismus* oder *linearer Monomorphismus*.

(iii) Falls f surjektiv ist, sagt man, f ist ein *Vektorraumepimorphismus* oder *linearer Epimorphismus*.

(iv) Falls f bijektiv ist, sagt man, f ist ein *Vektorraumisomorphismus* oder *linearer Isomorphismus*. Falls es einen Vektorraumisomorphismus zwischen den beiden K-Vektorräumen V und W gibt, sagt man, die Vektorräume V und W sind *isomorph*, in Zeichen: $V \cong W$.

Falls $V = W$ gilt, d.h. falls $f\colon V \to V$ eine lineare Abbildung ist, benutzt man auch folgende Sprechweisen.

(i) Man sagt, f ist ein *Vektorraumendomorphismus* oder *linearer Endomorphismus*.

(ii) Falls f bijektiv ist, sagt man, f ist ein *Vektorraumautomorphismus* oder *linearer Automorphismus*.

Hilfssatz 4.1.5. *Seien V und W Vektorräume und $f\colon V \to W$ ein Vektorraumisomorphismus. Dann ist auch $f^{-1}\colon W \to V$ linear, insbesondere ist f^{-1} ein Vektorraumisomorphismus.*

Beweis. Wir zeigen, dass $f^{-1}\colon W \to V$ linear ist. Seien $w_1, w_2 \in W$ zwei Vektoren und $\lambda_1, \lambda_2 \in K$. Da f bijektiv ist, gibt es Vektoren $v_1, v_2 \in V$ mit $f(v_1) = w_1$ und $f(v_2) = w_2$. Es folgt $f^{-1}(\lambda_1 w_1 + \lambda_2 w_2) = f^{-1}(\lambda_1 f(v_1) + \lambda_2 f(v_2)) = f^{-1}(f(\lambda_1 v_1 + \lambda_2 v_2)) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 f^{-1}(w_1) + \lambda_2 f^{-1}(w_2)$. □

Definition 4.1.6. Sei $f\colon V \to W$ eine lineare Abbildung. Die Teilmenge

$$f^{-1}(\{0\}) = \{v \in V \mid f(v) = 0\},$$

bestehend aus allen Vektoren in V, die auf den Nullvektor abgebildet werden, heißt der *Kern* von f und wird mit $\ker(f)$ bezeichnet.

Satz 4.1.7. *Der Kern einer linearen Abbildung $f\colon V \to W$ ist ein Untervektorraum von V. Das Bild einer linearen Abbildung $f\colon V \to W$ ist ein Untervektorraum von W. Eine lineare Abbildung ist genau dann injektiv, wenn ihr Kern nur aus dem Nullvektor besteht.*

Beweis. Der Kern einer linearen Abbildung $f\colon V \to W$ enthält immer den Nullvektor von V und ist somit nicht leer. Seien $v_1, v_2 \in \ker(f)$, und $\lambda \in K$. Dann gilt $f(v_1 + v_2) = f(v_1) + f(v_2) = 0$ und außerdem $f(\lambda v) = \lambda f(v) = 0$.

Wegen $f(0) = 0$ enthält $\mathrm{Bild}(f)$ zumindest den Nullvektor in W und ist somit nicht leer. Seien $v_1, v_2 \in V$, und $\lambda \in K$. Dann gilt $f(v_1) + f(v_2) = f(v_1 + v_2) \in \mathrm{Bild}(f)$ und außerdem $\lambda f(v) = f(\lambda v) \in \mathrm{Bild}(f)$.

Angenommen, der Kern von f enthält einen von Null verschiedenen Vektor v. Dann gilt $f(v) = 0 = f(0)$, was zeigt, dass f nicht injektiv ist. Umgekehrt, falls f nicht injektiv ist, gibt es zwei Vektoren $v_1, v_2 \in V$ mit $v_1 \neq v_2$ und $f(v_1) = f(v_2)$. Daraus folgt dass $0 = f(v_1) - f(v_2) = f(v_1 - v_2)$. Also ist $v_1 - v_2$ ein von Null verschiedener Vektor im Kern von f. □

4.2 Lineare Abbildungen und Basen

Satz 4.2.1. *Seien V und W zwei K-Vektorräume, sei $f\colon V \to W$ eine lineare Abbildung und sei B eine Basis von V. Die Abbildung f ist genau dann ein Vektorraumisomorphismus, falls f die Basis B bijektiv auf eine Basis von W abbildet.*

Beweis. Angenommen f ist ein Isomorphismus. Dann müssen wir zeigen, dass $f(B)$ linear unabhängig und ein Erzeugendensystem von W ist. Angenommen, $f(B)$ ist linear abhängig. Dann gibt es $w_1, \ldots, w_k \in f(B)$ und $\lambda_1, \ldots, \lambda_k \in K$, so dass

$$0 = \lambda_1 w_1 + \ldots + \lambda_k w_k,$$

wobei nicht alle λ_j gleich Null sind. Auf beide Seiten dieser Gleichung können wir die (nach Satz 4.1.5) lineare Abbildung f^{-1} anwenden und erhalten somit

$$0 = \lambda_1 f^{-1}(w_1) + \ldots + \lambda_k f^{-1}(w_k),$$

Daraus folgt, dass die Elemente $f^{-1}(w_1), \ldots, f^{-1}(w_k) \in B$ linear abhängig sind, Widerspruch! Sei $w \in W$ beliebig. Da B eine Basis von V ist, gibt es $v_1, \ldots, v_n \in$

B und $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$f^{-1}(w) = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Daraus folgt durch Anwenden von f auf beide Seiten der Gleichung, dass

$$w = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)$$

gilt.

Sei nun $f(B)$ eine Basis von W und $f|_B \colon B \to f(B)$ bijektiv. Wir zeigen, dass f injektiv und surjektiv ist. Angenommen f ist nicht injektiv. Dann gibt es nach Satz 4.1.7 einen von Null verschiedenen Vektor $v \in \ker(f)$. Da B eine Basis ist, gibt es paarweise verschiedene Vektoren $b_1, \ldots, b_n \in B$ und Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$, die nicht alle gleich Null sind, so dass $v = \lambda_1 b_1 + \ldots + \lambda_n b_n$. Nun folgt

$$0 = f(v) = f(\lambda_1 b_1 + \ldots + \lambda_n b_n) = \lambda_1 f(b_1) + \ldots + \lambda_n f(b_n),$$

im Widerspruch zur linearen Unabhängigkeit der $f(b_1), \ldots, f(b_n)$. Sei nun $w \in W$. Dann gibt es $b_1, \ldots, b_n \in V$ und Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$ so, dass $w = \lambda_1 f(b_1) + \ldots + \lambda_n f(b_n)$. Es folgt $w = f(\lambda_1 b_1 + \ldots + \lambda_n b_n) \in \mathrm{Bild}(f)$. $\square$

Satz 4.2.2. *Seien V und W zwei Vektorräume und $f \colon V \to W$ eine lineare Abbildung.*

(i) Sei E ein Erzeugendensystem von V. Dann ist die Abbildung f bereits durch ihre Werte auf den Elementen von E eindeutig festgelegt. Insbesondere ist also eine lineare Abbildung durch ihre Werte auf einer Basis eindeutig festgelegt.

(ii) Umgekehrt, gibt man die Werte auf einer Basis B von V vor, so gibt es eine eindeutig bestimmte lineare Abbildung, die auf der Basis B die vorgegebenen Werte annimmt, d.h. sei B eine Basis von V und $f_0 \colon B \to W$ eine Abbildung, dann gibt es eine eindeutig bestimmte lineare Abbildung $f \colon V \to W$, so dass $f(b) = f_0(b)$ für alle $b \in B$ gilt.

Beweis. (i) Sei $v \in V$ beliebig. Wir müssen zeigen, dass $f(v)$ bereits durch die Werte von f auf E eindeutig festgelegt ist. Da E ein Erzeugendensystem von V ist, gibt es Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$, so dass $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ gilt für $v_1, \ldots, v_n \in E$. Dann gilt $f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)$, was zeigt, dass $f(v)$ bereits durch die Werte von f auf den Elementen von E eindeutig festgelegt ist.

(ii) Wir müssen nur die Existenz der Abbildung f zeigen, die Eindeutigkeit folgt aus (i). Wir definieren eine Abbildung $f\colon V \to W$ wie folgt. Für jedes $v \in V$ gibt es endlich viele Vektoren $b_1, \ldots, b_n \in B$ und zugehörige Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$, so dass $v = \lambda_1 b_1 + \ldots \lambda_n b_n$ gilt. Aus Satz 3.5.3 folgt, dass es nur eine Möglichkeit gibt, v als Linearkombination endlich vieler Vektoren aus B darzustellen. Somit ist durch

$$f(v) := \lambda_1 f_0(b_1) + \ldots + \lambda_n f_0(b_n)$$

eine wohldefinierte Abbildung $f\colon V \to W$ gegeben. Es bleibt zu zeigen, dass f linear ist. Seien $v, w \in V$ und $\alpha, \beta \in K$. Dann lassen sich v und w beide als Linearkombination von Vektoren $b_1, \ldots, b_n \in B$ darstellen, $v = \lambda_1 b_1 + \ldots \lambda_n b_n$, $w = \mu_1 b_1 + \ldots \mu_n b_n$. Dann gilt

$$\begin{aligned} f(\alpha v + \beta b) &= f((\alpha\lambda_1 + \beta\mu_1)b_1 + \ldots + (\alpha\lambda_n + \beta\mu_n)b_n) = \\ &= (\alpha\lambda_1 + \beta\mu_1)f_0(b_1) + \ldots + (\alpha\lambda_n + \beta\mu_n)f_0(b_n) = \\ &= \alpha f(v) + \beta f(w). \end{aligned}$$

□

Definition 4.2.3. Der *Rang einer linearen Abbildung* ist die Dimension des Bildes.

Wir schreiben $\mathrm{rk}(f)$ für den Rang einer linearen Abbildung f. Nach Definition gilt $\mathrm{rk}(f) = \dim(\mathrm{Bild}(f))$. Nach Hilfssatz 3.6.5 ist eine lineare Abbildung $f\colon V \to W$ für einen endlich-dimensionalen Vektorraum W genau dann surjektiv, wenn $\mathrm{rk}(f) = \dim(W)$ gilt.

Satz 4.2.4 (Dimensionsformel für lineare Abbildungen)**.** *Sei V ein endlich-dimensionaler Vektorraum und sei $f\colon V \to W$ eine lineare Abbildung. Dann gilt*

$$\dim(V) = \mathrm{rk}(f) + \dim(\ker(f)).$$

Beweis. Da V endlich-dimensional ist, ist auch der Kern von f endlich dimensional. Es gibt also eine Basis $v_1, \ldots, v_m$ von $\ker(f)$. Nach dem Basisergänzungssatz können wir diese zu einer Basis $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ von V ergänzen. Es genügt jetzt, zu zeigen, dass $\mathrm{rk}(f) = n - m$ gilt. Dazu zeigen wir, dass $f(v_{m+1}), \ldots, f(v_n)$ eine Basis von $\mathrm{Bild}(f)$ ist. Sei $w \in \mathrm{Bild}(f)$, d.h. es gibt ein $v \in V$ mit $f(v) = w$. Dann gibt es Koeffizienten $\lambda_1, \ldots, \lambda_n \in K$, so dass $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ gilt. Daraus folgt $w = f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = \lambda_{m+1} f(v_{m+1}) + \ldots + \lambda_n f(v_n)$, wobei die letzte Gleichheit daraus folgt, dass $v_1, \ldots, v_m$ im Kern von f liegen. Dies zeigt, dass die Vektoren $f(v_{m+1}), \ldots, f(v_n)$ ein Erzeugendensystem von $\mathrm{Bild}(f)$ sind. Es bleibt zu zeigen, dass diese Vektoren linear unabhängig sind. Angenommen, es gilt

$$\lambda_{m+1} f(v_{m+1}) + \ldots + \lambda_n f(v_n) = 0$$

für $\lambda_{m+1}, \ldots, \lambda_n \in K$. Dann folgt

$$f(\lambda_{m+1} v_{m+1} + \ldots + \lambda_n v_n) = 0,$$

d.h. der Vektor $\lambda_{m+1} v_{m+1} + \ldots + \lambda_n v_n$ liegt im Kern von f. Da $v_1, \ldots, v_m$ im Kern liegen und jeder Vektor in V auf genau eine Weise als Linearkombination der $v_1, \ldots, v_n$ dargestellt werden kann, folgt $\lambda_{m+1} = \ldots = \lambda_n = 0$. $\square$

Folgerung 4.2.5. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen der selben Dimension. Dann gilt: Die Abbildung f ist genau dann surjektiv, wenn sie injektiv ist.*

Beweis. Nach Hilfssatz 3.6.5 ist f genau dann surjektiv, wenn $\mathrm{rk}(f) = \dim(W)$ gilt. Im Spezialfall $\dim(V) = \dim(W)$ gilt dies nach der Dimensionsformel in Satz 4.2.4 wiederum genau dann, wenn der Kern von f der Nullvektorraum ist. Nach Satz 4.1.7 ist dies dazu äquivalent, dass f injektiv ist. $\square$

4.3 Lineare Abbildungen und Matrizen

Nach Satz 4.2.2 ist eine lineare Abbildung $V \to W$ insbesondere eindeutig durch ihre Werte auf einer Basis von V festgelegt. Dies ermöglicht es, lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen durch Matrizen darzustellen. Dazu führen wir den Begriff der angeordneten Basis ein, d.h. wir weisen den Basisvektoren eine Reihenfolge zu. Ist $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren in einem n-dimensionalen Vektorraum V, so dass die Teilmenge $\{v_1, \ldots, v_n\}$ eine Basis ist, so sagen wir $(v_1, \ldots, v_n) \in V^n$ ist eine *(angeordnete) Basis* von V.

Definition 4.3.1. Sei V ein Vektorraum mit Basis $B = (v_1, \ldots, v_n)$. Dann hat jedes Element $v \in V$ eine eindeutige Darstellung als Linearkombination

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Die Koeffizienten $(\lambda_1, \ldots, \lambda_n)$, als n-Tupel von Elementen des Körpers K, aufgefasst, heißen *Koordinaten* des Vektors bezüglich der Basis B. Die Abbildung $\Phi_B \colon V \to K^n$, die jedem Vektor $v \in V$ seine Koordinaten bezüglich der Basis B zuordnet, nennen wir *Koordinatenabbildung* von V bezüglich B.

Die Koordinatenabbildung in der obigen Definition ist gegeben durch

$$\Phi_B \colon V \to K^n, \quad \lambda_1 v_1 + \ldots + \lambda_n v_n \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Wie man sich leicht klarmacht, ist Φ_B eine lineare Abbildung, nach Satz 4.2.1 sogar ein Vektorraumisomorphismus. Wählt man die Standardbasis $e_1, \ldots, e_n$ von K^n, so ist $\Phi_B = \mathrm{id}_{K^n}$, die identische Abbildung auf K^n.

Definition 4.3.2. Seien V und W zwei Vektorräume über dem selben Körper K. Sei $(v_1, \ldots, v_n)$ eine Basis von V und sei $(w_1, \ldots, w_m)$ eine Basis von W. Sei $f \colon V \to W$ eine lineare Abbildung. Die Matrix $A \in M_{m,n}(K)$, deren j-te Spalte jeweils das Bild $f(v_j)$ des j-ten Basisvektors v_j in Koordinaten bezüglich $(w_1, \ldots, w_m)$ enthält, heißt *die (darstellende) Matrix* von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$.

Merkregel: "Die Spalten der Matrix stellen die Bilder der Basisvektoren dar."

Beispiel 4.3.3. Sei $A \in M_{m,n}(K)$. Dann ist durch die Matrixmultiplikation $x \mapsto Ax$ für Spaltenvektoren $x \in K^n$ eine lineare Abbildung $K^n \to K^m$ definiert. Wir erhalten für das Bild eines beliebigen Spaltenvektors $x \in K^n$

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Beispiel 4.3.4. Sei V der $\mathbb{R}$-Vektorraum der reellen Polynome vom Grad höchstens zwei. Eine Basis von V ist gegeben durch die Monome

$$t \mapsto 1, \quad t \mapsto t, \quad t \mapsto t^2.$$

Wir betrachten den linearen Endomorphismus $\varphi\colon V \to V$, der gegeben ist durch die Ableitung $f \mapsto f'$. Dann ist

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

die darstellende Matrix von φ bezüglich der obigen Basis.

Der folgende Satz zeigt, dass die Matrixmultiplikation genau der Verkettung von linearen Abbildungen entspricht.

Satz 4.3.5. *Seien die endlich-dimensionalen K-Vektorräume V, W, X und die linearen Abbildungen $g\colon V \to W$ und $f\colon W \to X$ gegeben. Sei in V, W, X jeweils eine Basis gewählt und seien A und B jeweils die darstellenden Matrizen von f bzw. g. Dann ist das Matrizenprodukt $A \cdot B$ die darstellende Matrix von $f \circ g$.*

Beweis. Sei $v_1, \dots, v_n$ eine Basis von V. Dann sind die Spalten Be_j von B die Bilder dieser Basisvektoren unter g, bezüglich der in W gewählten Basis dargestellt, wobei e_j-den j-ten kanonischen Basisvektor von K^n bezeichnet. Sei $\Phi\colon X \to K^p$, wobei $p = \dim(X)$, die Koordinatenabbildung bezüglich der in X gewählten Basis. Es folgt $\Phi(f(g(v_j))) = A(Be_j) = (AB)e_j$, d.h. die j-te Spalte von AB ist das Bild von v_j unter $f \circ g$, dargestellt bezüglich der in X gewählten Basis. $\square$

Sei E die $n \times n$-Einheitsmatrix. Wegen

$$\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ist E die darstellende Matrix der Identität $\mathrm{id}_V : V \to V$ für jeden n-dimensionalen Vektorraum, wenn man in V ein- und dieselbe Basis als Basis des Definitionsbereichs und des Wertebereichs wählt.

Definition 4.3.6. Eine Matrix heißt *invertierbar*, wenn sie die darstellende Matrix einer bijektiven linearen Abbildung ist. Die darstellende Matrix ihrer Umkehrabbildung bezeichnen wir mit A^{-1}, sie wird die zu A inverse Matrix oder die *Inverse* von A genannt.

Aus der Definition folgt mit Satz 4.2.1, dass nur quadratische Matrizen invertierbar sein können. Aus Satz 4.3.5 folgt, dass für eine invertierbare $n \times n$-Matrix stets

$$AA^{-1} = A^{-1}A = E$$

gilt.

Hilfssatz 4.3.7. *Seien $A, B \in M_n(K)$ und sei E die $n \times n$-Einheitsmatrix. Dann sind die folgenden Aussagen äquivalent.*

(i) $AB = E$.

(ii) $BA = E$.

(iii) *A is invertierbar und B ist die Inverse von A.*

(iv) *B is invertierbar und A ist die Inverse von B.*

Beweis. Gelte (i), d.h. sei $AB = E$. Es folgt dass die durch $x \mapsto Bx$ gegebene lineare Abbildung $K^n \to K^n$ injektiv ist, denn sonst wäre die Verkettung

$x \mapsto ABx = x$ nicht injektiv. Nach Folgerung 4.2.5 ist $x \mapsto Bx$ dann auch surjektiv, somit ist B die darstellende Matrix einer bijektiven linearen Abbildung, also invertierbar. Nach Satz 4.2.1 bildet B somit die kanonische Basis von K^n auf eine Basis ab (nämlich die, die durch die Spalten von B gegeben ist). Da $x \mapsto Ax$ diese (wegen $AB = E$) wieder auf die kanonische Basis von K^n abbildet, ist auch A nach Satz 4.2.1 invertierbar. Da eine lineare Abbildung durch ihre Werte auf einer Basis berets eindeutig festgelegt ist, folgt außerdem, dass A mit der darstellenden Matrix der Umkehrabbildung von $x \mapsto Bx$ übereinstimmt, also die Inverse von B ist. Dies zeigt (ii), (iii) und (iv). Analog folgen (i), (iii) und (iv) aus (ii). Direkt aus der Definition folgt, dass sowohl (iii) als auch (iv) die Aussagen (i) und (ii) implizieren. □

Hilfssatz 4.3.8. *Sind $A, B \in M_n(K)$ invertierbare Matrizen. Dann ist sind auch das Produkt AB und die Inverse A^{-1} von A invertierbar und es gilt*

$$(AB)^{-1} = B^{-1}A^{-1} \quad \textit{und} \quad (A^{-1})^{-1} = A.$$

Beweis. Wegen der Assoziativität der Matrizenmultiplikation gilt

$$(AB)(B^{-1}A^{-1}) = A((BB^{-1})A^{-1}) = A(EA^{-1}) = AA^{-1} = E,$$

Mit Hilfssatz 4.3.7 folgt, dass $B^{-1}A^{-1}$ die Inverse zu AB ist. Aus $AA^{-1} = E$ folgt mit Hilfssatz 4.3.7, dass A^{-1} invertierbar ist und A seine Inverse. □

Hilfssatz 4.3.8 zeigt, dass die Menge

$$\mathrm{GL}_n(K) := \{A \in M_n(K) \mid A \text{ ist invertierbar}\}$$

eine Gruppe mit Matrixmultiplikation als Verknüpfung (diese ist assoziativ) und der $n \times n$-Einheitsmatrix als neutralem Element ist. Die Gruppe $\mathrm{GL}_n(K)$ wird als *allgemeine lineare Gruppe* vom Grad n über K bezeichnet.

Kapitel 5

Lineare Gleichungssysteme

5.1 Lineare Gleichungssysteme und Matrizen

Definition 5.1.1. Sei K ein Körper und seien $m, n \in \mathbb{N}$. Für $1 \leq i \leq m$ und $1 \leq j \leq n$ seien die Elemente $a_ij \in K$ und $b_i \in K$ gegeben. Dann nennen wir

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

ein *System von m linearen Gleichungen in den n Unbekannten* $x_1, \ldots, x_n$. Für den Fall, dass $b_1 = \ldots = b_m = 0$ gilt, nennen wir das System *homogen*, sonst *inhomogen*. Die Menge aller n-Tupel $(x_1, \ldots, x_n) \in K^n$, für die alle m Gleichungen simultan erfüllt sind, nennen wir die *Lösungsmenge* des Systems.

Sei $A \in M_{m,n}(K)$ die Matrix mit den Einträgen a_{ij} und sei b der Spaltenvektor mit den Einträgen b_i. Fassen wir die Unbekannten als einen Spaltenvektor x auf, so können wir das obige lineare Gleichungssystem äquivalent schreiben als

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

oder, in Kurzform, als

$$Ax = b.$$

Zu einem linearen Gleichungssystem $Ax = b$ nennen wir $Ax = 0$ das *zugehörige homogene System*. Die Lösungsmenge des homogenen Systems $Ax = 0$ ist nichts anderes als der Kern der linearen Abbildung

$$K^n \to K^m \colon x \mapsto Ax.$$

Dies zeigt mit Satz 4.1.7, dass die Lösungsmenge eines homogenen linearen Gleichungssystems stets ein Untervektorraum von K^n ist. Der Nullvektor in K^n ist also stets eine Lösung eines homogenen Gleichungssystems, wir nennen den Nullvektor die *triviale Lösung*. Ist $\xi \in K^n$ ein Element aus der Lösungsmenge des inhomogenen Systems $Ax = b$, gilt also $A\xi = b$, so nennen wir ξ auch eine *spezielle Lösung* von $Ax = b$.

Satz 5.1.2. *Sei ξ eine spezielle Lösung des in homogenen Gleichungssystems $Ax = b$. Dann gilt:*

(i) Für jede spezielle Lösung η des inhomogenen Systems $Ax = b$ ist $\xi - \eta$ eine Lösung des zugehörigen homogenen Systems $Ax = 0$.

(ii) Umgekehrt, für jede Lösung v des homogenen Systems $Ax = 0$, d.h. jeden Vektor $v \in K^n$ mit $Av = 0$ ist auch $\xi + v$ eine Lösung des inhomogenen Systems $Ax = b$.

Beweis. Aus $A\xi = A\eta = b$ folgt $A(\xi - \eta) = A\xi - A\eta = 0$, was (i) zeigt. Aus $A\xi = b$ und $Av = 0$ folgt $A(\xi + v) = A\xi + Av = b$, was (ii) zeigt. $\square$

Beispiele 5.1.3. Wir betrachten verschiedene Gleichungssystem für $K = \mathbb{R}$.

(i) Wir betrachten das Gleichungssystem

$$\begin{array}{rrcl} 3x_1 & +3x_2 & = & 9, \\ & 2x_2 & = & 4. \end{array}$$

Aus der zweiten Gleichung folgt sofort $x_2 = 2$. Setzen wir dies in die ersten Gleichung ein, erhalten wir $3x_1 + 6 = 9$ und somit $x_1 = 1$. Die Lösung dieses Gleichungssystems ist also eindeutig bestimmt, die einzige (spezielle) Lösung ist $(x_1, x_2) = (1, 2)$. Die Lösungsmenge des zugehörigen homogenen Systems enthält somit nur den Nullvektor.

(ii) Wir betrachten das Gleichungssystem

$$\begin{array}{rrcl} 2x_1 & +4x_2 & = & 8, \\ 3x_1 & +6x_2 & = & 12, \end{array}$$

Dieses System kann man äquivalent umformen, indem man z.B. die zweite Gleichung mit $\frac{2}{3}$ multipliziert. Dann sieht man, dass die beiden Gleichungen äquivalent zueinander sind. Die Lösungsmenge des zugehörigen homogenen Systems besteht aus allen $(x_1, x_2) \in \mathbb{R}^2$ mit $x_1 = -2x_2$, d.h. aus dem Untervektorraum von $\mathbb{R}^2$, der von $(-2, 1)$ aufgespannt wird. Eine spezielle Lösung des Systems ist gegeben durch $(x_1, x_2) = (2, 1)$. Die Lösungsmenge des inhomogenen Systems ist somit

$$\{(2 - 2t, 1 + t) \mid t \in \mathbb{R}\}.$$

(iii) Wir ändern das Gleichungssystem in (ii) ab, indem wir auf der rechten Seite der zweiten Gleichung die 12 durch eine 7 ersetzen:

$$\begin{array}{rrcl} 2x_1 & +4x_2 & = & 8, \\ 3x_1 & +6x_2 & = & 7, \end{array}$$

Multiplizieren wir die zweite Gleichung mit $\frac{2}{3}$, so erhalten wir die Gleichung $2x_1 + 4x_2 = \frac{14}{3}$, die der ersten widerspricht. Es gibt somit überhaupt keine Lösungen des inhomogenen Systems.

Wir sehen an diesen Beispielen, dass lineare Gleichungssysteme eine eindeutig bestimmte, unendlich viele oder auch gar keine Lösung haben können.

Definition 5.1.4. Sei A eine Teilmenge des Vektorraums V. Wir sagen A ist ein *affiner Unterraum* von V, falls für A eine der beiden folgenden Eigenschaften gilt.

(i) A ist die leere Menge.

(ii) Es gibt eine Vektor $p \in V$ und einen Untervektorraum $U \subseteq V$, so dass $A = p + U$, d.h. für jedes $v \in A$ gilt $v - p \in U$.

Der Punkt p ist nur eindeutig bestimmt, falls $U = \{0\}$ gilt. Ansonsten kann p durch jeden Punkt $q \in U$ ersetzt werden, da $p + U = q + U$ wegen $p - q \in U$ gilt. Der Untervektorraum U ist jedoch eindeutig bestimmt, denn aus $p + U = q + W$ folgt $p - q \in U$ und $p - q \in W$, woraus wiederum $W \subseteq U$ und $U \subseteq W$ folgt. Für einen nicht leeren affinen Unterraum $p + U$ definiert man die *Dimension* durch $\dim(A) := \dim(U)$.

Satz 5.1.2 zeigt, dass die Lösungsmenge L eines linearen Gleichungssystems stets ein affiner Unterraum von K^n ist; nämlich von der Form $\xi + \ker(A)$, wobei ξ eine spezielle Lösung des Gleichungssystems ist und wir $\ker(A)$ als den Kern der linearen Abbildung $x \mapsto Ax \colon K^n \to K^m$ definieren.

5.2 Der Gaußalgorithmus

Wir kommen nun zu einem Verfahren, mit dem lineare Gleichungssysteme praktisch gelöst werden können, dem Gaußalgorithmus. Er beruht auf der folgenden Beobachtung.

Satz 5.2.1. *Die folgenden Umformungen eines linearen Gleichungssystems sind* Äquivalenzumformungen*, d.h. sie ändern nichts an der Lösungsmenge des Systems.*

(GA1) Das Vertauschen zweier Gleichungen, allgemeiner beliebiges Umsortieren der Gleichungen.

(GA2) Das Multiplizieren einer Gleichung mit einem Skalar ungleich Null.

(GA3) Das Addieren des λ-fachen einer Gleichung zu einer anderen Gleichung, wobei $\lambda \in K$ ein Skalar ist.

Beweis. Diese Aussagen gelten offensichtlich. □

Beispiel 5.2.2. Bevor wir den Gaußalgorithmus einführen, wollen wir ihn an einem einfachen Beispiel demonstrieren. Dazu betrachten wir das folgende inhomogene Gleichungssystem im $\mathbb{R}^6$, das wir in die sogenannte *Stufenform* überführen wollen.

$$\begin{array}{rrrrrcl} & 10x_3 & -3x_4 & -x_5 & & = & 15. \\ & 5x_3 & +2x_4 & & -2x_6 & = & 7, \\ 3x_2 & & & +x_5 & & = & 12. \end{array}$$

Man beachte, dass die Variable x_1 nicht vorkommt. Im ersten Schritt wenden wir (GA1) an bringen wir die dritte Zeile nach oben

$$\begin{array}{rrrrrcl} 3x_2 & & & +x_5 & & = & 12, \\ & 5x_3 & +2x_4 & & -2x_6 & = & 7, \\ & 10x_3 & -3x_4 & -x_5 & & = & 15. \end{array}$$

Diese Umformung ist streng genommen natürlich überflüssig, wir machen sie aus eher formalen Gründen, um das System am Ende in eine möglichst übersichtliche Form zu bringen. Der zweite Umformungsschritt besteht darin, dass wir nach (GA3) das Zweifache der zweiten Zeile von der dritten abziehen.

$$\begin{array}{rrrrrcl} 3x_2 & & & +x_5 & & = & 12, \\ & 5x_3 & +2x_4 & & -2x_6 & = & 7, \\ & & -7x_4 & -x_5 & +4x_6 & = & 1. \end{array}$$

Damit ist das Gleichungssystem auf *Stufenform* gebracht. Das Gleichungssystem lässt sich jetzt durch *Rückwärtseinsetzen* lösen. Sind nämlich x_5 und x_6 vorgegeben, lässt sich mit Hilfe der dritten Gleichung x_4 bestimmen. Danach kann man aus der zweiten Gleichung x_3 bestimmen und schließlich aus der ersten x_2. Die Variable x_1 kommt in den Gleichungen nicht vor, sie kann beliebig gewählt werden. Wir verdeutlichen das wie folgt. Wir setzen zunächst

$$x_6 := r, \quad x_5 := s,$$

wobei r und s für beliebige reelle Parameter stehen. Dann ergibt sich aus der dritten Gleichung

$$x_4 = \frac{1}{7}(-s + 4r - 1) = -\frac{1}{7}s + \frac{4}{7}r - \frac{1}{7},$$

aus der zweiten

$$x_3 = -\frac{1}{5}(2x_4 - 2r - 7) = \frac{2}{35}s + \frac{6}{35}r + \frac{51}{35}.$$

und schließlich aus der ersten

$$x_2 = -\frac{1}{3}s + 4.$$

Die erste Variable $x_1 := t$ ist frei wählbar. Damit haben wir die Lösungsmenge bestimmt:

$$\left\{ \begin{pmatrix} t \\ -\frac{1}{3}s + 4 \\ \frac{2}{35}s + \frac{6}{35}r + \frac{51}{35} \\ -\frac{1}{7}s + \frac{4}{7}r - \frac{1}{7} \\ s \\ r \end{pmatrix} \in \mathbb{R}^6 \;\middle|\; r, s, t \in \mathbb{R} \right\}.$$

Eine andere Möglichkeit, die Lösungsmenge anzugeben, ist die folgende. Eine spezielle Lösung des inhomogenen Systems ist, wir setzen z.B. $r = s = t = 0$:

$$\begin{pmatrix} 0 \\ 4 \\ \frac{51}{35} \\ -\frac{1}{7} \\ 0 \\ 0 \end{pmatrix}.$$

Eine Basis des Lösungsraums des homogenen Systems ist:

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -\frac{1}{3} \\ \frac{2}{35} \\ \frac{1}{7} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \frac{6}{35} \\ \frac{4}{7} \\ 0 \\ 1 \end{pmatrix} \right).$$

Um ein lineares Gleichungssystem zu lösen, kann man die obigen Umformungen auch anhand der *erweiterten Matrix* des Gleichungssystems durchführen. In unserem Beispiel wäre das die Matrix

$$(A|b) = \left(\begin{array}{cccccc|c} 0 & 0 & 10 & -3 & 1 & 0 & 15 \\ 0 & 0 & 5 & 2 & 0 & -2 & 7 \\ 0 & 3 & 0 & 0 & 1 & 0 & 12 \end{array}\right).$$

An ihr führen wir die folgenden *Zeilenumformungen* aus, die den obigen Umformungen des Gleichungssystems entsprechen.

$$\left(\begin{array}{cccccc|c} 0 & 0 & 10 & -3 & 1 & 0 & 15 \\ 0 & 0 & 5 & 2 & 0 & -2 & 7 \\ 0 & 3 & 0 & 0 & 1 & 0 & 12 \end{array}\right) \rightsquigarrow$$

$$\left(\begin{array}{cccccc|c} 0 & 3 & 0 & 0 & 1 & 0 & 12 \\ 0 & 0 & 5 & 2 & 0 & -2 & 7 \\ 0 & 0 & 10 & -3 & 1 & 0 & 15 \end{array}\right) \rightsquigarrow$$

$$\left(\begin{array}{cccccc|c} 0 & 3 & 0 & 0 & 1 & 0 & 12 \\ 0 & 0 & 5 & 2 & 0 & -2 & 7 \\ 0 & 0 & 0 & -7 & 1 & 4 & 1 \end{array}\right)$$

In der folgenden Definition wollen wir die im obigen Beispiel durchgeführten Umformungsschritte formal genau beschrieben.

Definition 5.2.3. Sei durch $Ax = b$ mit $A \in M_{m,n}(K)$ und $b \in K^m$ ein lineares Gleichungssystem auf dem K^n gegeben. Dann nennen wir $(A|b) \in M_{m,n+1}$ die zu dem Gleichungssystem gehörende *erweiterte Matrix*. Der *Gaußalgorithmus* besteht darin, die folgenden Umformungsschritte an der Matrix $(A|b)$ durchzuführen.

(i) Wir beginnen mit der am weitesten links stehenden Spalte von A, die nicht ausschließlich Nullen enthält. Sei dies die j-te Spalte. In dieser wählen wir den am weitesten oben stehenden Eintrag, der von Null verschieden ist. Falls dieser nicht in der ersten Zeile steht, sondern in der j-ten mit $j \neq 1$, so vertauschen wir die j-te Zeile mit der ersten in $(A|b)$. Danach können wir annehmen, dass die Matrix $(A|b)$ die Form

$$(A|b) = \left(\begin{array}{ccccccc|c} 0 & \dots & 0 & a_{1j} & * & \dots & * & * \\ 0 & \dots & 0 & a_{2j} & * & \dots & * & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a_{mj} & * & \dots & * & * \end{array}\right),$$

hat, wobei das Symbol $*$ als Abkürzung für irgendwelche Elemente aus K steht, über die wir momentan nichts aussagen wollen.

(ii) Nun ziehen wir

- das $\frac{a_{2j}}{a_{1j}}$-fache der ersten Zeile von der zweiten ab,
- das $\frac{a_{3j}}{a_{1j}}$-fache der ersten Zeile von der dritten, usw.

  $\vdots$

- und schließlich das $\frac{a_{mj}}{a_{1j}}$-fache der ersten Zeile von der letzten.

Danach hat die Matrix die Form die Matrix (A, b) die Form

$$(A|b) = \left(\begin{array}{ccccccc|c} 0 & \dots & 0 & a_{1j} & * & \dots & * & * \\ 0 & \dots & 0 & 0 & * & \dots & * & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 & * & \dots & * & * \end{array}\right).$$

(iii) Wir wenden dasselbe Verfahren nun auf die unteren $m-1$ Zeilen der obigen Matrix an, solange bis wir sie in die folgende Gestalt

$$\left(\begin{array}{cccccccccc|c} 0 & \dots & 0 & \alpha_{1j_1} & * & & \dots & & & * & * \\ 0 & \dots & \dots & \dots & 0 & \alpha_{2j_2} & * & \dots & & * & * \\ \vdots & & & & & \ddots & \ddots & \ddots & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \alpha_{rj_r} & * \; \dots & * & * \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \beta_{r+1} \\ \vdots & & & & & & & & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \beta_m \end{array}\right).$$

überführt haben. Die Elemente $\alpha_{1j_1}, \alpha_{2j_2}, \dots, \alpha_{rj_r}$ nennen wir *Pivotelemente*, die Variablen $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ heißen *Pivotvariablen.*

Die Pivotelemente einer Matrix in Zeilenstufenform sind alle von Null verschieden. Bringt man ein Gleichungssystem mit dem Gaußalgorithmus auf Stufenform, so ändert sich nach Satz 5.2.1 nichts an der Lösungsmenge des Systems. Aus der Stufenform lässt sich jedoch durch Rückwärtseinsetzen leicht die Lösungsmenge bestimmen: Sind die Elemente $b_{r+1}, \ldots, b_m$ nicht alle gleich Null, so ist die Lösungsmenge leer. Andernfalls bestimmt man die Lösungsmenge wie folgt. Die Variablen x_j, die keine Pivotvariablen sind, können beliebig vorgegeben werden. Dann kann man, bei der letzten Pivotvariable x_{j_r} beginnend, nacheinander aus den ersten r Gleichungen die Werte der Pivotvariablen $x_{j_r}, \ldots, x_{j_2}, x_{j_1}$ ermitteln.

Definition 5.2.4. Die folgenden Umformungen einer Matrix $A \in M_{m,n}(K)$ werden *elementare Zeilenumformungen* genannt.

(i) Das Vertauschen von zwei Zeilen.

(ii) Das Multiplizieren einer Zeile mit einem Skalar ungleich Null.

(iii) Das Addieren des λ-fachen einer Zeile zu einer anderen.

Der Gaußalgorithmus besteht also darin, gewisse elementare Zeilenumformungen an der erweiterten Matrix $(A|b)$ durchzuführen.

Definition 5.2.5. Sei $A \in M_{m,n}(K)$ eine Matrix. Wir sagen, A hat *Zeilenstufenform*, wenn folgendes gilt: Es gibt eine aufsteigende Folge von Indizes $1 \leq j_1 < j_2 < \ldots < j_r \leq n$, so dass folgendes gilt:

(i) Die Einträge $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ sind ungleich Null.

(ii) Für $k = 1, \ldots, r$ sind jeweils alle Einträge a_{ij} mit $i \geq k$ and $j \leq j_k - 1$ gleich Null.

(iii) Die untersten $m - r$ Zeilen sind Nullzeilen.

Dann nennen wir die Einträge $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ die *Pivotelemente* von A.

Eine Matrix in Zeilenstufenform hat also die Form

$$\begin{pmatrix} 0 & \dots & 0 & \alpha_{1j_1} & * & & \dots & & * \\ 0 & \dots\dots\dots\dots\dots\dots & & & 0 & \alpha_{2j_2} & * & \dots & * \\ \vdots & & & & \ddots & \ddots & \ddots & & \vdots \\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots & & & & 0 & \alpha_{rj_r} & * \ \dots & * \\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & & & & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & & & & 0 \end{pmatrix}.$$

mit $a_{1j_1}, a_{2j_2}, \dots, a_{rj_r} \neq 0$.

Definition 5.2.6 (Rang)**.** Sei $f\colon V \to W$ eine lineare Abbildung. Dann nennen wir die Dimension des Bildes den *Rang* $\mathrm{rk}(f)$ der linearen Abbildung, in Zeichen $\mathrm{rk}(f) := \dim(\mathrm{Bild}(f))$. Sei $A \in M_{m,n}(K)$ eine Matrix. Dann bezeichnen wir mit φ_A die lineare Abbildung $\varphi_A\colon K^n \to K^m\colon x \mapsto Ax$ und wir nennen das Bild bzw. den Kern von φ_A auch das *Bild* von A bzw. den *Kern* von A. Entsprechend bezeichnen wir $\mathrm{rk}(\varphi_A)$ auch als den *Rang* $\mathrm{rk}(A)$ der Matrix A. Sei

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Dann bezeichnen wir

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

als die *Spaltenvektoren* von A und

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix}, \dots, \begin{pmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{pmatrix}$$

als die *Zeilenvektoren* von A. Als *Zeilenrang* von A bezeichnen wir die Dimension des Aufspanns der Zeilenvektoren von A in K^n, als *Spaltenrang* die Dimension des Aufspanns der Spaltenvektoren in K^m.

Hilfssatz 5.2.7. *Sei $A \in M_{m,n}(K)$ eine Matrix in Zeilenstufenform. Dann sind die Zeilenvektoren von A, die ungleich Null sind, linear unabhängig.*

Beweis. Seien $z_1, \dots, z_r$ die von Null verschiedenen Zeilenvektoren von A und gelte $\lambda_1 z_1 + \ldots + \lambda_r z_r = 0$. Seien $a_{1j_1}, a_{2j_2}, \dots, a_{rj_r}$ die Pivotelemente von A. Dann steht hat der Vektor auf der linken Seite in der j_1-ten Komponente den Eintrag $\lambda_1 a_{1j_1}$. Da $a_{1j_1} \neq 0$, folgt, dass $\lambda_1 = 0$ ist. Nun folgt, dass der Vektor auf der linken Seite in der j_2-ten Komponente den Eintrag $\lambda_2 a_{1j_2}$, woraus $\lambda_2 = 0$ folgt, usw. Dies zeigt $\lambda_1 = \ldots = \lambda_r = 0$. □

Satz 5.2.8. *Sei $A \in M_{m,n}(K)$. Dann gelten die folgenden Aussagen.*

(i) *Führt man eine elementare Zeilenumformung an einer Matrix durch, so bleiben der Kern der Matrix sowie der Rang der Matrix unverändert. Insbesondere ändert das Anwenden des Gaußalgorithmus nichts am Kern oder am Rang der Matrix.*

(ii) *Der Rang einer Matrix in Zeilenstufenform ist gleich der Anzahl der Pivotelemente bzw. -variablen und somit gleich der Anzahl der von Null verschiedenen Zeilen.*

(iii) *Der Zeilenrang und der Spaltenrang von A sind beide gleich dem Rang von A.*

Man beachte, dass das Bild einer Matrix unter elementaren Zeilenumformungen im Allgemeinen nicht unverändert bleibt (sondern nur die Dimension des Bildes).

Beweis von Satz5.2.8. (i) Wir haben bereits in Satz 5.2.1 notiert, dass elementare Zeilenumformungen Äquivalenzumformungen des homogenen Gleichungssystems $Ax = 0$ sind. Daher ändern sie nichts am Kern der Matrix. Aus der Dimensionsformel für lineare Abbildungen (Satz 4.2.4) folgt damit, dass sich bei Zeilenumformungen auch die Dimension des Bildes der Matrix unverändert bleibt. (Das Bild selbst kann sich ändern.)

(ii) Sei r die Anzahl der Pivotelemente. Wir bemerken zuerst, dass das Bild der linearen Abbildung φ_A von den Spaltenvektoren von A aufgespannt wird, denn für

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

gilt $Ax = x_1 s_1 + \ldots + x_n s_n$, wobei $s_1, \ldots, s_n$ die Spaltenvektoren von A sind. Es gilt, dass $\mathrm{rk}(A)$ kleiner oder gleich r ist, denn das Bild ist im Aufspann der ersten r kanonischen Basisvektoren $e_1, \ldots, e_r$ enthalten. Andererseits sind die Spalten $s_{j_1}, s_{j_2}, \ldots, s_{j_r}$ linear unabhängige Vektoren im Bild von A. (Dies zeigt man analog wie im Beweis von Hilfssatz 5.2.7.) Also gilt $\mathrm{rk}(A) = r$.

(iii) Wir haben in (ii) gesehen, dass der Aufspann der Spaltenvektoren gleich dem Bild von A ist. Also ist der Spaltenrang von A gleich dem Rang von A. Andererseits ändern elementare Zeilenumformungen, insbesondere das Anwenden des Gaußalgorithmus, nichts am Aufspann der Zeilenvektoren. Somit ändert sich der Zeilenrang nicht, wenn man die Matrix mit dem Gaußalgorithmus auf Zeilenstufenform bringt. Nach Hilfssatz 5.2.7 sind die von Null verschiedenen Zeilen einer Matrix in Zeilenstufenform linear unabhängig und eine Matrix in Zeilenstufenform mit r Pivotelementen hat somit Zeilenrang r, was nach (ii) gleich dem Rang von A ist. □

Bemerkung 5.2.9. In Definition 5.2.3 wurde für den Gaußalgorithmus der genaue Ablauf festgelegt. Insbesondere beim praktischen Rechnen ist es jedoch oft vorteilhaft, davon abzuweichen. Man kann z.B. zusätzliche Zeilenvertauschungen durchführen oder, was in Definition 5.2.3 nicht benutzt wird, Zeilen mit von Null verschiedenen Skalaren multiplizieren. Will man ein lineares Gleichungssystem $Ax = b$ für verschiedene rechte Seiten $b \in \{b_1, \ldots, b_k\}$, kann man die Matrix

$$(A|b_1|\ldots|b_k)$$

betrachten, um so verschiedene rechte Seiten simultan zu behandeln. Auch bei dem im nächsten Abschnitt beschriebenen Verfahren zum Invertieren von Matrizen wird eine von Definition 5.2.3 abweichende Variante des Gaußalgorithmus benutzt.

5.3 Weitere Anwendungen des Gaußalgorithmus

Wir wollen einige weitere Anwendungen des Gaußalgorithmus beschreiben. Die erste ist ein Verfahren, mit dem man eine Basis eines von m Vektoren im K^n erzeugten Untervektorraums bestimmen kann.

Satz 5.3.1. *Sei $A \in M_{m,n}(K)$ eine Matrix und sei B die Matrix, die sich aus A durch Anwenden des Gaußalgorithmus ergibt. Dann bilden die von Null verschiedenen Zeilenvektoren von B eine Basis des von den Zeilenvektoren von A aufgespannten Untervektorraums von K^n.*

Beweis. Das Anwenden des Gaußalgorithmus ändert nichts am Aufspann der Zeilenvektoren, also spannen die Zeilenvektoren von A und die von B den selben Untervektorraum U von K^n auf. Um diesen aufzuspannen, genügen natürlich die von Null verschiedenen Zeilenvektoren von B. Nach Hilfssatz 5.2.7 sind diese linear unabhängig, also eine Basis von U. $\square$

Beispiel 5.3.2. Wir verdeutlichen das Verfahren an einem Beispiel. Seien im $\mathbb{R}^4$ die Vektoren

$$\begin{pmatrix} -1 \\ -2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -3 \\ -1 \\ 0 \end{pmatrix}$$

gegeben und sei $U \subseteq \mathbb{R}^4$ der Spann dieser Vektoren. Wir tragen diese Vektoren als Zeilen in eine Matrix ein und führen den Gaußalgorithmus durch.

$$\begin{pmatrix} -1 & -2 & -1 & 1 \\ 0 & -1 & 0 & 2 \\ 2 & 2 & 1 & 2 \\ 0 & -1 & 1 & 2 \\ -2 & -3 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & -2 & -1 & 1 \\ 0 & -1 & 0 & 2 \\ 0 & -2 & -1 & 4 \\ 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -2 \end{pmatrix} \rightsquigarrow$$

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

Eine Basis von U ist also durch die Vektoren

$$\begin{pmatrix} -1 \\ -2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

gegeben.

Eine weitere Anwendung des Gaußalgorithmus ist zu entscheiden, ob eine gegebene quadratische Matrix invertierbar ist und gegebenenfalls die Berechnung der Inversen. Um zu entscheiden, ob eine quadratische Matrix invertierbar ist, genügt es, den Gaußalgorithmus auf die Matrix anzuwenden. Nach Folgerung 4.2.5 ist eine quadratische Matrix genau dann invertierbar, wenn ihr Kern trivial ist, d.h. nur aus dem Nullvektor besteht; dies ist genau dann der Fall, wenn das Anwenden des Gaußalgorithmus auf die quadratische Matrix keine Nullzeilen liefert. Wir wollen eine Variante des Gaußalgorithmus einführen, mit der man die Inverse der Matrix gegebenenfalls auch berechnen kann. Um die Korrektheit des Verfahrens beweisen zu können, benötigen wir die folgende Definition.

Definition 5.3.3. Sei n eine natürliche Zahl und K ein Körper. Wir definieren die folgenden quadratischen Matrizen aus $M_n(K)$.

(i) Für $1 \le i < j \le n$ setzen wir

$$E_{ij} := \begin{pmatrix}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & 0 & & & & 1 & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & 1 & & & & 0 & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{pmatrix}
\begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \end{matrix}$$

d.h. wir setzen E_{ij} gleich der Matrix, die aus der $n \times n$-Einheitsmatrix durch Vertauschen der i-ten und der j-ten Zeile entsteht.

(ii) Für $\lambda \in K \setminus \{0\}$ setzen wir $E_i(\lambda)$ gleich der Matrix, die aus der $n \times n$-Einheitsmatrix durch Multiplikation der i-ten Zeile mit dem Skalar λ hervorgeht, d.h.

$$E_i(\lambda) := \begin{pmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & & & & \\
& & & \lambda & & & \\
& & & & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{pmatrix}
\begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{matrix} .$$

(iii) Für $\lambda \in K$ und $i, j \in \{1, \ldots, n\}$, $i \neq j$ setzen wir $E_{ij}(\lambda)$ gleich der Matrix, die aus der $n \times n$-Einheitsmatrix durch Addition des λ-fachen der i-ten Zeile zur j-ten Zeile hervorgeht, d.h. es gilt

$$E_{ij}(\lambda) := \begin{pmatrix}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & 1 & & & & 0 & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & \lambda & & & & 1 & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{pmatrix}
\begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \end{matrix}$$

für $i < j$. Für $i > j$ erscheint der Eintrag λ über der Diagonalen und die Matrix sieht wie folgt aus.

$$E_{ij}(\lambda) := \left(\begin{array}{ccc|c|ccc|c|ccc}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
\hline
&&& 1 &&&& \lambda &&& \\
\hline
&&&& 1 &&&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
\hline
&&& 0 &&&& 1 &&& \\
\hline
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\
\end{array}$$

Die so definierten Matrizen werden als die $n \times n$-*Elementarmatrizen* bezeichnet.

Hilfssatz 5.3.4. *Elementarmatrizen sind invertierbar. Außerdem gilt für eine Matrix $A \in M_{n,m}(K)$:*

(i) Multipliziert man A mit E_{ij} von links, so bewirkt dies die Vertauschung der i-ten mit der j-ten Zeile von A, d.h. das Produkt $E_{ij}A$ entsteht aus der Matrix A, indem man die i-te und j-ten Zeile miteinander vertauscht.

(ii) Multipliziert man A mit $E_i(\lambda)$ von links, so entsteht das Produkt $E_i(\lambda)A$ aus der Matrix A, indem man die i-te Zeile von A mit λ multipliziert.

(iii) Multipliziert man A mit $E_{ij}(\lambda)$ von links, so entsteht das Produkt $E_{ij}(\lambda)A$ aus der Matrix A, indem man das λ-fache der i-ten Zeile von A zur j-ten Zeile von A addiert.

Zusammengefasst: Das Multiplizieren einer $n \times m$-Matrix von links mit den Elementarmatrizen $E_{ij}, E_i(\lambda), E_{ij}(\lambda)$ bewirkt jeweils genau die in Definition 5.2.4 (i), (ii), (iii) beschriebenen elementaren Zeilenumformungen an A.

Beweis. Das Elementarmatrizen invertierbar sind, sieht man direkt an der Form der Matrizen, diese sind nämlich entweder bereits auf Zeilenstufenform oder man kann sie durch eine elementare Zeilenumformung auf Zeilenstufenform bringen. Die Aussagen (i), (ii), (iii) überlegt man sich leicht anhand der Definition der Matrizenmultiplikation. $\square$

Wie man sich leicht überlegt, gilt für die Inversen der Elementarmatrizen

$$E_{ij}^{-1} = E_{ij}, \quad E_i(\lambda)^{-1} = E_i(\lambda^{-1}) \quad \text{und} \quad E_{ij}(\lambda)^{-1} = E_{ij}(-\lambda).$$

Satz 5.3.5. *Sei $A \in M_n(K)$ eine quadratische Matrix und sei E die $n \times n$-Einheitsmatrix. Lässt sich die $n \times 2n$-Matrix $(A|E)$ durch elementare Zeilenumformungen auf die Gestalt $(E|B)$ mit $B \in M_n(K)$ bringen, so ist A invertierbar und B die Inverse von A.*

Beweis. Elementare Zeilenumformungen an $(A|E)$ lassen sich durch Linksmultiplikation mit Elementarmatrizen ausdrücken. Seien $M_1, \dots, M_k \in M_n(K)$ Elementarmatrizen derart, dass $M_1 \cdot \ldots \cdot M_k \cdot (A|E) = (E|B)$ gilt. Dies ist äquivalent zu $M_1 \cdot \ldots \cdot M_k \cdot A = E$ und $M_1 \cdot \ldots \cdot M_k = B$, was zeigt, dass A invertierbar ist und dass B die (eindeutig bestimmte) Inverse zu A ist. $\square$

Daraus erhält man ein Verfahren, mit dem man entscheiden kann, ob eine gegebene quadratische Matrix invertierbar ist, und mit dem man gegebenenfalls die Inverse bestimmen kann. Sei $A \in M_n(K)$ eine quadratische Matrix. Anstelle von A betrachtet man die Matrix $(A|E)$ und führt solange elementare Zeilenumformungen an ihr durch, bis die linke Hälfte der Matrix auf Zeilenstufenform gebracht ist. Wenn dabei in der linken Hälfte der Matrix Nullzeilen entstehen, ist A nicht invertierbar.

Beispiele 5.3.6. Wir verdeutlichen das Verfahren an zwei Beispielen.

(i) Wir zeigen zunächst: Die Matrix

$$T := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in M_3(\mathbb{R})$$

ist nicht invertierbar. Dazu wenden wir den Gaußalgorithmus an:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Die Matrix T hat also Rang 2 und ist somit nicht invertierbar.

(ii) Nun zeigen wir, dass die folgende Matrix invertierbar ist und bestimmen ihre Inverse.

$$A := \begin{pmatrix} 2 & 0 & 1 \\ 4 & 0 & 1 \\ 6 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{R})$$

Dazu führen wir an der Matrix $(A|E)$ die folgenden Zeilenumformungen durch.

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 & 1 & 0 \\ 6 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -2 & 1 & 0 \\ 0 & 1 & -3 & -3 & 0 & 1 \end{array}\right) \rightsquigarrow$$

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -3 & 0 & 1 \\ 0 & 0 & -1 & -2 & 1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 3 & -3 & 1 \\ 0 & 0 & -1 & -2 & 1 & 0 \end{array}\right) \rightsquigarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 3 & -3 & 1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right)$$

Die Matrix A ist also invertierbar und

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

ist die Inverse.

5.4 Basiswechsel

Definition 5.4.1. Sei V ein endlich-dimensionaler K-Vektorraum mit der Basis $B = (v_1, \ldots, v_n)$. Sei $B' = (v_1', \ldots, v_n')$ eine weitere Basis von V. Dann lässt sich jeder Vektor von B' in eindeutiger Weise als Linearkombination der Vektoren $v_1, \ldots, v_n$ darstellen, d.h. es gibt eine Matrix $T = (t_{ij}) \in M_n(K)$, so dass gilt

$$v_j' = \sum_{i=1}^{n} t_{ij} v_i.$$

Die so definierte Matrix nennen wir die *Transformationsmatrix* oder *Übergangsmatrix* für den Übergang von der Basis B' zur Basis B.

Die Matrix T in der vorangehenden Definition erhält man, indem man die Vektoren $v_1', \ldots, v_n'$, bezüglich der Basis B darstellt und die so definierten Elemente von K^n als Spaltenvektoren der Matrix T verwendet. Mit anderen Worten, die Spaltenvektoren von T sind $\Phi_B(v_1'), \ldots, \Phi_B(v_n')$. Es gilt dann

$$\Phi_B(v) = T \cdot \Phi_{B'}(v) \text{ für alle } v \in V.$$

Eine andere Art, die Matrix T zu interpretieren, ist die folgende. Betrachte die identische Abbildung $\mathrm{id}_V \colon V \to V$. Dann ist die Transformationsmatrix die darstellende Matrix von id_V, wenn für V als Definitionsbereich die Basis B' und für V als Wertebereich die Basis B verwendet wird. Insbesondere folgt, dass die Matrix T invertierbar ist, da sie die darstellende Matrix einer bijektiven linearen Abbildung ist.

Satz 5.4.2. *Seien nun V und W zwei endlich-dimensionale K-Vektorräume und sei $f \colon V \to W$ eine lineare Abbildung. Sei B eine Basis von V und C eine Basis von W. Sei außerdem A die darstellende Matrix der linearen Abbildung f bezüglich der Basen B und C. Sei T die Transformationsmatrix für den Übergang von B' zu B und sei S die Transformationsmatrix für den Übergang von C' zu C. Dann ist die darstellende Matrix von f bezüglich der Basen B' und C' gegeben durch*

$$S^{-1} A T.$$

Beweis. Sei $v \in V$ und sei $x = \Phi_{B'}(v)$ die Darstellung von v in Koordinaten bezüglich der Basis B'. Dann gilt $S^{-1}ATx = S^{-1}A\,\Phi_B(v) = S^{-1}\Phi_C(f(v)) = \Phi_{C'} \circ \Phi_C^{-1}(\Phi_C(f(v))) = \Phi_{C'}(f(v))$. $\square$

Beispiel 5.4.3. Wir wollen den Basiswechsel an einem Beispiel verdeutlichen. Sei $A \in M_{3,4}(\mathbb{R})$ die Matrix

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

Wir nehmen hier an, dass A die darstellende Matrix einer linearen Abbildung f bezüglich der kanonischen Basen von $\mathbb{R}^4$ bzw. $\mathbb{R}^3$ ist. Nun führen wir eine neue angeordnete Basis

$$B' = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)$$

von $\mathbb{R}^4$ ein und die neue Basis

$$C' = \left(\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \right)$$

von $\mathbb{R}^3$. Dann sind die Transformationsmatrizen gegeben durch

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad S = \begin{pmatrix} 2 & 0 & 1 \\ 4 & 0 & 1 \\ 6 & 1 & 0 \end{pmatrix}$$

und die darstellende Matrix von f bezüglich B' und C' ist gegeben durch

$$S^{-1}\,A\,T = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 3 & -3 & 1 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 2 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 3 & -3 & 1 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 3 & 1 & 1 \\ 2 & 2 & 1 & 2 \\ 2 & 2 & 2 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \\ 5 & 5 & 2 & -3 \\ 4 & 4 & 1 & 0 \end{pmatrix},$$

wobei wir die Inverse von S schon in Beispiel 5.3.6 (ii) berechnet haben.

Definition 5.4.4. Zwei Matrizen $A, B \in M_{m,n}(K)$ heißen *äquivalent*, wenn es invertierbare Matrizen $S \in M_m(K)$ und $T \in M_n(K)$ gilt, so dass

$$B = S^{-1}AT$$

gilt.

Zwei Matrizen sind somit genau dann äquivalent, wenn sie dieselbe lineare Abbildung $K^n \to K^m$ bezüglich verschiedener Wahlen von Basen in K^n bzw. K^m darstellen.

Satz 5.4.5. *Sei $A \in M_{m,n}(K)$ eine Matrix. Dann ist A äquivalent zu einer Matrix der Form*

$$\left(\begin{array}{ccc|ccc} 1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right), \tag{5.4.1}$$

wobei die Anzahl der Einsen gleich dem Rang von A ist. Außerdem sind zwei Matrizen aus $M_{m,n}(K)$ genau dann äquivalent, wenn sie den gleichen Rang haben.

Beweis. Sei $v_1, \dots, v_k$ eine Basis des Kerns von A. Wir ergänzen sie (nach dem Basisergänzungssatz 3.6.1) zu einer Basis $v_1, \dots, v_k, v_{k+1}, \dots, v_n$ von $\mathbb{K}^n$. Setze nun $w_1 := A\,v_{k+1}, \dots, w_{n-k} := A\,v_n$. Wir zeigen, dass diese Vektoren in K^m sind linear unabhängig sind. Angenommen, es gibt $\lambda_1, \dots, \lambda_{n-k} \in K$, so dass

$$\begin{aligned} 0 &= \lambda_1 w_1 + \dots + \lambda_{n-k} w_{n-k} = \lambda_1 f(v_{k+1}) + \dots + \lambda_{n-k} f(v_n) = \\ &= f(\lambda_1 v_{k+1} + \dots + \lambda_{n-k} v_n), \end{aligned}$$

denn liegt $\lambda_1 v_{k+1} + \ldots + \lambda_{n-k} v_n$ im Kern von A, woraus wegen der eindeutigen Darstellung jedes Vektors in K^n bezüglich der Basis $v_1, \ldots, v_n$ folgt, dass $\lambda_1 = \ldots = \lambda_{n-k} = 0$ gilt. Nun ergänzen wir $w_1, \ldots, w_{n-k}$ zu einer angeordneten Basis $w_1, \ldots, w_m$ von K^m und betrachten die darstellende Matrix der linearen Abbildung $x \mapsto Ax$ bezüglich der angeordneten Basen $(v_{k+1}, \ldots, v_n, v_1 \ldots, v_k)$ und $(w_1, \ldots, w_n)$. Diese Matrix hat die in der Behauptung angegebene Form. Da sich die Dimension des Bildes bei einer Basistransformation nicht ändert, ist die Anzahl der Einsen durch den Rang von A gegeben.

Angenommen, $A, B \in M_{m,n}(K)$ sind zwei äquivalente Matrizen, d.h. es gibt invertierbare Matrizen $S \in M_m(K)$ und $T \in M_n(K)$ mit $B = S^{-1}AT$. Dann gilt $\mathrm{rk}(B) = \mathrm{rk}(S^{-1}AT) = \mathrm{rk}(A)$, denn $x \mapsto Tx$ und $y \mapsto S^{-1}y$ sind bijektive lineare Abbildungen, deshalb ist gilt einerseits $\mathrm{Bild}(A) = \mathrm{Bild}(AT)$ und andererseits, da $y \mapsto S^{-1}y$ das Bild von AT bijektiv auf das Bild von $S^{-1}AT$ abbildet, $\mathrm{rk}(S^{-1}AT) = \mathrm{rk}(AT)$. Also haben A und B den gleichen Rang.

Umgekehrt, falls $A, B \in M_{m,n}(K)$ denselben Rang r haben, dann gibt es $S, R \in M_m(K)$ und $T, U \in M_n(K)$, sodass $S^{-1}AT$ und $R^{-1}BU$ beide die Form (5.4.1) mit derselben Anzahl von Einsen haben, also gleich sind. Dann folgt

$$B = RS^{-1}ATU^{-1} = (SR^{-1})^{-1}A(TU^{-1})$$

mit den invertierbaren Matrizen $SR^{-1} \in M_m(K)$ und $TU^{-1} \in M_n(K)$, d.h. A und B sind äquivalent. $\square$

Bemerkung 5.4.6. Ist $f\colon V \to V$ ein linearer Endomorphismus des endlich-dimensionalen Vektorraums V, sind B und B' zwei Basen von V und T die Übergangsmatrix von B nach B' und A die darstellende Matrix von f bezüglich B, so ist die darstellende Matrix von f bezüglich B' nach Satz 5.4.2 gegeben durch

$$T^{-1}\,AT.$$

Man nennt zwei quadratische Matrizen $A, B \in M_n(K)$ *ähnlich*, wenn es eine invertierbare Matrix $T \in M_n(K)$ gibt, so dass $B = T^{-1}AT$ gilt. Im Gegensatz zur Äquivalenz von Matrizen ist die Frage, wann zwei quadratische Matrizen ähnlich sind, wesentlich schwerer zu entscheiden. Wir werden uns mit dieser Frage in einem späteren Kapitel befassen.

Bemerkung 5.4.7. Die Äquivalenz und die Ähnlichkeit von Matrizen sind Äquivalenzrelationen auf der Menge der Matrizen $M_{m,n}(K)$ bzw. der Menge der quadratischen Matrizen $M_n(K)$, siehe Definition 1.4.1.

Wir beweisen nur die erste Behauptung. Seien E_m die $m \times m$-Einheitsmatrix und E_n die $n \times n$-Einheitsmatrix. Wegen $A = E_m A E_n$ ist jede Matrix zu sich selbst äquivalent, also ist die Relation reflexiv. Sei nun A äquivalent zu B und B äquivalent zu C. Dann gibt es invertierbare Matrizen $S, R \in M_m(K)$ und $T, U \in M_n(K)$, so dass $B = S^{-1}AT$ und $C = R^{-1}BU$ gilt. Dann folgt auch $A = SBT^{-1}$. Dies zeigt die Symmetrie der Relation. Außerdem folgt

$$C = R^{-1}S^{-1}ATU,$$

was die Transitivität zeigt. Ganz analog kann man den Beweis für die zweite Behauptung führen.

Kapitel 6

Mehr über Vektorräume und lineare Abbildungen

6.1 Der Vektorraum $\mathrm{Hom}(V, W)$

Seien A und B zwei Menge. Dann bezeichnen wir die Menge aller Abbildungen von A nach B mit

$$\mathrm{Abb}(A, B) := \{F\colon A \to B\}$$

Satz 6.1.1. *Sei M eine Menge und W ein K-Vektorraum. Dann ist die Menge*

$$\mathrm{Abb}(M, W) := \{F\colon M \to W\}$$

aller Abbildungen von M nach W wieder ein K-Vektorraum, wenn die Addition durch

$$(f + g)(m) := f(m) + g(m) \text{ \textit{für alle} } f, g \in \mathrm{Abb}(M, W), m \in M$$

und die Skalarmultiplikation durch

$$(\lambda \cdot f)(m) := \lambda \cdot f(m) \text{ \textit{für alle} } \lambda \in K, f \in \mathrm{Abb}(M, W), m \in M$$

definiert sind.

Beweis. Wir überlegen uns zunächst, dass $\mathrm{Abb}(M, W)$ mit der oben definierten Addition eine abelsche Gruppe ist. Die Addition ist assoziativ, da die Addition im Körper K assoziativ ist:

$$\begin{aligned}((f+g)+h)(m) &= (f+g)(m)+h(m) = f(m)+g(m)+h(m) = \\ &= f(m)+(g+h)(m) = (f+(g+h))(m)\end{aligned}$$

Das neutrale Element ist durch die Abbildung gegeben, die alle Elemente von M auf die Null in W abbildet. Für jede Abbildung $f \in \mathrm{Abb}(M, W)$ ist die Abbildung $x \mapsto -f(x)$ ein inverses Element. Die Kommutativität der Addition folgt aus der Kommutativität der Addition in W:

$$(f+g)(m) = f(m)+g(m) = g(m)+f(m) = (g+f)(m).$$

Es bleibt, die vier Axiome für die Skalarmultiplikation in Definition 3.1.1 nachzuprüfen. Seien $\lambda, \mu \in K$, $f, g \in \mathrm{Abb}(M, W)$, $m \in M$. Wir rechnen einzeln nach, dass die Eigenschaften (V1) – (V4) erfüllt sind:

$$\begin{aligned}((\lambda\cdot\mu)\cdot f)(m) &= (\lambda\cdot\mu)\cdot f(m) = \lambda\cdot(\mu\cdot f(m)) = (\lambda\cdot(\mu\cdot f))(m),\\ ((\lambda+\mu)\cdot f)(m) &= (\lambda+\mu)\cdot f(m) = \lambda\cdot f(m)+\mu\cdot f(m) = \\ &= (\lambda\cdot f+\mu\cdot f)(m),\\ (\lambda\cdot(f+g))(m) &= \lambda f(m)+\lambda g(m) = (\lambda\cdot f+\lambda\cdot g)(m),\\ (1\cdot f)(m) &= 1\cdot f(m) = f(m).\end{aligned}$$

□

Definition 6.1.2. Seien V und W zwei K-Vektorräume. Dann bezeichnen wir mit $\mathrm{Hom}(V, W)$ die Menge aller linearen Abbildungen von V nach W. Im Fall $V = W$ bezeichnen wir mit $\mathrm{End}(V) := \mathrm{Hom}(V, V)$ die Menge aller linearen Abbildungen von V nach V. Mit $\mathrm{Aut}(V)$ bezeichnen wir die Menge der Vektorraumautomorphismen von V.

Satz 6.1.3. *Seien V und W zwei K-Vektorräume. Dann ist die Teilmenge*

$$\mathrm{Hom}(V, W) \subseteq \mathrm{Abb}(V, W)$$

ein Untervektorraum. Insbesondere ist $\mathrm{Hom}(V, W)$ *ein K-Vektorraum.*

Beweis. Die Menge $\mathrm{Hom}(V, W)$ ist nicht leer, denn sie enthält in jedem Fall die Nullabbildung. Es bleibt zu zeigen, dass $\mathrm{Hom}(V, W)$ abgeschlossen unter Addition und Skalarmultiplikation ist. Seien $f, g \colon V \to W$ lineare Abbildungen. Dann ist auch $f+g$ linear, denn es gilt für $v, w \in V$ und $\lambda \in K$

$$\begin{aligned}(f+g)(v+w) &= f(v+w)+g(v+w) = f(v)+f(w)+g(v)+g(w) = \\ &= (f+g)(v)+(f+g)(w)\end{aligned}$$

und

$$\begin{aligned}(f+g)(\lambda\cdot v) &= f(\lambda\cdot v)+g(\lambda\cdot v) = \lambda f(v)+\lambda g(v) = \\ &= \lambda(f(v)+g(v)) = \lambda\cdot(f+g)(v)\end{aligned}$$

Damit haben wir die beide Kriterien (UV1) und (UV2) aus der Definition 3.2.1 nachgewiesen. □

Sind V und W zwei endlich-dimensionale K-Vektorräume, so lässt sich jede lineare Abbildung von V nach W, also jedes Element von $\mathrm{Hom}(V, W)$, nach der Wahl von Basen in eindeutiger Weise als eine Matrix in $M_{m,n}(K)$ mit $n = \dim(V)$ und $m = \dim(W)$ darstellen. Wir wollen zeigen, dass durch diese Zuordnung sogar ein Vektorraumisomorphismus gegeben ist. Man überlegt sich leicht, dass $M_{m,n}(K)$, mit der komponentenweisen Addition und Skalarmultiplikation, ein mn-dimensionaler K-Vektorraum ist. Eine Basis von $M_{m,n}(K)$ ist etwa durch die Matrizen M_{ij} gegeben, wobei wir mit M_{ij} die $m \times n$-Matrix bezeichnen, die eine Eins als Eintrag in der i-ten Zeile und j-ten Spalte hat und sonst nur Nullen enthält.

Satz 6.1.4. *Seien V und W zwei endlich-dimensionale K-Vektorräume und seien $B = (b_1, \dots, b_n)$ sowie $C = (c_1, \dots, c_m)$ Basen von V bzw. W. Die Abbildung $\mathrm{Hom}(V, W) \to M_{m,n}(K)$, die eine lineare Abbildung von V nach W auf ihre darstellende Matrix bezüglich B und C abbildet, ist ein Isomorphismus von Vektorräumen.*

Beweis. Sei $\Phi_C^B\colon \mathrm{Hom}(V, W) \to M_{m,n}(K)$, die Abbildung, die einen linearen Homomorphismus von V nach W auf seine darstellende Matrix bezüglich B und C abbildet. Wir zeigen zuerst die Linearität von Φ_C^B. Seien $f, g \in \mathrm{Hom}(V, W), \lambda, \mu \in K, v \in V, x = \Phi_B(v) \in \mathbb{K}^n$ und seien $A = \Phi_C^B(f), D = \Phi_C^B(g)$. Dann gilt

$$\begin{aligned}\Phi_C^B(\lambda f + \mu g)(x) &= \Phi_C(\lambda f(v) + \mu g(v)) = \lambda \cdot \Phi_C(f(v)) + \mu \cdot \Phi_C(g(v)) = \\ &= \lambda A x + \mu D x = (\lambda A + \mu D)x = \\ &= (\lambda \Phi_C^B(f) + \mu \Phi_C^B(g))(v)\end{aligned}$$

Da eine lineare Abbildung $f\colon V \to W$ nach Satz 4.2.2 durch ihre Werte auf der Basis B bereits eindeutig bestimmt ist und die Spalten der darstellenden Matrix die Werte der Basisvektoren $(b_1, \dots, b_n)$ festlegen, ist durch jede Matrix in $A \in M_{m,n}(K)$ eine Abbildung $V \to W$ eindeutig festgelegt, die A als darstellende Matrix bezüglich B und C hat. Andererseits folgt daraus auch, dass zwei lineare Abbildungen genau dann gleich sind, wenn sie dieselbe darstellende Matrix besitzen. Dies zeigt die Surjektivität und die Injektivität von Φ_C^B. □

Der Vektorraum $\mathrm{End}(V)$, den wir im Spezialfall $V = W$ aus $\mathrm{Hom}(V, W)$ erhalten, hat noch eine weitere Struktur. Wir können zwei lineare Endomorphismen von einem Vektorraum V mit Hilfe der Verkettung von Abbildungen miteinander verknüpfen. Der folgende Satz besagt, dass wir damit einen Ring erhalten, den *Endomorphismenring* von V.

Satz 6.1.5. *Sei V ein K-Vektorraum. Dann bildet $\mathrm{End}(V)$, mit der üblichen Addition von linearen Abbildungen und der Verkettung von Abbildungen als Multiplikation, einen unitären Ring.*

Beweis. Die Verkettung zweier linearer Abbildungen ist wieder eine lineare Abbildung siehe Beispiel 4.1.3 (iv). Wir müssen nachweisen, dass die Axiome (R1) bis (R7) in Definition 2.2.1 gelten. Wir haben bereits gezeigt, dass $\mathrm{End}(V)$ einen Vektorraum bildet, daher ist klar, dass die übliche Addition von linearen Abbildung aus $\mathrm{End}(V)$ eine abelsche Gruppe macht; dies zeigt (R1) bis (R4). Wir haben uns schon in Abschnitt 1.5 überlegt, dass die Verkettung von Abbildungen assoziativ ist, was (R5) zeigt. Die Distributivgesetze (R6) und (R7) können wir direkt nachrechnen. Seien dazu $f, g, h \in \mathrm{End}(V)$ und $v \in V$. Dann gilt

$$\begin{aligned}(f \circ (g+h))(v) &= f((g+h)(v)) = f(g(v)+h(v)) = \\ &= f(g(v)) + f(h(v)) = f \circ g(v) + f \circ h(v) = \\ &= (f \circ g + f \circ h)(v)\end{aligned}$$

und

$$\begin{aligned}((f+g) \circ h)(v) &= (f+g)(h(v)) = f(h(v)) + g(h(v))) = \\ &= f \circ h(v) + g \circ h(v) = (f \circ h + g \circ h)(v).\end{aligned}$$

Schließlich ist id_V offensichtlich ein neutrales Element für die Multiplikation in $\mathrm{End}(V)$. □

Folgerung 6.1.6. *Sei V ein endlich-dimensionaler K-Vektorraum und sei $B = (b_1, \ldots, b_n)$ eine Basis von V. Die Abbildung $\Phi_B^B \colon \mathrm{End}(V) \to M_n(K)$, die eine lineare Abbildung von V nach V auf ihre darstellende Matrix bezüglich B abbildet, ist ein Isomorphismus von Vektorräumen und außerdem ein Isomorphismus von Ringen.*

Beweis. Sei $\Phi_B^B \colon \mathrm{End}(V) \to M_n(K)$ die Abbildung, die einen linearen Endomorphismus von V auf seine darstellende Matrix bezüglich B abbildet. Dass Φ_B^B ein Isomorphismus von Vektorräumen ist, folgt aus Satz 6.1.4 im Spezialfall $V = W$. Dies beinhaltet insbesondere, dass

$$\Phi_B^B(f+g) = \Phi_B^B(f) + \Phi_B^B(g) \quad \text{für alle} \quad f, g \in \mathrm{End}(V).$$

gilt. Dass auch

$$\Phi_B^B(f \circ g) = \Phi_B^B(f) \cdot \Phi_B^B(g) \quad \text{für alle} \quad f, g \in \mathrm{End}(V).$$

gilt, folgt aus Satz 4.3.5 im Spezialfall $V = W = X$. □

Satz 6.1.7. *Sei V ein Vektorraum. Dann bildet die Menge $\mathrm{Aut}(V)$ der Vektorraumautomorphismen von V eine Gruppe mit der Verkettung von Abbildungen als Multiplikation.*

Beweis. Im Beweis von Satz 6.1.5 haben wir bereits gezeigt, dass die Verkettung zweier linearer Abbildungen linear ist. Da die Verkettung zweier bijektiver Abbildungen bijektiv ist, folgt, dass die Verkettung zweier Vektorraumautomorphismen von V wieder ein Vektorraumautomorphismus von V ist. Somit ist durch die Verkettung von Abbildungen eine Verknüpfung auf $\mathrm{Aut}(V)$ gegeben. Die Verkettung von Abbildungen ist assoziativ, wie schon in Abschnitt 1.5 gezeigt wurde. Ein neutrales Element ist durch die identische Abbildung $\mathrm{id}_V \colon V \to V$ gegeben. Schließlich ist nach Hilfssatz 4.1.5 auch die Umkehrabbildung eines Vektorraumautomorphismus wieder ein Vektorraumautomorphismus. □

Folgerung 6.1.8. *Sei V ein endlich-dimensionaler K-Vektorraum und sei $B = (b_1, \dots, b_n)$ eine Basis von V. Die Abbildung*

$$\Phi_B^B|_{\mathrm{Aut}(V)} \colon \mathrm{Aut}(V) \to \mathrm{GL}_n(K),$$

die einen Vektorraumautomorphismus von V auf seine darstellende Matrix bezüglich B abbildet, ist ein Gruppenisomorphismus.

Beweis. Die Abbildung $\Phi_B^B \colon \mathrm{End}(V) \to M_n(K)$ ist bijektiv, was wir uns im Beweis von Satz 6.1.4 bereits überlegt haben. Aus der Definition der Invertierbarkeit von Matrizen folgt, dass Φ_B^B die Menge $\mathrm{Aut}(V)$ bijektiv auf $\mathrm{GL}_n(K)$ abbildet. Satz 4.3.5 zeigt, dass es sich um einen Gruppenhomorphismus handelt. □

6.2 Linearformen und Dualraum

Für jeden Körper K ist auch $K = K^1$ selbst ein K-Vektorraum. In diesem Abschnitt werden lineare Abbildungen von einem K-Vektorraum V nach K betrachten.

Definition 6.2.1. Sei V ein K-Vektorraum. Eine *Linearform* auf V ist eine lineare Abbildung $V \to K$.

Beispiele 6.2.2. (i) Sei $V = K^n$ und seien $\alpha_1, \ldots, \alpha_n \in K$. Dann ist durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_x \end{pmatrix} \mapsto \alpha_1 x_1 + \ldots \alpha_n x_n$$

eine Linearform gegeben.

(ii) Sei V der Vektorraum der Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ und $t_0 \in \mathbb{R}$ eine reelle Konstante. Dann ist durch die *Auswertungsabbildung*

$$f \mapsto f(t_0)$$

eine Linearform auf V gegeben.

(iii) Sei $[a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall und V nun der Vektorraum der stetigen Funktionen $f\colon [a, b] \to \mathbb{R}$. Dann ist durch das Integral

$$f \mapsto \int_a^b f(t)dt$$

eine Linearform auf V gegeben. Die Linearität folgt aus den Eigenschaften

$$\int_a^b f(t) + g(t)dt = \int_a^b f(t)dt + \int_a^b g(t)dt$$

und

$$\int_a^b c \cdot f(t)dt = c \cdot \int_a^b f(t)dt \text{ für } c \in \mathbb{R}$$

des Integrals.

Definition 6.2.3. Sei V ein Vektorraum. Die Menge aller Linearformen auf V heißt der *Dualraum von* V und wird mit V^* bezeichnet.

Da

$$V^* = \operatorname{Hom}(V, K)$$

gilt, ist klar, dass der Dualraum eines K-Vektorraums V selbst wieder ein K-Vektorraum ist.

Satz 6.2.4. *Sei V ein endlich-dimensionaler Vektorraum. Dann gilt $V \cong V^*$, d.h. der Vektorraum V ist zu seinem Dualraum V^* isomorph.*

Da endlich-dimensionale Vektorräume bereits dann isomorph sind, wenn ihre Dimensionen übereinstimmen, genügt es zum Beweis von Satz 6.2.4 zu zeigen, dass zu einer Basis von V eine Basis von V^* mit gleich vielen Elementen existiert.

Hilfssatz 6.2.5. *Sei $(v_1, \ldots, v_n)$ eine Basis des Vektorraums V. Dann ist durch*

$$\omega_i(v_j); = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

für $i, j \in \{1, \ldots, n\}$ eine Basis von V^* definiert.

Beweis. Zunächst bemerken wir, dass durch die obige Festlegung tatsächlich n Linearformen $\omega_1, \ldots, \omega_n \in V^*$ gegeben sind, da lineare Abbildungen durch ihre Werte auf einer Basis bereits eindeutig bestimmt sind. Sei nun $\omega \in V$ eine Linearform auf V und seien $\alpha_j := \omega(v_j)$. Dann gilt

$$\omega = \alpha_1\omega_1 + \ldots + \alpha_n\omega_n,$$

denn $(\alpha_1\omega_1 + \ldots + \alpha_n\omega_n)(v_j) = \alpha_j$, d.h. die Linearformen ω und $\alpha_1\omega_1 + \ldots + \alpha_n\omega_n$ liefern auf den Vektoren der Basis $(v_1, \ldots, v_n)$ dieselben Werte und stimmen somit überein. Dies zeigt, dass sich jede Linearform in V^* als Linearkombination der ω_i schreiben lässt. Umgekehrt, gilt

$$\alpha_1\omega_1 + \ldots + \alpha_n\omega_n = 0,$$

so folgt durch Einsetzen von v_1 bis v_n in die Linearform auf der linken Seite, dass alle α_j gleich Null sind, was die lineare Unabhängigkeit zeigt. □

Die Basis $(\omega_1, \ldots, \omega_n)$ von V^* wird die zu $(v_1, \ldots, v_n)$ *duale Basis* genannt.

Beweis von Satz 6.2.4. Aus Hilfssatz 6.2.5 folgt, dass es zu jeder Basis von V eine Basis von V^* mit der gleichen Anzahl von Elementen gibt. □

Bemerkung 6.2.6. Ist V unendlich-dimensional und v_i mit $i \in I$ eine Basis von V, so lassen sich zwar die Elemente ω_i analog wie oben definieren, aber sie bilden kein Erzeugendensystem von V^*. Denn die Linearform ω auf V, die durch

$$\omega(v_i) = 1 \quad \text{für alle} \quad i \in I$$

definiert ist lässt sich nicht als endliche Linearkombination der ω_i darstellen.

Definition 6.2.7. Sei V ein K-Vektorraum und V^* sein Dualraum. Dann heißt die Abbildung $V \times V^* \to K, (v, \omega) \mapsto \omega(v)$ die *duale Paarung*.

Die duale Paarung ist eine bilineare Abbildung, d.h. es gilt, dass für jedes fest gewählte $\omega \in V^*$ ist die Abbildung $V \to K, v \mapsto \omega(v)$ linear ist (das ist klar) und dass für jeden fest gewählten Vektor $v \in V$ die Abbildung $V^* \to K, \omega \mapsto \omega(v)$ linear ist (dies folgt sofort aus der Definition der Addition und skalaren Multiplikation in V^*).

Bemerkung 6.2.8. Für den Fall $V = K^n$ lässt sich die duale Paarung mit Hilfe der Matrixmultiplikation beschreiben. Dazu betrachten wir die *kanonische Basis* von K^n, den wir als *Spaltenvektoren*, darstellen wollen, d.h. wir setzen $K^n = M_{n,1}(K)$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dann können wir $(K^n)^* = \mathrm{Hom}(K^n, K)$ als *Zeilenvektoren* darstellen, d.h. wir setzen $(K^n)^* = M_{1,n}(K)$. Denn die Menge der linearen Abbildungen von K^n nach K können wir mit den $1 \times n$-Matrizen mit Einträgen in K identifizieren. Dann ist die zu $e_1, \ldots, e_n$ duale Basis gegeben durch

$$e_1^* = (1, 0, \ldots, 0),\ e_2^* = (0, 1, 0, \ldots, 0),\ \ldots,\ e_n^* = (0, \ldots, 0, 1).$$

Die duale Paarung ist dann durch Matrixmultiplikation gegeben, für $v \in V = M_{n,1}(K), \omega = (\omega_1, \ldots, \omega_n) \in V^* = M_{n,1}(K)$ gilt

$$\omega(v) = (\omega_1, \ldots, \omega_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \omega_1 v_1 + \ldots + \omega_n v_n.$$

Seien nun B und B' Basen des endlich-dimensionalen Vektorraums V und sei T die Transformationsmatrix für den Übergang von B' nach B, d.h. es gilt $\Phi_B(v) = T\Phi_{B'}(v)$ für alle $v \in V$. Sei der Zeilenvektor $A \in M_{1,n}(K)$ die darstellende Matrix der Linearform $\omega \in V^*$ bezüglich B. Dann ist die darstellende Matrix von ω bezüglich B' durch die Matrix AT gegeben. Das sieht man direkt, denn $A\Phi_B(V) = AT\Phi_{B'}(v)$, es folgt aber auch aus Satz 5.4.2.

Definition 6.2.9 (Transponierte Matrix)**.** Sei $A \in M_{m,n}(K)$ eine Matrix mit den Einträgen a_{ij}, wobei für die Indizes $i \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, n\}$ gilt. Dann ist die *Transponierte* von A, bezeichnet mit $A^t \in M_{n,m}(K)$, definiert als die Matrix mit den Einträgen a_{ji}.

Die Matrix A^t entsteht also aus A durch Vertauschen von Zeilen- und Spaltenindex der Einträge.

Beispiel 6.2.10. Sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Dann ist

$$A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Lemma und Definition 6.2.11. Seien V und W Vektorräume, seien V^* und W^* ihre Dualräume. Sei $f\colon V \to W$ eine lineare Abbildung. Dann ist durch

$$f^*\colon W^* \to V^*,\ \omega \mapsto \omega \circ f$$

eine lineare Abbildung gegeben. Man nennt f^* die zu f *duale Abbildung*, die Linearform $f * \omega$ wird auch der *Pullback* von ω unter f genannt.

Beweis. Es ist klar, dass $\omega \circ f$, als Verkettung von linearen Abbildungen, wieder linear ist, also $\omega \circ f \in V^*$ gilt.

Wir zeigen noch, dass f^* linear ist. Seien $\alpha, \beta \in W^*$ und $v \in V$. Dann gilt

$$f^*(\alpha+\beta)(v) = (\alpha+\beta) \circ f(v) = \alpha(f(v)) + \beta(f(v)) = (f^*\alpha + f^*\beta)(v).$$

Sei nun $\lambda \in K$

$$f^*(\lambda \cdot \alpha)(v) = (\lambda \cdot \alpha) \circ f(v) = \lambda \cdot \alpha(f(v)) = (\lambda \cdot f^*\alpha)(v).$$

$\square$

Seien nun V und W endlich-dimensional und sei $A \in M_{m,n}(K)$ die darstellende Matrix von f bezüglich der Basen B und C von V bzw. W. Seien B^* und C^* die zu B bzw. C dualen Basen. Dann ist die darstellende Matrix von f^* bezüglich B^* und C^* durch die Transponierte A^t von A gegeben.

Sei durch $V^{**} := (V^*)^*$ der *Bidualraum* von V definiert. Dann ist durch

$$i\colon V \to V^{**},\ v \mapsto (\alpha \mapsto \alpha(v))$$

eine lineare Abbildung definiert, denn

$$\begin{aligned} i(\lambda v + \mu w) &= (\alpha \mapsto \alpha(\lambda v + \mu w)) = \\ &= (\alpha \mapsto \lambda\alpha(v) + \mu\alpha(w)) = \lambda\, i(v) + \mu\, i(w). \end{aligned}$$

Diese Abbildung ist injektiv, d.h. es gilt $\ker(i) = \{0\}$, da es zu jedem $v \in V$ eine Linearform $\alpha \in V^*$ gibt, die auf v nicht verschwindet, wie man sich leicht überlegt. Falls V endlich-dimensional ist, ist i nach der Dimensionsformel damit sogar ein linearer Isomorphismus. Man sagt, V^{**} ist zu V *kanonisch isomorph*, falls $\dim(V) < \infty$. Dabei bedeutet das Wort "kanonisch", dass man den durch i gegebenen Isomorphismus angeben kann, ohne eine Wahl (z.B. einer Basis) zu treffen.

6.3 Quotientenraum und Homomorphiesatz

Definition 6.3.1. Sei V ein K-Vektorraum und sei $U \subseteq V$ ein Untervektorraum. Dann bezeichnen wir die Menge

$$V/U := \{v + U \mid v \in V\}$$

der affinen Unterräume der Form $v+U$ mit $v \in V$ als den *Quotienten(vektor)raum* V/U (gesprochen "V modulo U"). Die Elemente von $v + U$, also die Vektoren der Form $v + u$ mit $u \in U$ bezeichnen wir als *Vertreter* oder *Repräsentanten* des Elements $v + U \in V/U$. Die Abbildung

$$\pi \colon V \to V/U, \quad v \mapsto v + U$$

bezeichnen wir als *kanonische Projektion*.

Der Raum V wird also in eine disjunkte Vereinigung von parallelen affinen Unterräumen zerlegt. Wir wollen nun V/U mit einer Vektorraumstruktur versehen. Dazu bemerken wir zunächst, dass zwei Elemente $v + U, v' + U \in V/U$ genau dann gleich sind, wenn $v - v' \in U$ gilt.

Wir definieren eine Addition auf V/U durch

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

und eine skalare Multiplikation durch

$$\lambda(v + U) := (\lambda v) + U$$

für $\lambda \in K$. Diese Definitionen hängen nicht von der Wahl des Repräsentanten ab, denn falls $v-v' \in U$ und $w-w' \in U$ gilt, so folgt, dass auch $(v+w)-(v'+w') \in U$ und $\lambda v - \lambda v' \in U$ gilt, und damit

$$(v + U) + (w + U) = (v' + U) + (w' + U)$$

sowie

$$\lambda(v + U) = \lambda(v' + U).$$

Satz 6.3.2. *Mit der oben eingeführten Addition und Skalarmultiplikation ist V/U ein K-Vektorraum und $\pi\colon V \to V/U$ ist ein linearer Epimorphismus, für den* $\ker(\pi) = U$ *gilt.*

Beweis. Bevor wir zeigen, dass V/U ein Vektorraum ist, bemerken wir, dass π die Eigenschaften einer linearen Abbildung besitzt. Nämlich gilt

$$\pi(v + w) = \pi(v) + \pi(w) \quad \text{und} \quad \pi(\lambda v) = \lambda\, \pi(v),$$

was direkt aus der Definition der Vektorraumoperationen auf V/U folgt. Mit Hilfe von π lässt sich nun die Gültigkeit der Vektorraumaxiome für V/U aus der Gültigkeit der Vektorraumaxiome für V herleiten, wir zeigen dies nur am Beispiel der Assoziativität der Addition:

$$\begin{aligned} \pi(v) + (\pi(w) + \pi(x)) &= \pi(v) + \pi(w + x) = \pi(v + (w + x)) = \\ &= \pi((v + w) + x) = \pi(v + w) + \pi(x) = (\pi(v) + \pi(w)) + \pi(x). \end{aligned}$$

Da wir nun V/U als K-Vektorraum erkannt haben und π die Eigenschaften einer linearen Abbildung hat, folgt nun, dass π tatsächlich eine lineare Abbildung zwischen den Vektorräumen V und V/U ist. Offensichtlich ist π surjektiv und es gilt $\pi(v) = 0$ genau dann, wenn $v \in U$. $\square$

Satz 6.3.3. *Sei $f\colon V \to W$ eine lineare Abbildung und $U \subseteq V$ ein Untervektorraum und sei $\pi\colon V \to V/U$ die kanonische Projektion. Genau dann gibt es eine lineare Abbildung $\tilde{f}\colon V/U \to W$ mit $\tilde{f} \circ \pi = f$, wenn $U \subseteq \ker(f)$ gilt. In diesem Fall ist $\tilde{f}$ eindeutig bestimmt.*

Bemerkung 6.3.4. Dass die Verkettung $\tilde{f} \circ \pi$ die Abbildung f ergibt, lässt sich dadurch ausdrücken, dass das folgende Diagramm kommutiert, d.h. wenn man dieselbe Abbildung von V nach W, egal, welchen der beiden Wege durch das Diagramm man wählt.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/U \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

Beweis von Satz 6.3.3. Angenommen, es gibt eine lineare Abbildung $\tilde{f}$ mit den angegebenen Eigenschaften und es gilt $v \in U$. Dann gilt $f \circ \pi(u) = f(0) = 0$, also liegt v im Kern von f.

Umgekehrt, gilt $U \subseteq \ker(f)$, dann setzen wir $\tilde{f}(v + U) := f(v)$. Die Abbildung ist wohldefiniert, denn falls $u \in U$, folgt $f(v + u) = f(v) + f(u) = f(v)$. Das so definierte $\tilde{f}$ ist linear, denn

$$\begin{aligned}\tilde{f}(\lambda(v + U) + \mu(w + U)) &= \tilde{f}((\lambda v + U) + (\mu w + U)) = \\ &= \tilde{f}((\lambda v + \mu w) + U)) = f(\lambda v + \mu w) = \lambda f(v) + \mu f(w) = \\ &= \lambda \tilde{f}(v + U) + \mu \tilde{f}(w + U).\end{aligned}$$

Dies zeigt die Existenz einer Abbildung $\tilde{f}$ mit den gewünschten Eigenschaften. Andererseits kann es nur eine solche Abbildung geben, denn für $v + U \in V/U$ muss $\tilde{f}(v + U) = f(v)$ gelten. □

Satz 6.3.5 (Homomorphiesatz)**.** *Sei $f \colon V \to W$ eine lineare Abbildung dann ist durch $\tilde{f} \colon V/\ker(f) \to \mathrm{Bild}(f)$, wobei $\tilde{f}$ wie in Satz 6.3.3 definiert ist, ein linearer Isomorphismus gegeben.*

Bemerkung 6.3.6. Die Aussage des Satzes lässt sich wie folgt umschreiben: Es gibt einen linearen Isomorphismus $V/\ker(f) \to \mathrm{Bild}(W)$, der das folgende Diagramm kommutativ macht, wobei $i \colon \mathrm{Bild}(f) \to W, w \mapsto w$ die Inklusionsabbildung bezeichnet.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/\ker(f) \arrow[r, "\tilde{f}"] & \mathrm{Bild}(f) \arrow[u, "i"'] \end{tikzcd}$$

Beweis von Satz 6.3.5. Sei $\tilde{f}$ die wie in Satz 6.3.3 definierte Abbildung, wobei wir $U = \ker(f)$ setzen. Da $\mathrm{Bild}(f) = \mathrm{Bild}(\tilde{f})$ gilt, definiert dies auch eine surjektive lineare Abbildung $V/\ker(f) \to \mathrm{Bild}(f)$. Sei $v \in V$, so dass $\tilde{f}(v + \ker(f)) =$

$f(v) = 0$, dann folgt $v \in \ker(f)$, also $v + \ker(f) = 0 \in V/\ker(f)$. Also ist $\tilde{f}$ auch injektiv. $\square$

Kapitel 7

Determinanten

7.1 Alternierende multilineare Abbildungen

Definition 7.1.1. Sei V ein K-Vektorraum. Eine Abbildung

$$\alpha\colon \underbrace{V \times \ldots V}_{m \text{ Faktoren}} \to K$$

heißt *Multilinearform*, m*-lineare Abbildung* oder kurz m*-Form*, falls sie in jedem ihrer Argumente linear ist, d.h. falls für alle $i \in \{1, \ldots, m\}$, alle $v_1, \ldots, v_n, w \in V$ und alle $\lambda \in K$ gilt

$$\begin{aligned}
&\alpha(v_1, \ldots, v_{i-1}, v_i + w, v_{i+1}, \ldots, v_m) = \\
&\quad = \alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_m) + \alpha(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_m)
\end{aligned}$$

und

$$\alpha(v_1, \ldots, v_{i-1}, \lambda v_i, v_{i+1}, \ldots, v_m) = \lambda\, \alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_m).$$

Beispiele 7.1.2. (i) Linearformen auf V, d.h. Elemente des Dualraums V^* sind 1-Formen.

(ii) Für eine $n \times n$-Matrix A ist durch $(x, y) \mapsto x^t A y$ ist eine bilineare Abbildung $K^n \times K^n \to K$ gegeben.

(iii) Durch die Abbildung $(v_1, \ldots, v_m) \mapsto \alpha(v_1) \cdot \ldots \cdot \alpha(v_m)$ ist für jeden Vektorraum V und jede Linearform $\alpha \in V^*$ eine Multilinearform $V^m \to K$ gegeben.

Hilfssatz 7.1.3. *Sei $\alpha\colon V \times \ldots \times V \to K$ eine m-lineare Abbildung und sei $b_1, \ldots, b_n$ eine Basis von V. Dann ist die Abbildung α bereits eindeutig durch ihre Werte auf allen Kombinationen von Basisvektoren, also durch die Werte*

$$\alpha(b_{i_1}, \ldots, b_{i_m}) \quad \textit{für} \quad i_1, \ldots, i_m \in \{1, \ldots, n\}$$

bestimmt.

Beweis. Seien $v_1, \ldots, v_m$ Vektoren aus V. Jeder Vektor v_j kann als Linearkombination der Basisvektoren $b_1, \ldots b_m$ dargestellt werden, d.h. es gilt

$$v_j = \sum_{i=1}^{n} a_{ji} b_i \quad \text{für} \quad a_{ji} \in K.$$

Der Wert von α auf dem m-Tupel von Vektoren $(v_1, \ldots, v_m)$ ergibt sich damit zu

$$\begin{aligned} \alpha(v_1, \ldots, v_m) &= \alpha(\sum_{i_1=1}^{n} a_{1i_1} b_{i_1}, \ldots, \sum_{i_m=1}^{n} a_{mi_m} b_{i_m}) = \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} a_{1i_1} \cdot \ldots \cdot a_{mi_m} \cdot \alpha(b_{i_1}, \ldots, b_{i_m}), \end{aligned}$$

was die Behauptung zeigt. $\square$

Definition 7.1.4. Eine Multilinearform α heißt *alternierend*, wenn sie Null ergibt, sobald zwei ihrer Argumente gleich sind, d.h. falls stets

$$\alpha(v_1, \ldots, v_m) = 0$$

gilt, wenn es zwei verschiedene Indizes $i, j \in \{1, \ldots, n\}$ gibt, so dass $v_i = v_j$ gilt.

Beispiel 7.1.5. Ist β eine Bilinearform $V \times V \to K$, so ist durch

$$\alpha(v, w) := \beta(v, w) - \beta(w, v)$$

eine alternierende Bilinearform $V \times V \to K$ definiert.

Hilfssatz 7.1.6. *Sei $\alpha\colon V \times \ldots \times V \to K$ eine alternierende Multilinearform. Dann gilt:*

(i) Vertauscht man zwei Argumente, so multipliziert sich der Wert von α mit -1, d.h. es gilt

$$\begin{aligned}&\alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_m) = \\ &\quad = -\alpha(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_m).\end{aligned}$$

(ii) Addiert man das λ-fache eines Arguments zu einem anderen Argument, so ändert sich der Wert von α nicht, d.h. es gilt

$$\begin{aligned}&\alpha(v_1, \ldots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \ldots, v_m) = \\ &\quad = \alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_m)\end{aligned}$$

für $i \neq j$.

(iii) Sind die Vektoren $v_1 \ldots, v_m$ linear abhängig, so gilt $\alpha(v_1 \ldots, v_m) = 0$.

Beweis. (i) Wir berechnen

$$\begin{aligned}0 &= \alpha(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_m) = \\ &= 0 + \alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_m) \\ &\quad + \alpha(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_m) + 0.\end{aligned}$$

(ii) Es gilt

$$\begin{aligned}&\alpha(v_1, \ldots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \ldots, v_m) = \\ &\quad = \alpha(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_m) \\ &\quad + \lambda \cdot \alpha(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_m),\end{aligned}$$

was die Behauptung zeigt, denn der letzte Term ist Null, da er zweimal das Argument v_j enthält.

(iii) Sind die Vektoren $v_1 \dots, v_m$ linear abhängig, dann lässt sich einer der Vektoren als Linearkombination der restlichen ausdrücken, etwa $v_i = \sum_{j=1}^m \lambda_j v_j$, wobei $\lambda_i = 0$ gilt, damit folgt

$$\alpha(v_1, \dots, v_{i-1}, \textstyle\sum_{j=1}^m \lambda_j v_j, v_{i+1}, \dots, v_m) = \\ = \displaystyle\sum_{j=1}^m \lambda_j \, \alpha(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_m).$$

Im letzten Term sind alle Summanden der Summe gleich Null, denn für $j \neq i$ kommen jeweils zwei gleiche Argumente in α vor und für $j = i$ gilt $\lambda_j = 0$.

□

Wir wollen in diesem Kapitel die Determinante einführen. Dies ist eine alternierende n-Form auf einem n-dimensionalen Vektorraum. In Hilfssatz 7.1.3 haben wir gezeigt, dass α bereits durch die Werte $\alpha(b_{i_1}, \dots, b_{i_n})$ für eine Basis $b_1, \dots b_n$ eindeutig bestimmt ist. Für eine alternierende Multilinearform sind diese Werte aber schon gleich Null, falls unter den Indizes $(i_1, \dots, i_n)$ der selbe Wert zweimal erscheint. Das bedeutet, dass $\alpha(b_{i_1}, \dots, b_{i_n}) = 0$ gilt, falls die Abbildung $(1, \dots, n) \mapsto (i_1, \dots, i_n)$ keine Permutation, siehe Abschnitt 2.1, ist. Außerdem gilt, dass sich höchstens das Vorzeichen ändert, wenn wir zwei der Argumente in $\alpha(b_{i_1}, \dots, b_{i_n})$ vertauschen. Durch Umordnen der Vektoren $(b_{i_1}, \dots, b_{i_n})$ können wir also den Wert von $\alpha(b_{i_1}, \dots, b_{i_n})$ aus dem von $\alpha(b_1, \dots, b_n)$ berechnen.

7.2 Die Signatur einer Permutation

Wir erinnern uns, dass eine Permutation $\sigma \in \mathrm{Sym}(n)$ eine *Umordnung* der Zahlen $1, \dots, n$ ist, also eine bijektive Abbildung

$$\{1, \dots, n\} \to \{1, \dots, n\}.$$

Zum Beispiel wird durch

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

die Abbildung

$$1 \mapsto 4; \quad 2 \mapsto 2; \quad 3 \mapsto 1; \quad 4 \mapsto 3$$

bezeichnet.

Definition 7.2.1. Sei $\sigma \in \mathrm{Sym}(n)$ eine Permutation und seien $1 \leq i < j \leq n$. Das Paar (i, j) heißt *Fehlstand* von σ, falls $\sigma(i) > \sigma(j)$ gilt. Eine Permutation heißt *gerade*, falls die Anzahl ihrer Fehlstände gerade ist, andernfalls *ungerade*. Wir setzen

$$\mathrm{sgn}(\sigma) := \begin{cases} +1, & \sigma \text{ gerade;} \\ -1, & \sigma \text{ ungerade.} \end{cases}$$

Der Wert $\mathrm{sgn}(\sigma)$ wird die *Signatur* der Permutation genannt.

Beispiel 7.2.2. Für die oben angegebene Permutation τ sind genau

$$(1, 2), (1, 3), (1, 4), (2, 3)$$

die Fehlstände. Es handelt sich also um eine gerade Permutation.

Hilfssatz 7.2.3. *Für jede Permutation $\sigma \in \mathrm{Sym}(n)$ gilt:*

$$\mathrm{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Beweis. Für jedes Paar (i, j) mit $1 \leq i < j \leq n$ enthält das Produkt auf der rechten Seite den Faktor

$$\frac{\sigma(i) - \sigma(j)}{i - j},$$

der genau dann positiv ist, wenn (i, j) ein Fehlstand ein von σ ist und sonst negativ. Somit ist das Vorzeichen des Produkts genau dann positiv, wenn σ eine gerade Permutation ist, und sonst negativ. $\square$

Satz 7.2.4. *Die Abbildung* $\operatorname{sgn}\colon \operatorname{Sym}(n) \to \{\pm 1\}$ *ist ein Gruppenhomomorphismus, wobei* $\operatorname{Sym}(n)$ *die Verkettung von Abbildungen als Verknüpfung hat und* $\{\pm 1\}$ *die übliche Multiplikation.*

Beweis. Es gilt

$$\begin{aligned} \operatorname{sgn}(\sigma \circ \tau) &= \prod_{1 \le i < j \le n} \frac{\sigma(\tau(i)) - \sigma((j))}{i - j} = \\ &= \prod_{1 \le i < j \le n} \frac{\sigma(\tau(i)) - \sigma((j))}{\tau(i) - \tau(j)} \cdot \prod_{1 \le i < j \le n} \frac{\tau(i) - \tau(j)}{i - j} = \\ &= \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau). \end{aligned}$$

Der linke Faktor in der vorletzten Zeile enthält dieselben Faktoren wie das zu $\operatorname{sgn}(\sigma)$ gehörige Produkt, nur in anderer Reihenfolge, da τ eine Permutation ist. □

Es folgt, dass die Verkettung zweier gerader Permutationen (oder zweier ungerader Permutationen) eine gerade Permutation ist und dass die Verkettung zweier Permutation mit verschiedener Signatur eine ungerade Permutation ist. Nach Satz 2.1.6 ist die Menge der geraden Permutationen eine Untergruppe von $\operatorname{Sym}(n)$, denn sie ist der Kern des Homomorphismus sgn.

Definition 7.2.5. Eine Permutation $\sigma \in \operatorname{Sym}(n)$ heißt *Transposition*, wenn es $i, j \in \{1, \ldots, n\}$ gibt, so dass gilt

$$\sigma(k) = \begin{cases} \text{j}, & \text{falls } k = i; \\ \text{i}, & \text{falls } k = j; \\ \text{k}, & \text{sonst.} \end{cases}$$

Die Transposition, die i auf j abbildet und umgekehrt, bezeichnen wir mit $\tau_{i,j}$.

Hilfssatz 7.2.6. *Transpositionen sind ungerade Permutationen.*

Beweis. Die Transposition

$$\tau_{i,j} = \begin{pmatrix} 1 & \ldots & i-1 & i & i+1 & \ldots & j-1 & j & j+1 & \ldots & n \\ 1 & \ldots & i-1 & j & i+1 & \ldots & j-1 & i & j+1 & \ldots & n \end{pmatrix}$$

mit $i < j$ hat genau die Fehlstände

$$(i, i+1), \ldots, (i, j-1), (i, j), (i+1, j), \ldots, (j-1, j),$$

somit eine ungerade Anzahl. $\square$

Satz 7.2.7. *Jede Permutation aus* $\mathrm{Sym}(n)$ *lässt sich als ein Produkt von* $n-1$ *oder weniger Transpositionen schreiben.*

Beweis. Wir zeigen dies durch Induktion über n. Für $n = 1$ ist nichts zu zeigen. Sei nun $\sigma \in \mathrm{Sym}(n)$, $n \geq 2$. Wir unterscheiden zwei Fälle. Ist $\sigma(n) = n$, wähle $\pi := \mathrm{id}_{\{1,\ldots,n\}}$. Gilt $\sigma(n) = i \neq n$, wähle $\pi := \tau_{i,n}$. Nun gilt in beiden Fällen $\pi \circ \sigma(n) = n$. Nach Induktionsvoraussetzung lässt sich nun

$$\pi \circ \sigma|_{\{1,\ldots,n-1\}}$$

als Produkt von höchstens $n-2$ Transpositionen $\rho_1 \circ \ldots \circ \rho_k \in \mathrm{Sym}(n-1)$, $k \leq n-2$ schreiben, die wir mit der Vereinbarung $\rho_j(n) = n$ auch als Elemente von $\mathrm{Sym}(n)$ auffassen können. Dies zeigt, dass $\sigma = \pi^{-1} \circ \rho_1 \circ \ldots \circ \rho_k = \pi \circ \rho_1 \circ \ldots \circ \rho_k$ das Produkt von $n-1$ oder weniger Transpositionen ist. $\square$

Aus Satz 7.2.4 folgt: Schreibt man eine (un)gerade Permutation als Produkt von Transpositionen, so ist die Anzahl der Faktoren (un)gerade.

7.3 Die Determinante

Hilfssatz 7.3.1. *Sei* V *ein Vektorraum und sei* $b_1, \ldots, b_n$ *eine Basis von* V*. Sei* $\alpha\colon V^n \to K$ *eine alternierende* n*-Form. Seien* $v_1, \ldots, v_n$ *Vektoren aus* V*, die als Linearkombinationen* $v_j = \sum_{i=1}^n a_{ij} b_i$ *mit* $a_{ij} \in K$ *dargestellt werden. Dann gilt*

$$\alpha(v_1, \ldots, v_n) = \sum_{\sigma \in \mathrm{Sym}(n)} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \cdot \mathrm{sgn}(\sigma) \cdot \alpha(b_1, \ldots, b_n).$$

Insbesondere ist die n*-Form bereits durch den Wert* $\alpha(b_1, \ldots, b_n)$ *eindeutig festgelegt.*

Beweis. Wir formen um:

$$\begin{aligned}
\alpha(v_1,\dots,v_n) &= \sum_{i_1=1}^{n}\dots\sum_{i_n=1}^{n} a_{1i_1}\cdot\ldots\cdot a_{ni_n}\cdot\alpha(b_{i_1},\dots,b_{i_n}) = \\
&= \sum_{\sigma\in\mathrm{Sym}(n)} a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)}\cdot\alpha(b_{\sigma(1)},\dots,b_{\sigma(n)}) = \\
&= \sum_{\sigma\in\mathrm{Sym}(n)} a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)}\cdot\mathrm{sgn}(\sigma)\cdot\alpha(b_1,\dots,b_n)
\end{aligned}$$

Die erste Umformung folgt aus Lemma 7.1.3. Die zweite ist möglich, da α eine alternierende n-Form ist; daher sind alle Terme $\alpha(b_{i_1},\dots,b_{i_n})$ gleich Null, für welche die Abbildung $(1,\dots,n)\mapsto(i_1,\dots,i_n)$ keine Permutation ist und wir können uns darauf beschränken, über die wiederholungsfreien n-Tupel $(i_1,\dots,i_n)$ zu summieren. Die letzte Umformung besteht darin, die Vektoren $b_{\sigma(1)},\dots,b_{\sigma(n)}$ im Argument von α in die "richtige" Reihenfolge $b_1,\dots,b_n$ zu bringen. Nach Satz 7.2.7 erhalten wir dabei den Faktor $\mathrm{sgn}(\sigma)$. $\square$

Sei nun A eine $n\times n$-Matrix mit Einträgen in K. Indem wir die Spaltenvektoren von A betrachten, können wir die Matrix auch als ein n-Tupel von Vektoren im K^n auffassen. Wir definieren also:

$$\alpha\begin{pmatrix} a_{11} & \dots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \dots & a_{nn}\end{pmatrix} := \alpha\left(\begin{pmatrix} a_{11}\\ \vdots\\ a_{n1}\end{pmatrix},\dots,\begin{pmatrix} a_{1n}\\ \vdots\\ a_{nn}\end{pmatrix}\right).$$

Wir nennen eine alternierende n-Form α *normiert*, wenn Sie auf der Einheitsmatrix den Wert 1 annimmt, bzw. wenn gilt:

$$\alpha(E)=\alpha(e_1,\dots,e_n)=1,$$

Nach dem vorangehenden Satz gibt es nur eine solche alternierende n-Form auf dem K^n. Diese nennen wir die *Determinante*.

Mit Satz 7.3.1 erhalten wir die folgende explizite Formel für die Determinante:

Definition 7.3.2. Sei $A = (a_{ij}) \in M_n(K)$. Dann nennen wir

$$\det(A) := \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \in K$$

die *Determinante* von A.

Die in unserer Definition der Determinante verwendete Formel wird auch die *Leibnizsche Formel* für die Determinante genannt.

Beispiel 7.3.3. Wir wollen die Formeln für die Determinante für $n = 1, 2, 3$ explizit ausschreiben.

(i) Für $n = 1$ hat die Summe nur einen Summanden, die identische Permutation und es gilt

$$\det(a_{11}) = a_{11}.$$

(ii) Für $n = 2$ haben wir zwei Summanden, die identische Abbildung auf $\{1, 2\}$, und die Transposition, die 1 mit 2 vertauscht. Wir erhalten

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

Beispielsweise gilt

$$\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2.$$

(iii) Für den Fall $n = 3$, notieren wir uns zunächst alle Elemente von $\mathrm{Sym}(3)$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Die ersten drei Elemente sind hier gerade Permutationen, die letzten drei ungerade. Damit erhalten wir:

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

Diese Formel wird auch *Regel von Sarrus* genannt, man kann sie sich durch die folgende, sogenannte *Jägerzaunregel* leicht einprägen.

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

Man erweitert dabei die 3×3-Matrix, indem man die beiden ersten Spalten noch einmal rechts anfügt. Man bildet jeweils die Produkte von den drei Einträgen, durch die ein Pfeil verläuft, addiert die Summe der derjenigen drei Produkte, die zu den nach rechts unten zeigenden Pfeilen gehören und subtrahiert davon die Summe der derjenigen drei Produkte, die zu den nach rechts oben zeigenden Pfeilen gehören.

Warnung: Für $n \times n$-Matrizen mit $n \geq 4$ ist eine entsprechende Regel *nicht* gültig.

Wir zeigen nun, dass sich die Determinante $\det(A)$ nicht ändert, wenn man die Matrix A durch ihre Transponierte A^t ersetzt, dass also die Determinanten von

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix}$$

gleich sind.

Hilfssatz 7.3.4. *Sei $A \in M_n(K)$. Dann gilt für die Transponierte A^t von A:*

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

Beweis. Wir bemerken zunächst, dass für jede Permutation $\sigma \in \mathrm{Sym}(n)$ gilt:

$$a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = a_{1\sigma(1)^{-1}} \cdot \ldots \cdot a_{n\sigma(n)^{-1}},$$

denn die beiden Produkte stimmen bis auf die Reihenfolge der Faktoren überein. Außerdem folgt aus Satz 7.2.4, dass $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ gilt. Damit erhalten wir

$$\begin{aligned}
\det(A^t) &= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} = \\
&= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)} = \\
&= \sum_{\sigma^{-1} \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} = \\
&= \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}.
\end{aligned}$$

Die letzte Umformung ändert nur die Reihenfolge der Summation. □

Satz 7.3.5 (Determinantenmultiplikationssatz)**.** *Seien* $A, B \in M_n(K)$. *Dann gilt*

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

Beweis. Sei $A \in M_n(K)$. Wir zeigen: Die beiden Abbildungen $(K^n)^n \to K$, gegeben durch einerseits

$$(v_1, \ldots, v_n) \mapsto \det(Av_1, \ldots, Av_n)$$

und andererseits

$$(v_1, \ldots, v_n) \mapsto \det(A) \cdot \det(v_1, \ldots, v_n)$$

sind beide alternierende Multilinearformen. Die Linearität für die erste Abbildung daraus, dass $v_i \mapsto Av_i$ lineare Abbildungen sind; gilt außerdem $v_i = v_j$ für $i \neq j$, so folgt $\det(Av_1, \ldots, Av_n) = 0$. Die zweite Abbildung ist lediglich die Determinante von $v_1, \ldots, v_n$, mit dem Skalar $\det(A)$ multipliziert. Setzt man die kanonische Basis $e_1, \ldots, e_n$ ein, so erhält man in beiden Fällen den Wert $\det(A)$. Nach Lemma 7.3.1 stimmen die beiden Abbildungen also überein. Die Behauptung folgt, indem man in beiden Abbildungen die Spalten von B für die Vektoren $v_1, \ldots, v_n$ einsetzt. □

Folgerung 7.3.6. *Eine Matrix $A \in M_n(K)$ ist genau dann invertierbar, wenn* $\det(A) \neq 0$ *ist. In diesem Fall gilt* $\det(A^{-1}) = \det(A)^{-1}$.

Beweis. Ist A invertierbar, so folgt aus dem Determinantenmultiplikationssatz

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

Ist A nicht invertierbar, so folgt aus Hilfssatz 7.1.6 (iii), dass $\det(A) = 0$ gilt. $\square$

Satz 7.3.7. *Seien $A \in M_k(K)$, $B \in M_{n-k}(K)$ und $C \in M_{k,n-k}(K)$. Dann gilt*

$$\det \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array} \right) = \det(A) \cdot \det(B)$$

Beweis. Sei

$$D := \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array} \right)$$

Dann gilt nach der Definition der Determinante,

$$\det(D) = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \cdot d_{1\sigma(1)} \cdot \ldots \cdot d_{n\sigma(n)}$$

Wir betrachten die Summanden in der Summe auf der rechten Seite nun einzeln. Sei $\sigma \in \mathrm{Sym}(n)$ eine Permutation. Dann gibt es zwei Fälle. Entweder die Abbildung $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ lässt die Teilmenge $\{1, \ldots, k\} \subset \{1, \ldots, n\}$ invariant oder nicht, d.h. entweder für alle $i \in \{1, \ldots, k\}$ gilt $\sigma(i) \leq k$ oder für mindestens ein $i \in \{1, \ldots, k\}$ gilt $\sigma(i) \geq k + 1$.

Im ersten Fall gilt, da die Abbildung σ bijektiv ist, dass σ auch die Teilmenge $\{k + 1, \ldots, n\}$ auf sich abbildet. In diesem Fall gibt es also Permutationen $\rho \in \mathrm{Sym}(k)$ und $\tau \in \mathrm{Sym}(n - k)$, so dass

$$\sigma(i) = \begin{cases} \rho(i), & \text{falls } i \leq k; \\ \tau(i - k) + k, & \text{falls } i \geq k + 1. \end{cases}$$

gilt. Außerdem gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\rho) \cdot \operatorname{sgn}(\tau)$, was man z.B. an der Formel in Hilfssatz 7.2.3 sehen kann, denn Paare (i, j) mit $i \leq k < j$ sind keine Fehlstände von σ. Damit erhalten wir für den für den zu σ gehörenden Summanden

$$\begin{aligned} &\operatorname{sgn}(\sigma) \cdot d_{1\sigma(1)} \cdot \ldots \cdot d_{n\sigma(n)} = \\ &\qquad = \operatorname{sgn}(\rho) \cdot a_{1\rho(1)} \cdot \ldots \cdot a_{k\rho(k)} \cdot \operatorname{sgn}(\tau) \cdot b_{1\tau(1)} \cdot \ldots \cdot b_{n-k,\tau(n-k)}. \end{aligned} \tag{7.3.1}$$

Im Fall, dass es mindestens ein $i \in \{1, \ldots, k\}$ gibt mit $\sigma(i) \geq k + 1$ muss es auch mindestens ein $i \in \{k + 1, \ldots, n\}$ geben mit $\sigma(i) \leq k$. Wegen der Gestalt der Matrix D folgt, dass der Matrixeintrag $d_{i\sigma(i)}$ gleich Null ist. Daher können wir den zu σ gehörenden Summanden aus der Summe weglassen und die Summe erstreckt sich nur noch über die Summanden auf der rechten Seite in (7.3.1) mit $\rho \in \operatorname{Sym}(k)$, $\tau \in \operatorname{Sym}(n - k)$. Dies zeigt $\det(D) = \det(A) \cdot \det(B)$. □

Folgerung 7.3.8. *Sei $A \in M_n(K)$ eine Matrix, die obere Dreiecksgestalt hat, d.h. es gilt $A_{ij} = 0$ für $i > j$. Dann gilt für die Determinante*

$$\det(A) = \det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{pmatrix} = a_{11} \cdot \ldots \cdot a_{nn},$$

d.h. die Determinante einer oberen Dreiecksmatrix ist das Produkt ihrer Diagonalelemente. (Wegen $\det(A^t) = \det(A)$ gilt dasselbe für untere Dreiecksmatrizen.)

Beweis. Folgt durch Induktion aus Satz 7.3.7. □

Ein alternativer direkter Beweis ergibt sich daraus, dass man mit einem ähnlichen Argument wie im Beweis von Satz 7.3.7 einsieht, dass in der Formel für die Definition der Determinante der einzige Summand, der nicht verschwindet, der zur identischen Permutation gehörige ist.

Bemerkung 7.3.9 (Determinantenberechnung mit dem Gaußalgorithmus). Hilfssatz 7.1.6 zeigt, dass sich der Wert der Determinante nicht ändert, wenn man das λ-fache einer Zeile (oder Spalte) zu einer anderen Zeile (oder Spalte) addiert und dass sich der Werte der Determinante nur um den Faktor -1 ändert, wenn man zwei Zeilen (oder Spalten) vertauscht. Dies zeigt, dass wir auch den Gaußalgorithmus verwenden können, um eine Determinante zu berechnen: Wir bringen eine quadratische Matrix mit Hilfe der gerade genannte elementaren Zeilenumformungen auf obere Dreiecksgestalt; die Determinante der so entstandenen Matrix ist das Produkt der Diagonalelmente und sie unterscheidet sich von der Determinante der ursprünglichen Matrix höchstens um ein Vorzeichen; wir müssen uns beim Berechnen also merken, ob wir eine gerade oder ungerade Anzahl von Zeilenvertauschungen vorgenommen haben.

(Ein alternativer Beweis für die Gültigkeit dieses Verfahrens ergibt sich aus dem Determinantenmultiplikationssatz, wenn man die Umformungsschritte des Gaußalgorithmaus durch Multiplikation mit Elementarmatrizen realisiert.)

Beispiele 7.3.10. Wir wollen die Bemerkung 7.3.9 beschriebene Methode zur Determinantenberechnung an einigen Beispielen veranschaulichen. Durch eine Zeilenumformung erhält man etwa

$$\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 8 & 3 & 4 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 8 & 3 & 4 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} = 0.$$

Ein weiteres Beispiel:

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

$$= -\det\begin{pmatrix} 1 & 3 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} = (-1)\cdot 3\cdot(-\tfrac{1}{3}) = 1.$$

Ebenso erhält man durch Zeilenvertauschungen für die folgende $n \times n$-Matrix

$$A = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix},$$

dass $\det(A) = (-1)^{\lfloor \frac{n}{2} \rfloor}$ gilt. Dabei bezeichnet die Schreibweise $\lfloor \frac{n}{2} \rfloor$, dass $\frac{n}{2}$ auf die nächstkleinere ganze Zahl abgerundet wird. Es sind nämlich genau $\lfloor \frac{n}{2} \rfloor$ Zeilenvertauschungen sind erforderlich, um aus A die $n \times n$-Einheitsmatrix zu machen.

Wir wollen noch eine weitere, sehr nützliche Formel für die Berechnung von Determinanten herleiten.

Satz 7.3.11 (Laplacescher Entwicklungssatz)**.** *Sei $A \in M_n(K)$ eine quadratische Matrix und sei ein Spaltenindex $j \in \{1, \dots, n\}$ gewählt. Dann gilt*

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

wobei

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

die $(n-1) \times (n-1)$-Matrix bezeichnet, die aus A entsteht, indem man die i-te Zeile und die j-te Spalte streicht.

Beweis. Seien $s_1, \dots, s_n$ die Spalten von A. Dann gilt für die j-te Spalte

$$s_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Dies setzen wir nun ein in

$$\begin{aligned} \det(A) &= \det(s_1, \dots, s_n) = \\ &= \det(s_1, \dots, s_{j-1}, \textstyle\sum_{i=1}^{n} a_{ij} e_i, s_{j+1}, \dots, s_n) = \\ &= \sum_{i=1}^{n} a_{ij} \det(s_1, \dots, s_{j-1}, e_i, s_{j+1}, \dots, s_n). \end{aligned}$$

Um den Beweis zu beenden, genügt es nun, zu zeigen, dass

$$\det(s_1, \ldots, s_{j-1}, e_i, s_{j+1}, \ldots, s_n) = (-1)^{i+j} \det(A_{ij})$$

gilt. Um zu sehen, dass diese Formel richtig ist, führen $j-1$ Spaltenvertauschungen so durch, dass wir die Matrix $\det(e_i, s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_n)$ erhalten. Auf analoge Weise formen wir die wir die so entstandene Matrix durch $i-1$ Zeilenvertauschungen so um, dass wir schließlich die Matrix

$$\left(\begin{array}{c|c} 1 & * \ldots * \\ \hline & \\ 0 & A_{ij} \\ & \end{array}\right)$$

erhalten. Diese hat nach Satz 7.3.7 die Determinante $\det(A_{ij})$. Durch die Spalten- und Zeilenvertauschungen haben wir den Faktor $(-1)^{i-1} \cdot (-1)^{j-1} = (-1)^{i+j}$, was den Beweis abschließt. □

Folgerung 7.3.12. *Wegen* $\det(A^t) = \det(A)$ *gilt auch die Entwicklungsformel für die Entwicklung nach einer Zeile*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

In den nachfolgenden Berechnungen verwenden wir die Kurzschreibweise

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

für die Determinante einer Matrix.

Beispiel 7.3.13. Im folgenden Beispiel wird die Determinante im ersten Schritt durch Entwickeln nach der zweiten Spalte berechnet.

$$\begin{vmatrix} 1 & 0 & 2 & 1 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 1 & 3 \\ 2 & 1 & 1 & 1 \end{vmatrix} = +1 \cdot \begin{vmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \end{vmatrix} = -2 \cdot \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} =$$

$$= -2 \cdot (6-1) - 1 \cdot (1-2) = -10 + 1 = -9$$

Beim zweiten Umformungsschritt wird dann nach der zweiten Zeile entwickelt.

7.4 Die Cramersche Regel und die Formel für die Inverse einer Matrix

Definition 7.4.1. Sei A eine $n \times n$-Matrix. Die Terme

$$(-1)^{i+j} \det(A_{ij})$$

heißen *Kofaktoren* der Matrix A. Die Matrix

$$\tilde{A} := (\tilde{a}_{ij}) = (-1)^{i+j} \det(A_{ij})$$

heißt *Ko(faktor)matrix* von A. Die Transponierte der Ko(faktor)matrix heißt *Adjunkte* der Matrix A.

Hilfssatz 7.4.2. *Sei $A \in M_n(K)$ eine quadratische Matrix. Dann gilt*

$$\tilde{A}^t \cdot A = \det(A) \cdot E.$$

Beweis. Wir berechnen den Eintrag in der i-ten Zeile und k-ten Spalte der Matrix, die durch das Matrizenprodukt auf der linken Seite gegeben ist. Nach Definition der Matrizenmultiplikation gilt

$$\sum_{j=1}^{n} (-1)^{i+j} \det(A_{ji}) \cdot a_{jk}.$$

Falls $i = k$ ist das genau die Formel für die Entwicklung der Determinanten von A^t nach der i-ten Spalte. Die Einträge auf der Diagonalen der Matrix $\tilde{A}^t \cdot A$ sind also alle gleich $\det(A)$. Falls $i \neq k$ ist dies die Entwicklung der Determinanten der Matrix, die aus A^t entsteht, wenn man die i-te Spalte von A^t durch die k-te Spalte von A^t ersetzt. Die Einträge außerhalb der Diagonalen der Matrix $\tilde{A}^t \cdot A$ sind also alle gleich Null. □

Daraus erhält man eine explizite Formel für die Inverse einer Matrix.

Satz 7.4.3. *Ist $A \in M_n(K)$ invertierbar, so gilt*

$$A^{-1} = \frac{1}{\det(A)} \cdot \tilde{A}^t.$$

Beweis. Folgt direkt aus Lemma 7.4.2. $\square$

Beispiele 7.4.4. (i) Sei die 2×2-Matrix

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

invertierbar. Wir erhalten für die Komatrix von A dann

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

und somit für die Inverse

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(ii) Wir wollen nun auch eine Formel für die Inverse einer beliebigen 3×3-Matrix herleiten. Sei

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \in M_3(K)$$

invertierbar. Dann gilt

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} +\begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & +\begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ & & \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & +\begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ & & \\ +\begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & +\begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{pmatrix} =$$

$$= \frac{1}{\det(A)} \begin{pmatrix} +ei - hf & -bi + hc & +bf - ec \\ -di + gf & +ai - gc & -af + dc \\ +dh - ge & -ah + gb & +ae - db \end{pmatrix}.$$

(iii) Wir berechnen die Inverse der Matrix

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

mit Hilfe von Satz 7.4.3. Wir lesen direkt ab, dass $\det(A) = 2$ gilt. Wir erhalten

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

Satz 7.4.5 (Cramersche Regel)**.** *Sei $A \in M_n(K)$ invertierbar und $b \in K^n$. Dann ist die (eindeutig bestimmte) Lösung x des Gleichungssystems $Ax = b$ gegeben durch*

$$x_i = \frac{\det(A_{i,b})}{\det(A)},$$

wobei $A_{i,b}$ die Matrix bezeichnet, die aus A entsteht, indem man die i-te Spalte von A durch den Vektor b ersetzt.

Beweis. Seien $s_1, \ldots, s_n$ die Spalten von A und sei x die Lösung von $Ax = b$. Dann gilt $\sum_{j=1}^n x_j s_j$. Es folgt

$$\begin{aligned} \det(A_{i,b}) &= \det(s_1, \ldots, s_{i-1}, b, s_{i+1}, \ldots, s_n) = \\ &= \det(s_1, \ldots, s_{i-1}, \textstyle\sum_{j=1}^n x_j s_j, s_{i+1}, \ldots, s_n) = \\ &= \textstyle\sum_{j=1}^n x_j \det(s_1, \ldots, s_{i-1}, s_j, s_{i+1}, \ldots, s_n) = \\ &= x_i \det(s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n). \end{aligned}$$

Die letzte Gleichheit gilt, da die alternierende Multilinearform $\det$ verschwindet, wenn zwei gleiche Argumente eingesetzt werden. $\square$

Beispiel 7.4.6. Seien

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in M_2(\mathbb{Q}) \quad \text{und} \quad b = \begin{pmatrix} 5 \\ 6 \end{pmatrix} \in \mathbb{Q}^2$$

gegeben. Da $\det(A) = -2$ gilt, ist A invertierbar und die eindeutig bestimmte Lösung der Gleichung $Ax = b$ ist gegeben durch

$$x_1 = \frac{\begin{vmatrix} 5 & 2 \\ 6 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}} = \frac{8}{-2} = -4 \quad \text{und} \quad x_2 = \frac{\begin{vmatrix} 1 & 5 \\ 3 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}} = \frac{-9}{-2} = \frac{9}{2}.$$

Kapitel 8

Eigenwerte

In diesem Kapitel geht es darum, darstellende Matrizen von Endomorphismen durch Basiswechsel auf eine möglichst einfache Gestalt zu bringen. “Möglichst einfach” bedeutet in diesem Zusammenhang, dass die darstellende Matrix möglichst viele Nullen enthalten soll, idealerweise z.B. Diagonalgestalt annimmt. Dazu benutzt man den Begriff des Eigenwerts.

8.1 Eigenwerte und Eigenvektoren

Definition 8.1.1. Sei $f\colon V \to V$ ein linearer Endomorphismus. Ein Skalar λ heißt *Eigenwert* von f, wenn es einen von Null verschiedenen Vektor $v \in V$ gibt, so dass gilt

$$f(v) = \lambda\, v.$$

In diesem Fall sagen wir, v ist ein *Eigenvektor von f (zum Eigenwert λ)*.

Beispiele 8.1.2. (i) Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch die Matrix

$$A = \begin{pmatrix} 1 & 2 & 2 \\ -3 & 6 & 3 \\ 5 & 3 & 2 \end{pmatrix}.$$

Dann ist der Vektor

$$v = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

ein Eigenvektor von f zum Eigenwert 3, denn es gilt

$$Av = \begin{pmatrix} 3 \\ -3 \\ 6 \end{pmatrix} = 3v.$$

(ii) Jeder von Null verschiedene Vektor im Kern eines linearen Endomorphismus ist ein Eigenvektor zum Eigenwert 0.

(iii) Sei $f = \lambda \cdot \mathrm{id}_V$. Dann ist jeder Vektor in $V \setminus \{0\}$ ein Eigenvektor von f zum Eigenwert λ.

(iv) Sei V der reelle Vektorraum der beliebig oft differenzierbaren Funktionen $\mathbb{R} \to \mathbb{R}$. Dann ist die Exponentialfunktion $\exp\colon \mathbb{R} \to \mathbb{R}$ ein Eigenvektor zum Eigenwert 1 des Ableitungsoperators $V \to V, f \mapsto f'$, denn es gilt bekanntlich

$$\frac{d}{dt}\exp(t) = \exp(t).$$

Definition 8.1.3. Sei $f\colon V \to V$ ein linearer Endomorphismus. Ein Untervektorraum $U \subseteq V$ heißt *invariant* (genauer f*-invariant*), wenn $f(U) \subseteq U$ gilt.

Dass ein linearer Endomorphismus einen Eigenvektor besitzt, ist gleichbedeutend damit, dass es einen eindimensionalen invarianten Untervektorraum gibt. Nicht jede lineare Abbildung besitzt einen Eigenvektor, z.B. lässt eine Drehung im $\mathbb{R}^2$, gegeben durch die Matrix

$$\begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}, \quad t \in \mathbb{R},$$

nur dann eine Gerade durch den Ursprung invariant, wenn t ein Vielfaches von π ist, wenn es sich also um die Punktspiegelung am Ursprung oder die identische Abbildung handelt.

Hilfssatz 8.1.4. *Ein linearer Endomorphismus $f\colon V \to V$ hat genau dann den Eigenwert $\lambda \in K$, wenn der Endomorphismus*

$$f - \lambda \cdot \mathrm{id}_V \colon V \to V$$

einen von Null verschiedenen Kern hat.

Beweis. Es gilt

$$f(v) = \lambda\, v \iff f(v) - \lambda \cdot v = 0 \iff (f - \lambda \cdot \mathrm{id}_V)(v) = 0. \qquad \square$$

Wir formulieren den Hilfssatz um für den Fall, dass der Endomorphismus durch eine Matrix gegeben ist.

Folgerung 8.1.5. *Sei $A \in M_n(K)$ eine quadratische Matrix. Dann hat die lineare Abbildung $K^n \to K^n, x \mapsto Ax$ genau dann den Eigenwert $\lambda \in K$, wenn die Matrix*

$$A - \lambda \cdot E$$

nicht invertierbar ist.

Beweis. Folgt direkt aus Hilfssatz 8.1.4. $\square$

Dies liefert uns eine Methode, mit der man in vielen Fällen entscheiden kann, ob eine lineare Abbildung Eigenwerte besitzt und diese gegebenenfalls auch bestimmen kann. Nämlich hat die Abbildung im obigen Hilfssatz genau dann den Eigenwert λ, wenn gilt

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

Beispiel 8.1.6. Sei
$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in M_2(\mathbb{R}).$$
Wir betrachten die Abbildung
$$\begin{aligned} t \mapsto \det\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - t \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) &= \begin{vmatrix} 1-t & 2 \\ 3 & 4-t \end{vmatrix} = \\ = (1-t)(4-t) - 6 = t^2 - 5t - 2, \end{aligned}$$
Die Nullstellen dieser Abbildung sind $\lambda_{1,2} = \frac{1}{2}(5 \pm \sqrt{33})$ und dies sind genau die Eigenwerte der obigen linearen Abbildung.

Definition 8.1.7. Sei $f\colon V \to V$ ein linearer Endomorphismus und $\lambda \in K$. Wir definieren
$$V^\lambda := \{v \in V \mid f(v) = \lambda\, v\}.$$
Falls λ ein Eigenwert von f ist, nennen wir V^λ den Eigenraum von f zum Eigenwert λ. Falls λ ein Eigenwert von f ist, dann schreibt man auch $E_\lambda(f)$ bzw. kurz E_λ anstatt V^λ.

Die Menge V^λ besteht aus der Menge aller Eigenvektoren von f zum Eigenwert λ, vereinigt mit $\{0\}$. Genau dann, wenn λ ein Eigenwert von f ist, gilt $V^\lambda \neq \{0\}$. Da V^λ der Kern der linearen Abbildung $f - \lambda \cdot \mathrm{id}_V$ ist, ist klar, dass V^λ stets ein Untervektorraum von V ist.

Beispiel 8.1.8. Im Beispiel 8.1.6 von oben sind die zu den beiden Eigenwerten $\lambda_{1,2}$ gehörigen Eigenräume gegeben durch
$$\ker\begin{pmatrix} -\frac{3}{2} - \frac{1}{2}\sqrt{33} & 2 \\ 3 & \frac{3}{2} - \frac{1}{2}\sqrt{33} \end{pmatrix} = \mathbb{R} \cdot \begin{pmatrix} 1 \\ \frac{3}{4} + \frac{1}{4}\sqrt{33} \end{pmatrix}$$
und
$$\ker\begin{pmatrix} -\frac{3}{2} + \frac{1}{2}\sqrt{33} & 2 \\ 3 & \frac{3}{2} + \frac{1}{2}\sqrt{33} \end{pmatrix} = \mathbb{R} \cdot \begin{pmatrix} 1 \\ \frac{3}{4} - \frac{1}{4}\sqrt{33} \end{pmatrix}.$$

8.2 Das charakteristische Polynom

Im folgenden wollen wir die Eigenwerte und Eigenvektoren von linearen Endomorphismen auf endlich-dimensionalen Vektorräumen studieren. Diese sind (nach Wahl einer Basis) durch quadratische Matrizen gegeben. Der Einfachheit halber identifizieren wir lineare Abbildungen $K^n \to K^n$ ab jetzt mit den darstellenden $n \times n$-Matrizen bezüglich der kanonischen Basis. Deshalb können von nun auch von den *Eigenwerten, Eigenvektoren etc. einer Matrix* $A \in M_n(K)$ sprechen.

Bemerkung 8.2.1. Im Folgenden werden wir Polynome benutzen, um lineare Endomorphismen zu studieren. Dabei ist es nötig die folgende Unterscheidung zu treffen. Man kann ein Polynom wie in Definition 8.2.2 unten

$$a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$$

als eine Folge von Koeffizienten $a_1, a_1, a_2, \ldots$ von denen nur endlich viele von Null verschieden sind, oder als eine *polynomiale Abbildung*, d.h. als die Funktion, die durch die Abbildungsvorschrift

$$t \mapsto a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0, \quad K \to K$$

gegeben ist. Im Fall eines endlichen Körpers unterscheiden sich die beiden Sichtweisen erheblich voneinander, denn wenn es nur endlich viele Elemente in K gibt, kann es auch nur endlich viele verschiedene Abbildungen $K \to K$ geben. Dennoch gibt es aber unendlich viele Polynome mit Koeffizienten in K. Wir wollen im Folgenden Polynome als Folgen von Koeffizienten auffassen, nicht als Abbildungsvorschriften. Im Falle eines unendlichen Körpers unterscheiden sich die beiden Sichtweisen nicht, denn jedes formale Polynom ist in diesem Fall durch die durch definierte Abbildungsvorschrift eindeutig festgelegt. (Dies kann man so einsehen: Die Differenz von zwei Polynomen, welche die gleiche Abbildungsvorschrift liefern, ist ein Polynom mit unendlich vielen Nullstellen. Mit Hilfe der Polynomdivision kann man aber zeigen, dass ein vom Nullpolynom verschiedenes Polynom n-ten Grades höchstens n verschiedene Nullstellen besitzt.)

Definition 8.2.2. Sei K ein Körper. Dann definieren wir

$$K[t] := \{a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0 \mid n \in \mathbb{N}_0, a_0, \ldots, a_n \in K\}$$

als die Menge der Polynome in der formalen Variablen t.

Wir können $K[t]$ mit der Menge der Folgen $a_0, a_1, a_2, \ldots$ in K, bei denen nur endliche viele Folgenglieder von Null verschieden sind, identifizieren.

Definition 8.2.3. Sei $A \in M_n(K)$ eine quadratische Matrix. Dann heißt die Matrix

$$A - tE = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} - \begin{pmatrix} t & & \\ & \ddots & \\ & & t \end{pmatrix}$$

mit Einträgen in $K[t]$ *charakteristische Matrix von A* und das Polynom

$$\chi_A(t) := \det(A - t \cdot E) \in K[t]$$

heißt *charakteristisches Polynom von A.*

Wir haben bereits gesehen:

Satz 8.2.4. *Sei $A \in M_n(K)$. Dann sind die Eigenwerte von A genau die Nullstellen des charakteristischen Polynoms. Ist $\lambda \in K$ ein Eigenwert von A, so ist der Eigenraum A zum Eigenwert λ gegeben durch den Kern der charakteristischen Matrix $A - \lambda\, E$.*

In den Beispielen 8.1.6 und 8.1.8 wurden das charakteristische Polynom und die charakteristische Matrix benutzt, um Eigenwerte und Eigenräume zu bestimmen.

Als unmittelbare Folgerung aus Satz 8.2.4 erhalten wir, dass es im Fall $K = \mathbb{C}$ stets Eigenwerte gibt.

Satz 8.2.5. *Sei A eine komplexe $n \times n$-Matrix, $n \geq 1$. Dann hat A einen Eigenwert.*

Beweis. Dies folgt direkt aus Satz 8.2.4 und dem Hauptsatz der Algebra (den wir nicht beweisen). Dieser besagt nämlich, dass jedes nicht konstante komplexe Polynom mindestens eine Nullstelle hat. $\square$

Für eine quadratische Matrix A definieren wir die *Spur* $\mathrm{tr}(A)$ als die Summe der Diagonalelemente, d.h.

$$\mathrm{tr}\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} := a_{11} + \ldots + a_{nn}.$$

Satz 8.2.6. *Sei $A \in M_n(K)$ und sei $\chi_A \in K[t]$ das charakteristische Polynom von A. Dann hat χ_A den* Grad *n, d.h. a_n ist unter den nicht-verschwindenden Koeffizienten derjenige mit dem höchstem Index. Außerdem gilt*

$$a_n = (-1)^n, \quad a_{n-1} = (-1)^{n-1}\,\mathrm{tr}(A) \quad \textit{und} \quad a_0 = \det(A).$$

Beweis. Wir verwenden die Leibnizsche Formel für die Determinante zum Beweis. Der einzige Summand in der Formel, der zur identischen Permutation gehört, ist $(a_{11} - t)(a_{22} - t) \cdot \ldots \cdot (a_{nn} - t)$. Beim Ausmultiplizieren dieses Produkts entsteht ein Summand $(-t)^n$, die anderen Terme haben niedrigere Grade. Die zu anderen Permutationen gehörenden Summanden sind ebenfalls Polynome von niedrigeren Graden. Dies zeigt $a_n = (-1)^n$.

Um $a_{n-1} = (-1)^{n-1}\,\mathrm{tr}(A)$ zu zeigen, überlegen wir uns zunächst noch, dass die Summanden in der Leibnizschen Formel, die nicht zur identischen Permutation gehören, Polynome vom Grad höchstens $n-2$. Denn bei einer nicht-identischen Permutation werden mindestens zwei Elemente aus $\{1, \ldots, n\}$ nicht auf sich selbst abgebildet. Der Koeffizient vor t^{n-1} ist also derselbe wie in dem Polynom $(a_{11} - t) \cdot \ldots \cdot (a_{nn} - t)$ und also gleich $(-t)^{n-1}a_{11} + \cdots + (-t)^{n-1}a_{nn}$.

Das konstante Glied a_0 ist der Wert, den das charakteristische Polynom bei $\lambda = 0$ annimmt. Dies zeigt $a_0 = \det(A)$. $\square$

Wir notieren noch einige nützliche Eigenschaften der Spur.

Hilfssatz 8.2.7 (Eigenschaften der Spur)**.** *Sei $A, B \in M_n(K)$. Dann gilt:*

(i) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

(ii) Die Spur von ähnlichen Matrizen ist gleich, d.h. ist B eine invertierbare Matrix, so gilt $\operatorname{tr}(B^{-1}AB) = \operatorname{tr}(A)$.

(iii) $\operatorname{tr}(A^t) = \operatorname{tr}(A)$.

Beweis. (i) Seien $A = (a_{ij})$ und $B = (b_{ij})$.
Dann folgt $\operatorname{tr}(AB) = \sum_{i,j=1}^{n} a_{ij}b_{ji} = \sum_{i,j=1}^{n} b_{ij}a_{ji} = \operatorname{tr}(BA)$.

(ii) Folgt aus (i), denn $\operatorname{tr}(B^{-1}AB) = \operatorname{tr}(ABB^{-1}) = \operatorname{tr}(A)$.

(iii) Folgt direkt aus der Definition der Spur. □

Wir bemerken, dass auch die Determinante von ähnlichen Matrizen gleich ist. Dies folgt direkt aus dem Determinantenmultiplikationssatz, denn: $\det(B^{-1}AB) = \det(B^{-1})\det(A)\det(B) = \det(A)\det(B)\det(B^{-1}) = \det(A)\det(BB^{-1}) = \det(A)$. Daraus können wir weiter folgern, dass alle Koeffizienten des charakteristischen Polynoms von je zwei ähnlichen Matrizen gleich sind, oder mit anderen Worten, dass das charakteristische Polynom sich bei einem Basiswechsel nicht ändert, nämlich so: $\chi_A(t) = \det(A - tE) = \det(B^{-1}(A - tE)B) = \det(B^{-1}AB - tB^{-1}EB) = \det(B^{-1}AB - tE) = \chi_{B^{-1}AB}(t)$.

Folgerung 8.2.8 (Charakteristisches Polynom einer 2×2-Matrix)**.** *Sei* $A \in M_2(K)$. *Dann gilt:*

$$\chi_A(t) = t^2 - \operatorname{tr}(A)t + \det(A).$$

Insbesondere folgt: Hat A zwei verschiedene Eigenwerte λ_1 und λ_2, dann gilt $\operatorname{tr}(A) = \lambda_1 + \lambda_2$ *und* $\det(A) = \lambda_1\lambda_2$.

Beweis. Folgt aus Satz 8.2.6 und $(t - \lambda_1)(t - \lambda_2) = t^2 - (\lambda_1 + \lambda_2)t + \lambda_1\lambda_2$. □

Beispiel 8.2.9. Wir betrachten die Matrix

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

Wegen $\mathrm{tr}(A) = 6$ und $\det(A) = 8$ gilt somit nach Satz 8.2.6, dass $\chi_A(t) = t^2 - 6t + 8$. Die beiden Nullstellen von χ_A, und somit die Eigenwerte von A, sind $\lambda_1 = 2$ und $\lambda_2 = 4$. Der Eigenraum zum Eigenwert 2 bestimmt sich zu

$$\ker \begin{pmatrix} 3-2 & 1 \\ 1 & 3-2 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \mathbb{R} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Der Eigenraum zum Eigenwert 4 bestimmt sich zu

$$\ker \begin{pmatrix} 3-4 & 1 \\ 1 & 3-4 \end{pmatrix} = \ker \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Matrix A ist also zu einer Diagonalmatrix mit den beiden Diagoanlelementen 2 und 4 ähnlich In der Tat gilt, wenn wir die Transformationsmatrix

$$T = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

verwenden, deren Spalten eine Basis aus Eigenvektoren von A bilden:

$$T^{-1}AT = \frac{1}{-2} \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

8.3 Diagonalisierbarkeit

Wir erinnern an den Begriff der Ähnlichkeit von Matrizen: zwei quadratische Matrizen A und B heißen *ähnlich*, wenn es eine invertierbare Matrix T gibt, so dass $B = T^{-1}AT$ gilt.

Definition 8.3.1. Eine quadratische Matrix heißt *diagonalisierbar*, wenn sie zu einer Diagonalmatrix ähnlich ist.

Satz 8.3.2. *Sei $A \in M_n(K)$. Dann sind äquivalent:*

(i) A ist diagonalisierbar.

(ii) Es gibt eine Basis von K^n aus Eigenvektoren von A.

(iii) Die Summe der Dimensionen aller Eigenräume von A ist n.

Beweis:. Ist A diagonalisierbar, dann gibt es $T \in M_n(K)$, so dass $T^{-1}AT$ eine Diagonalmatrix ist. Dann folgt für jedes $i \in \{1, \dots, n\}$, dass für ein $\lambda \in K$ gilt: $(T^{-1}AT)e_i = \lambda e_i$, was zu $A(Te_i) = \lambda(Te_i)$ ist, d.h. die Te_i sind Eigenvektoren von A Da T invertierbar ist, bilden somit $Te_1, \dots, Te_n$ eine Basis aus Eigenvektoren von A.

Umgekehrt, gibt es eine Basis von K^n aus Eigenvektoren von A, so hat die darstellende Matrix bezüglich dieser Basis Diagonalgestalt, denn jeder dieser Vektoren von A auf ein Vielfaches von sich abgebildet.

Gibt es eine Basis $v_1, \dots, v_n$ aus Eigenvektoren von A mit $Av_i = \lambda_i v$, so sind insbesondere die zu einem Eigenwert gehörigen Eigenvektoren linear unabhängig. Sie bilden auch ein Erzeugendensystem dieses Eigenraums: Denn sei $v \in E_\lambda$. Dann hat v eine eindeutig bestimmte Darstellung $v = \sum_{i=1}^n x_i v_i$. Es gilt $0 = Av - \lambda v = \sum_{i=1}^n (\lambda_i - \lambda) x_i v_i$. Wegen der linearen Unabhängigkeit der v_i folgt dann, dass $x_i = 0$ falls $\lambda_i \neq \lambda$, was zeigt, dass sich v alleine mit den Eigenvektoren zum Eigenwert λ als Linearkombination darstellen lässt. Diese bilden somit ein Erzeugendensystem und also auch eine Basis diese Eigenraums, was (iii) zeigt.

Schließlich, falls (iii) gilt, dann gibt es zu den Eigenwerten $\lambda_1, \dots, \lambda_k$ von A Basisvektoren $v_{ij} \in E_{\lambda_i}, j = 1, \dots, \dim(E_{\lambda_i})$, wobei $\dim(E_{\lambda_1}) + \dots + \dim(E_{\lambda_k}) = n$. Wir zeigen, dass die n Vektoren v_{ij} linear unabhängig sind: Sei $\sum x_{ij} v_{ij} = 0$ und sei $m \in \{1, \dots, k\}$. Betrachte die Abbildung, gegeben durch die Matrix $(A - \lambda_1 E) \dots (A - \lambda_{m-1} E)(A - \lambda_{m+1} E) \dots (A - \lambda_k E)$. Wird sie auf $0 = \sum x_{ij} v_{ij}$ angewendet, dann ergibt dies

$$0 = (\lambda_m - \lambda_1) \dots (\lambda_m - \lambda_{m-1})(\lambda_m - \lambda_{m+1}) \dots (\lambda_m - \lambda_k) \sum_{j=1}^{\dim(E_{\lambda_m})} x_{mj} v_{mj},$$

was zeigt, dass die x_{mj} alle Null sind, da die Vektoren v_{mj} linear unabhängig sind. Da dies für alle $m \in \{1, \dots, k\}$ gilt, folgt, dass alle Koeffizientenx_{ij} Null sind, womit wir die lineare Unabhängigkeit der n Vektoren v_{ij} gezeigt haben, diese bilden somit eine Basis des K^n aus Eigenvektoren von A. $\square$

Hilfssatz 8.3.3. *Sei A eine $n \times n$-Matrix. Seien $\lambda_1, \dots \lambda_k$ paarweise verschiedene Eigenwerte von A und $v_1, \dots, v_k$ dazu gehörende Eigenvektoren. Dann sind die Vektoren $v_1, \dots, v_k$ linear unabhängig.*

Beweis:. Wir beweisen die Aussage durch Induktion über k. Die Aussage ist richtig für $k = 1$, denn ein Eigenvektor ist ungleich Null nach Definition.

Seien nun $v_1, \dots, v_k$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ von A. Nach Induktionsannahme sind die $v_1, \dots, v_{k-1}$ linear unabhängig. Angenommen, die $v_1, \dots, v_k$ wären linear abhängig. Dann gibt es eine Linearkombination $x_1 v_1 + \cdots + x_k v_k = 0$, wobei die $x_i \in K$ Koeffizienten sind, die nicht alle gleich Null sind. Insbesondere ist hierbei $x_k \neq 0$, denn sonst wären die $v_1, \dots, v_{k-1}$ bereits linear abhängig gewesen. Damit gilt

$$v_k = -\frac{x_1}{x_k} v_1 - \dots - \frac{x_{k-1}}{x_k} v_{k-1} = w_1 + \dots + w_{k-1},$$

wobei wir $w_j := -\frac{x_j}{x_k} v_j$ für $j = 1, \dots, k-1$ gesetzt haben. Insbesondere sind die w_j nicht alle gleich Null, da $v_k \neq 0$. Anwenden von A auf beide Seiten der obigen Gleichung ergibt nun, da auch w_j als Vielfaches von v_j entweder wieder ein Eigenvektor zum Eigenwert λ_j oder gleich Null ist,

$$\lambda_k (w_1 + \dots + w_{k-1}) = \lambda_1 w_1 + \cdots + \lambda_{k-1} w_{k-1},$$

Daraus erhalten wir, dass

$$0 = (\lambda_1 - \lambda_k) w_1 + \cdots + (\lambda_{k-1} - \lambda_k) w_{k-1},$$

was eine nichtriviale Linearkombination ist, da die Eigenwerte λ_i alle paarweise verschieden sind und die w_j nicht alle gleich Null sind. Dies steht im Widerspruch zur Induktionsannahme, nach der die $v_1, \dots, v_{k-1}$ linear unabhängig sind.

Wir haben nun gezeigt, dass die $v_1, \dots, v_n$ eine Basis von K^n aus Eigenvektoren von A bilden. Nach Satz 8.3.2 ist A damit diagonalisierbar. $\square$

Daraus erhalten wir das folgende nützliche Kriterium.

Folgerung 8.3.4. *Hat eine $n \times n$-Matrix n paarweise verschiedene Eigenwerte, dann ist sie diagonalisierbar.*

Beweis:. Nach Hilfssatz 8.3.3 gibt es n linear unabhängige Vektoren in K^n, die alle Eigenvektoren von A sind. Nach Satz 8.3.2 ist A damit diagonalisierbar. □

Beispiele 8.3.5. (i) Sei

$$A = \begin{pmatrix} 1 & 4 & 6 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix} \in M_3(\mathbb{R}).$$

Dann ist A diagonalisierbar, dann man kann direkt ablesen, dass A die drei Eigenwerte $1, 2, 3$ hat.

(ii) Sei

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \in M_2(\mathbb{R}).$$

Dann ist A *nicht* diagonalisierbar, dann, wie man leicht sieht, hat A den einzigen Eigenwert λ, aber der Kern von

$$A - \lambda E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{R}).$$

ist nur eindimensional. Also kann es keine Basis von $\mathbb{R}^2$ aus Eigenwerten von A geben.

Der folgende Satz, den wir erst im nächsten Abschnitt beweisen, wird manchmal auch als "Spektralsatz" bezeichnet.

Satz 8.3.6. *Sei $A \in M_n(\mathbb{R})$ eine* symmetrische *reelle Matrix, d.h. es gilt $A^t = A$. Dann ist A diagonalisierbar.*

Beispiele 8.3.7. (i) Sei

$$A = \begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{pmatrix} \in M_3(\mathbb{R}).$$

Dann ist A diagonalisierbar, denn A ist symmetrisch.

(ii) Sei

$$A = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix} \in M_2(\mathbb{R}).$$

Dann ist A zwar symmetrisch, d.h. $A^t = A$, aber *nicht* diagonalisierbar, denn wegen $\det(A) = 0$ und $\mathrm{tr}(A) = 0$ folgt sofort, dass A nur den Eigenwert 0 hat. Die einzige Matrix, die aber zur Nullmatrix ähnlich ist, ist die Nullmatrix selbst.

Dies ist kein Widerspruch zu dem obigen Satz, da dieser nur für reelle Matrizen formuliert ist.

8.4 Euklidische Vektorräume und Hauptachsentransformation

In diesem Abschnitt wollen wir den Spektralsatz für reelle symmetrische Matrizen und eine geometrische Anwendung betrachten.

Dazu benötigen wir noch ein paar Vorbereitungen. Wir führen zunächst den Begriff des Skalarprodukts ein. Ein reeller Vektorraum V heißt *euklidischer Vektorraum*, falls es ein *Skalarprodukt* auf V gibt, d.h. eine bilineare Abbildung $\beta\colon V \times V \to \mathbb{R}$, die symmetrisch und positiv definit ist, d.h. es soll gelten $\beta(v, w) = \beta(w, v)$ und $\beta(v, v) \geq 0$ mit Gleichheit genau dann wenn $v = 0$. Wir schreiben für ein solches Skalarprodukt auch $\langle v, w\rangle$ statt $\beta(v, w)$.

Hat man ein solches Skalarprodukt, kann man geometrische Größen definieren, z.B. die *Länge* eines Vektors v als $\|v\| := \sqrt{\langle v, v\rangle}$ oder den *Winkel* zwischen zwei

Vektoren, indem man den Kosinus des Winkels zwischen zwei Vektoren der Länge eins als $\langle v, w\rangle$ setzt, bzw., allgemeiner für Vektoren $v, w \in V$ ungleich Null:

$$\angle(v, w) := \arccos\left(\frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}\right).$$

Wir sagen, dass zwei Vektoren aufeinander *senkrecht* stehen oder *orthogonal* zueinander sind, falls ihr Skalarprodukt Null ist (wegen $\cos(\frac{\pi}{2}) = 0$). Insbesondere steht der Nullvektor auf allen Vektoren senkrecht.

Ein Beispiel ist das Standardskalarprodukt auf $\mathbb{R}^n$, gegeben durch

$$\langle x, y\rangle = x_1 y_1 + \cdots + x_n y_n = x^t y.$$

Im Folgenden wollen wir stets annehmen, dass $\mathbb{R}^n$ mit dem Standardskalarprodukt versehen ist, bzw. wir betrachten im Rest des Abschnitts nur noch den Fall von $\mathbb{R}^n$, versehen mit dem Standardskalarprodukt $\langle \cdot, \cdot \rangle$.

Eine Basis $b_1, \ldots, b_n$ des $\mathbb{R}^n$ heißt *Orthonormalbasis*, falls jeder Vektor Einheitslänge hat und die Vektoren gegenseitig aufeinander senkrecht stehen. Dies kann man auch so ausdrücken:

$$\langle b_i, b_j\rangle = \begin{cases} 1, & \text{falls } i = j; \\ 0, & \text{sonst.} \end{cases}$$

Schreibt man die Vektoren einer Orthonormalbasis als Spalten in eine Matrix B, so gilt $B^t B = E$. Man nennt reelle quadratische Matrizen $B \in M_n(\mathbb{R})$ *orthogonal*, wenn $B^t B = E$ ist, was dazu äquivalent ist, dass die Spalten eine Orthonormalbasis bilden (was dazu äquivalent ist, dass die Zeilen der Matrix eine Orthonormalbasis bilden). Die Menge der orthogonalen reellen $n \times n$-Matrizen wird mit $\mathrm{O}(n)$ bezeichnet.

Nun können wir noch eine genauere Version des Spektralsatzes formulieren:

Satz 8.4.1. *Sei $A \in M_n(\mathbb{R})$ eine* symmetrische *reelle Matrix, d.h. es gilt $A^t = A$. Dann ist A* orthogonal diagonalisierbar*, d.h. es gibt eine orthogonale Matrix $B \in \mathrm{O}(n)$, so dass $B^{-1}AB$ eine reelle Diagonalmatrix ist, oder, mit anderen Worten, A besitzt eine Orthonormalbasis aus Eigenvektoren.*

Beweisskizze:. Da man A auch als komplexe Matrix auffassen kann, hat A jedenfalls einen komplexen Eigenwert λ, es gibt also ein $x \in \mathbb{C}^n$ ungleich Null mit $Ax = \lambda x$. Versehen wir beide Seiten dieser Gleichung mit komplexen Konjuagtionsstrichen, so erhalten wir $\bar{A}\bar{x} = \bar{\lambda}\bar{x}$, wegen $\bar{A} = A$ bedeutet dies: Der (komponentenweise) konjugierte Vektor $\bar{x}$ ist Eigenvektor von A zum Eigenwert $\bar{\lambda}$. Durch Transponieren und wegen $A^t = A$ erhält man $\bar{x}^t A = \bar{\lambda}\bar{x}^t$. Wir multiplizieren diese Gleichung von rechts mit dem Vektor x und erhalten $\bar{x}^t Ax = \bar{\lambda}\bar{x}^t x$. Daraus erhalten wir $\lambda\bar{x}^t x = \bar{\lambda}\bar{x}^t x$, woraus wegen $\bar{x}^t x > 0$ folgt dass $\lambda = \bar{\lambda}$. Der Eigenwert λ ist also reell.

Nun hat man einen Eigenvektorvektor x gefunden und man kann induktiv fortfahren, indem man feststellt, dass wegen $0 = \langle Ax, y\rangle = x^t A^t y = x^t Ay = \langle x, Ay\rangle$ die durch A gegebene Abbildung den Untervektorraum von $\mathbb{R}^n$, der aus den Vektoren senkrecht zu x besteht, auf sich abbildet.

Damit findet man eine Basis von Eigenvektoren, die zueinander orthogonal sind. Indem man diese auf Einheitslänge normiert, bekommt man die gewünschte Orthonormalbasis aus Eigenvektoren. □

Wegen $B^t B = E$ für eine orthogonale Matrix B folgt, dass orthogonale Matrizen invertierbar sind und für sie $B^{-1} = B^t$ gilt. Nun wollen wir eine Bilinearform $Q\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ betrachten.Wie führt man an einer Bilinearform einen Basiswechsel durch?

Bezüglich einer Basis $B = (b_1, \dots, b_n)$ des $\mathbb{R}^n$ ist Q durch die Werte auf den Basisvektoren gegeben, wir definieren: $Q(b_i, b_j) =: a_{ij}$ und nennen die so enstehende Matrix $A = (a_{ij})$ die ***Strukturmatrix*** von Q bezüglich der Basis $B = (b_1, \dots, b_n)$. Damit gilt für beliebige Vektoren $x, y \in \mathbb{R}^n$, dass

$$Q(x, y) = \Phi_B(x)^t A \Phi_B(y),$$

wenn $\Phi_B(x), \Phi_B(y) \in \mathbb{R}^n$ die Koordinatendarstellungen der Vektoren x, y bezüglich B sind. Ist T die Transformationsmatrix für den Übergang einer weiteren Basis B', so ist die Strukturmatrix von A bezüglich B' durch $T^t AT$ gegeben.

Ist unsere Bilinearform *symmetrisch* sein, d.h. gilt $Q(v,w) = Q(w,v)$, so ist auch ihre Strukturmatrix A symmetrisch und dann können wir nach dem obigen Satz einen orthogonalen Basiswechsel an A (aufgefasst als Endomorphismus $\mathbb{R}^n \to \mathbb{R}^n$) durchführen, wegen $B^{-1} = B^t$ bei einer orthogonalen Matrix B entspricht dies jedoch genau dem Basiswechsel der Strukturmatrix einer Bilinearform!

Sei nun q eine *quadratische Form* auf $\mathbb{R}^n$, d.h. ein Ausdruck der Form

$$q(x) = \sum_{i=1}^{n} a_i x_i^2 + \sum_{i<j}^{n} b_{ij} x_i x_j. \tag{8.4.1}$$

Dann definieren wir die *assoziierte symmetrische Bilinearform* durch ihre Strukturmatrix

$$A := \begin{pmatrix} a_1 & \frac{1}{2}b_{12} & \dots & \frac{1}{2}b_{1,n-1} & \frac{1}{2}b_{1n} \\ \frac{1}{2}b_{12} & a_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{1}{2}b_{1,n-1} & & \ddots & a_{n-1} & \frac{1}{2}b_{n-1,n} \\ \frac{1}{2}b_{1n} & \dots & \dots & \frac{1}{2}b_{n-1,n} & a_n \end{pmatrix}$$

Damit gilt $q(x) = x^t A x$.

Wir betrachten nun sogenannte quadratische Hyperflächen (auch Quadriken genannt) studieren, d.h. Gebilde der Art

$$\{x \in \mathbb{R}^n \mid q(x) = 1\},$$

wobei q wie in (8.4.1) ist.

Dazu bestimmen wir die Eigenwerte der symmetrischen Matrix A und diagonalisieren diese durch einen orthogonalen Basiswechsel.

Beispiel 8.4.2. Sei $n = 2$ und $q(x) = 3x_1^2 + 2x_1x_2 + 3x_2^2$. Wir betrachten die Quadrik $\{x \in \mathbb{R}^n \mid q(x) = 1\}$, die aus allen Punkten $(x_1, x_2) \in \mathbb{R}^2$ besteht mit

$$3x_1^2 + 2x_1x_2 + 3x_2^2 = 1.$$

Wir wollen die *Hauptachsentransformation* durchführen, d.h. wir wollen die Strukturmatrix der assozierten symmetrischen Bilinearform diagonalisieren. Diese Matrix ist

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

Die Eigenwerte davon haben wir bereits in Beispiel 8.2.9 bestimmt. Die dort gefundene Basis aus Eigenvektoren besteht in der Tat aus zwei aufeinander senkrecht stehenden Vektoren, diese lassen sich auf Einheitslänge skalieren. Somit ist

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Eine Orthogonalbasis aus Eigenvektoren. Stellt man die Matrix A bezüglich dieser Basis dar, wird sie nach Beispiel 8.2.9 zu

$$\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Die Quadrik nimmt also nach dem Basiswechsel die Form

$$2x_1^2 + 4x_2^2 = 1$$

an und man kann daran sehen, dass es sich um eine Ellipse handelt.

Allgemeiner kann so eine Quadrik, wenn sie *nicht ausgeartet* ist, d.h. wenn Null kein Eigenwert von A ist, folgende Formen annehmen: $\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$, wobei es sich um eine *Ellipse* handelt falls, $\lambda_1, \lambda_2 > 0$, um eine *Hyperbel*, falls $\lambda_1 \lambda_2 < 0$ und um eine *leere Quadrik*, falls $\lambda_1, \lambda_2 < 0$.

8.5 Jordansche Normalform und Klassifikation von Endomorphismen

In diesem Abschnitt betrachten wir ausschließlich komplexe Vektorräume endlicher Dimension, oder, mit anderen Worten, linear Endomorphismen von $\mathbb{C}^n$.

Diese kann man, auch wenn sie im Allgemeinen nicht diagonalisierbar sind, auf eine gewisse Normalform bringen, die sogenannte Jordansche Normalform. Da zwei qquadratische komplexe Matrizen genau dann die gleiche Normalform haben, wenn Sie ähnlich sind, kann man sich damit einen Gesamtüberblick über alle linearen Endomorphismen eines endlich-dimensionalen Vektorraums verschaffen, eine "Klassifikation" der Endomorphismen erreichen, also eine Einteilung in Ähnlichkeitsklassen.

Definition 8.5.1. Ein *Jordanblock*, *Jordankästchen* oder eine *elementare Jordanmatrix* ist eine $m \times m$-Matrix der Form

$$J_{m,\lambda} = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix},$$

d.h. eine $m \times m$-Matrix, wo auf der Diagonalen alle Einträge gleich λ sind, alle Einträge, deren Spaltenindex um eins größer als der Zeilenindex ist, gleich 1 und die restlichen Einträge gleich 0.

Beispiele 8.5.2.

$$J_{1,\lambda} = (\lambda), \qquad J_{2,\lambda} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad J_{3,\lambda} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Direkt an der Form der elementaren Jordanmatrix $J_{m,\lambda}$ kann man ablesen: Der einziege Eigenwert ist λ und die Matrix ist nicht diagonaliserbar, falls $m \geq 2$.

Definition 8.5.3. Eine komplexe quadratische Matrix heißt *Jordansche Normalform*, falls sie eine Blockdiagonalmatrix ist, bei der die Kästchen auf der Diagonalen elementare Jordanmatrizen (nicht notwendig mit verschiedenen λ) sind.

Satz 8.5.4 (Hauptsatz über die Jordansche Normalform)**.** *Jede komplexe quadratische Matrix ist ähnlich zu einer Jordanschen Normalform.*

Zwei Jordansche Normalformen sind genau dann zueinander ähnlich, wenn die Kästchen auf der Diagonalen bis auf die Reihenfolge gleich sind.

Aufgrund des zweiten Teils des Satzes kann man von der zu einer quadratischen komplexen Matrix gehörenden Jordanform sprechen, diese ist bis auf die Reihenfolge der Kästchen auf der Diagonalen eindeutig bestimmt.

Wir beweisen den Satz nicht, aber wir wollen uns ein paar Anwendungen überlegen.

Beispiele 8.5.5. Wir wollen uns in kleinen Dimensionen die möglichen Jordanschen Normalformen überlegen (bis auf Reihenfolge der Kästchen).

(i) $n = 2$. Sei $A \in M_2(\mathbb{C})$. Falls A zwei verschiedene Eigenwerte λ und μ hat, ist A diagonalisierbar und hat die Jordansche Normalform

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Hat A nur einen Eigenwert λ so ist A entweder gleich

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

oder ähnlich zu

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

(ii) Bei 3×3-Matrizen treten folgende Jordansche Normalformen auf:

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix}, \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix},$$
$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Wie kann man diese Formen unterscheiden? Es gilt, dass die Matrix $J_{m,\lambda} - \lambda E = J_{m,0}$ *nilpotent* ist, d.h. $J_{m,\lambda}^m = 0$. Genauer gilt: $\dim\ker(J_{m,0}^k) = k$ für $k =$

$1, \ldots, m$. Man kann deswegen die Anzahlen und Größen der Jordanblöcke der zu einer quadratischen komplexen Matrix A gehörenden Jordanschen Normalform zu einem Eigenwert λ an der Folge der Zahlen

$$\dim(\ker(A - \lambda E)^k))$$

ablesen.

Eine Anwendung der Jordanschen Normalform ist die Berechnung der Potenzen von quadratischen komplexen Matrizen A. Dazu bringt man A auf Jordansche Normalform $T^{-1}AT = J$. Wegen $(T^{-1}AT)^k = T^{-1}A^kT$, genügt es, die Potenzen einer Jordanschen Normalform zu bestimmen. Und wegen der Blockdiagonalgestalt der Jordanschen Normalform genügt es, ein Jordankästchen zu betrachten. Sei also $J = J_{m,\lambda}$ und sei $N = J_{m,0}$. Dann ist $J = \lambda E + N$. Hier gilt, dass N und λE kommutieren. Deswegen kann man hier sogar die binomische Formel anwenden und erhält $J^k = \sum_{i=0}^{k} \binom{k}{i} \lambda^i N^{k-i}$, wobei man die Konvention verwendet, dass $A^0 = E$ für eine quadratische Matrix A gelten soll.

Beispiel 8.5.6. Sei

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

Wir wollen A^3. Sei

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

Dann gilt

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

und $N^k = 0$ für $k \geq 3$. Zum Beispiel für A^7 erhalten wir:

$$A^7 = \lambda^7 E + \binom{7}{1}\lambda^6 N + \binom{7}{2}\lambda^5 N^2 = \begin{pmatrix} \lambda^7 & 7\lambda^6 & 21\lambda^5 \\ 0 & \lambda^7 & 7\lambda^6 \\ 0 & 0 & \lambda^7 \end{pmatrix}.$$