

Lineare Algebra und Analytische Geometrie

Christoph Helmberg

Skriptum zu den Vorlesungen
Wintersemester 2014/15
Sommersemester 2015

Fakultät für Mathematik
Technische Universität Chemnitz

Inhaltsverzeichnis

1 Grundlagen	1
1.1 Logik	1
1.2 Mengen	4
1.3 Abbildungen (Funktionen)	7
1.4 Mengen-Familien und Kartesische Produkte	11
1.5 Relationen	15
2 Algebraische Grundbegriffe	21
2.1 Halbgruppen, Monoide und Gruppen	21
2.2 Homomorphismen	30
2.3 Ringe	36
2.4 Körper	40
2.5 Polynome	46
3 Vektorräume	50
3.1 Vektorraum	50
3.2 Lineare Hülle, Linearkombinationen und lineare Unabhängigkeit	54
3.3 Basis und Dimension	59
3.4 Summen von Vektorräumen	62
3.5 Affine Unterräume, Quotientenvektorräume	64
4 Lineare Abbildungen	68
4.1 Definition und erste Eigenschaften	68
4.2 Bild, Fasern und Kern	73
4.3 Der Dualraum $V^* = \text{Hom}_K(V, K)$	79

5	Lineare Abbildungen und Matrizen	86
5.1	Koordinatensysteme und Matrizenrechnung	86
5.2	Koordinatentransformationen	93
5.3	Matrizenumformungen und Elementarmatrizen	98
5.4	Lineare Gleichungssysteme	101
6	Die Determinante	105
6.1	Definition und erste Eigenschaften	105
6.2	Minoren und der Entwicklungssatz von Laplace	111
6.3	Geometrische Interpretation im \mathbb{R}^n	114
7	Polynomdivision, Primfaktorzerlegung	117
8	Endomorphismen von Vektorräumen	123
8.1	Invariante Untervektorräume, Eigenwerte und Eigenvektoren	123
8.2	Das charakteristische Polynom	127
8.3	Das Minimalpolynom	130
8.4	F -zyklische Untervektorräume	133
8.5	Normalformen	138
9	Räume mit Skalarprodukt	143
9.1	Euklidische und Unitäre Räume	143
9.2	Orthogonalität und Orthogonalsysteme	147
9.3	Biorthogonale Basen und Adjungierte Abbildungen	151
9.4	Orthogonale und Unitäre Abbildungen	154
9.5	Normale Endomorphismen	160
9.6	Selbstadjungierte Endomorphismen	161
9.7	Positiv semidefinite Endomorphismen	164
10	Bilinear- und quadratische Formen	171
10.1	Bilinearformen	171
10.2	Symmetrische/Hermitesche Bilinearformen	173
10.3	Quadriken (im \mathbb{R}^n)	176

Kapitel 1

Grundlagen

1.1 Logik

„Schule des korrekten, klaren und folgerichtigen Denkens“

Eine **Aussage** ist ein Satz der Umgangssprache, der entweder **wahr** (W) oder **falsch** (F) ist.

Beispiele:

2010 hatte Chemnitz etwa 240000 Einwohner.	W
15 ist eine Primzahl.	F
13 ist eine schöne Zahl.	Keine Aussage (solange schön undefiniert ist).
Dieser Satz ist falsch.	Keine Aussage.

♡

Es gibt Aussagen, für die noch nicht bekannt ist, ob sie W oder F sind.

Beispiel: Es gibt unendlich viele Primzahlzwillinge (Paare $(p, p + 2)$ mit p und $p + 2$ Primzahl). ♡

Eine **Aussageform** erhält man, indem Objekte einer Aussage durch Variable/Parameter ersetzt werden. Eine Aussageform wird erst durch Einsetzen von Werten zu einer Aussage.

Beispiel:

x teilt y , $x|y$ ist eine Aussageform $A(x, y)$, deren Wahrheit von x und y abhängt.

$a^2 + b^2 = c^2$ ist eine Aussageform $B(a, b, c)$. ♡

Wir schreiben nun nur noch A, B, C für Aussagen und Aussageformen (und lassen meistens die Parameter weg)

Verknüpfungen von Aussagen (und Aussageformen)

<u>Negation</u>	$\neg A$	„nicht A “, „non A “,
<u>Konjunktion</u>	$A \wedge B$	„ A und B “, [beide müssen wahr sein]
<u>Disjunktion</u>	$A \vee B$	„ A oder B “, [mindestens eine muss wahr sein]
<u>Konditional</u>	$A \rightarrow B$	„wenn A , dann B “, „aus A folgt B “, „ A impliziert B “,
<u>Bikonditional</u>	$A \leftrightarrow B$	„ A genau dann, wenn B “, „ A äquivalent zu B “.

\neg und \wedge binden stärker als \vee ,
diese stärker als \rightarrow ,
diese stärker als \leftrightarrow ,
aber besser ist es zu klammern.

Die Wahrheit der Verknüpfung ergibt sich aus folgender Wahrheitstabelle:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$	$A \wedge \neg A$
W	W	F	W	W	W	W	F
W	F	F	F	W	F	F	F
F	W	W	F	W	W	F	F
F	F	W	F	F	W	W	F

Eine Aussage „aus A folgt B “ ($A \rightarrow B$) kann sehr wohl falsch sein.

Beachte aber, dass $A \rightarrow B$ immer wahr ist, wenn A falsch ist; jeder Schluss aus einer falschen Voraussetzung ist richtig!

Ab und zu schreibt man statt $A \rightarrow B$ auch $B \leftarrow A$, aber man darf es nicht mit $B \rightarrow A$ verwechseln.

Eine Aussage, die unabhängig vom Wahrheitswert der Teilaussagen immer wahr ist, heißt Tautologie oder aussagenlogisch wahr.

Beispiel: $\neg(A \wedge \neg A)$, das „Gesetz vom ausgeschlossenen Widerspruch“. ♡

Ü: Zeige folgende Tautologien:

- $(A \rightarrow B) \leftrightarrow (\neg A \vee B)$
- $\neg(A \wedge B) \leftrightarrow (\neg A) \vee (\neg B)$
- $\neg(A \vee B) \leftrightarrow (\neg A) \wedge (\neg B)$
- $(A \leftrightarrow B) \leftrightarrow ((A \rightarrow B) \wedge (B \rightarrow A))$
- $(A \leftrightarrow B) \leftrightarrow (\neg A \leftrightarrow \neg B)$
- $(A \wedge B) \leftrightarrow \neg((\neg A) \vee (\neg B))$

Beachte: Eigentlich genügen \neg und \vee (oder \neg und \wedge), um alle logischen Verknüpfungen formulieren zu können. ☒

In Beweisen verwendet man auch gerne folgende Tautologien als Schlusstechniken.

Ü: Zeige folgende Tautologien:

- $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$ Umkehrschluss
- $(\neg A \rightarrow F) \leftrightarrow A$ Indirekter Beweis
- $(A \wedge (A \rightarrow B)) \rightarrow B$ „Modus (ponendo) ponens“
- $((A \rightarrow B) \wedge \neg B) \rightarrow \neg A$ „Modus (tollendo) tollens“

☒

Verknüpft man zwei Aussageformen mit Parametern, entsteht daraus eine neue Aussageform, deren Parameter sich aus den Parameterbezeichnungen dieser zwei Aussageformen ergeben.

Beispiel: Seien $A(x, y), B(s, t)$ zwei Aussageformen. Durch $A(c, a) \wedge B(a, b)$ entsteht eine Aussageform $C(a, b, c)$ und durch $A(h, h) \vee B(h, h)$ eine Aussageform $D(h)$. \heartsuit

Quantoren

Will man Aussagen über alle möglichen Einsetzungen eines gegebenen „Diskursuniversums“ treffen, verwendet man dafür den Allquantor, in Zeichen „ \forall “, „für alle“, etwa für eine Aussageform $A(x)$

$$\forall x A(x)$$

Beispiel: Für $A(x, y) = „x \text{ hat die Farbe } y“$ mit x einem Haus und y einer Farbe wäre $\forall x A(x, \text{schwarz})$ die Aussage „Alle Häuser sind schwarz“. \heartsuit

Soll es mindestens eine Einsetzung geben, für die die Aussageform wahr wird, verwendet man den Existenzquantor, in Zeichen „ \exists “, „es existiert“, „es gibt“,

$$\exists x A(x, \text{schwarz}) \quad \text{wäre „Es gibt ein Haus, das schwarz ist“}.$$

$\forall x$ entspricht der Konjunktion über alle x , daher findet man auch

$$\bigwedge_x \quad \text{statt} \quad \forall x$$

Gibt es gar kein geeignetes x , für das die Aussageform zu überprüfen wäre (etwa in „Alle Atome im Vakuum sind leichter als Wasserstoff“), ist es vorteilhaft die Gesamtaussage als wahr gelten zu lassen, denn wahr ist neutral bezüglich der Konjunktion (wenn man eine Aussage mit W „und“-verknüpft, bleibt ihr Wahrheitswert gleich), man kann sie dann getrost über und-Verknüpfungen mit weiteren Aussagen verbinden.

$\exists x$ entspricht der Disjunktion über alle x , daher findet man auch

$$\bigvee_x \quad \text{statt} \quad \exists x$$

Gibt es gar kein geeignetes x , für das die Aussageform zu überprüfen wäre, kann es auch nicht mindestens eines geben, und die Gesamtaussage ist falsch. Vorteilhafterweise ist falsch neutral bezüglich der Disjunktion (wenn man eine Aussage mit F „oder“-verknüpft, bleibt ihr Wahrheitswert gleich), man kann sie also problemlos mit weiteren oder-Verknüpfungen verbinden.

Vorsicht! Die Negation führt zur Umkehr der Quantoren!

$$\neg(\forall x A(x)) \quad \text{ist äquivalent zu} \quad \exists x \neg A(x)$$

$$\neg(\exists x A(x)) \quad \text{ist äquivalent zu} \quad \forall x \neg A(x)$$

Es ist eine gute Übung, die entsprechenden Aussagen für $A(x, \text{schwarz})$ sprachlich zu formulieren. ∇

An den Anfang setzt man **Axiome**, das sind undefinierte Begriffe und Aussagen über diesen Begriffen, die als wahr vorausgesetzt werden.

Die **Theorie** zu diesen Axiomen besteht dann aus allen aussagenlogisch wahren Aussagen, die durch logische Schlussregeln daraus gewonnen werden können. (Die Mathematik besteht also „nur“ aus Tautologien.)

In Beweisen wird zur besseren Sicht- und Erkennbarkeit statt \rightarrow meist \Rightarrow und statt \leftrightarrow meist \Leftrightarrow geschrieben (die einfachen Pfeilvarianten kommen zu oft für andere Zwecke zum Einsatz). Bei jeder Aussage muss natürlich nachgewiesen werden, dass sie innerhalb des Axiomensystems eine Tautologie, also immer wahr ist.

Gilt $A \Rightarrow B$, sagt man

A ist **hinreichend** für B ,

B ist **notwendig** für A (sonst kann A nicht gelten).

Beispiel: Wenn die Taschenlampe brennt, dann ist die Batterie noch gut. Aber: Nur weil die Batterie noch gut ist, muss die Taschenlampe nicht brennen. B ist also i. A. (im Allgemeinen) keineswegs hinreichend für A . \heartsuit

Gilt $A \Leftrightarrow B$, sagt man auch

A ist **notwendig und hinreichend** für B .

Eine **Definition** führt einen Namen für ein Objekt oder eine Schreib-/Sprechweise für eine Aussage ein.

Wird ein Name für ein Objekt eingeführt, schreibt man gerne $:=$

Beispiel: $2 := 1 + 1$ $2 + 1 =: 3$ \heartsuit

Führt man einen Namen für eine Aussage(form) ein, verwendet man meist $:\Leftrightarrow$

Beispiel: $A(x) :\Leftrightarrow$ Das Haus x ist schwarz. \heartsuit

Dabei steht das zu definierende jeweils auf der Seite des Doppelpunkts.

1.2 Mengen

Definition 1.1 (Cantor) *Eine Menge ist eine Zusammenfassung von wohlunterscheidbaren Objekten unseres Denkens oder unserer Anschauung.*

Dass diese Definition nicht ganz ungefährlich ist, zeigt das RUSSELSches Paradoxon

„Die Menge aller Mengen, die sich nicht selbst enthalten“.

Wir nutzen dennoch diese „naive Mengenlehre“ und werden nur mit „ungefährlichen“ Mengen arbeiten.

Die in der Menge enthaltenen Objekte sind die Elemente der Menge.

Für „ a ist ein Element der Menge A “ schreiben wir $a \in A$ (auch $A \ni a$).

Für „ a ist nicht Element der Menge A “ schreiben wir $a \notin A$ (auch $A \not\ni a$).

Die Elemente einer Menge können durch Aufzählung,

$$\mathbb{N} := \{1, 2, 3, \dots\} \quad \text{die } \underline{\text{natürlichen Zahlen}}$$

[besser wäre die Einführung über die PEANO-Axiome, das dauert hier zu lange],

oder durch Beschreibung der Eigenschaften angegeben werden,

$$\mathbb{Z} := \{x : x \in \mathbb{N} \vee -x \in \mathbb{N} \vee x = 0\} \quad \text{die } \underline{\text{ganzen Zahlen}},$$

$$k + \mathbb{N} := \{n + k : n \in \mathbb{N}\} \quad \text{für } k \in \mathbb{N}$$

(manchmal schreibt man „|“ statt „:“, also etwa $\{n + k \mid n \in \mathbb{N}\}$).

Elemente, die mehrmals aufgezählt werden, sind dennoch nur „einmal“ enthalten (s. spätere Definition der Gleichheit von Mengen),

$$\{a, a, b, c, b\} \text{ ist das gleiche wie } \{a, b, c\}$$

Einige wichtige Mengenbezeichnungen:

\emptyset ... die leere Menge, sie enthält keine Elemente,

\mathbb{Q} ... die rationalen Zahlen (s. später),

\mathbb{R} ... die reellen Zahlen (s. Analysis),

\mathbb{C} ... die komplexen Zahlen (s. Analysis und später).

Eine Menge heißt Teilmenge einer Menge B , in Zeichen $A \subseteq B$ (manchmal auch $A \subset B$), falls jedes Element von A auch Element von B ist,

$$A \subseteq B :\Leftrightarrow (\forall e (e \in A \Rightarrow e \in B))$$

Insbesondere gilt für jede Menge A , dass $\emptyset \subseteq A$. Die leere Menge ist also Teilmenge jeder Menge.

Zwei Mengen A und B heißen gleich ($A = B$), falls $A \subseteq B$ und $B \subseteq A$ ist,

$$A = B :\Leftrightarrow ((A \subseteq B) \wedge (B \subseteq A)) \quad (\text{so zeigt man das auch!}).$$

Beispiel: Zeige für die übliche Addition: $\forall z \in \mathbb{Z} \quad \mathbb{Z} = z + \mathbb{Z} := \{z + a : a \in \mathbb{Z}\}$.

Beweis: s. Vorlesung □

Beachte: Ans Ende des Beweises setzt man gerne das \square -Symbol oder schreibt q.e.d. (für „quod erat demonstrandum“ [was zu beweisen war]). \heartsuit

Sind sie nicht gleich, schreibt man $A \neq B := \neg(A = B)$.

A heißt echte Teilmenge von B ($A \subset B$ oder $A \subsetneq B$) falls $(A \subseteq B) \wedge (A \neq B)$.

Beispiel: Für $k \in \mathbb{N}$ ist $k + \mathbb{N} \subset \mathbb{N}$ und $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$. \heartsuit

Die Potenzmenge einer Menge A ist die Menge aller Teilmengen von A,

$$2^A := \{B : B \subseteq A\} \quad (\text{manchmal auch } \mathcal{P}(A)).$$

Beispiel: $2^{\{0,1\}} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$, $2^\emptyset = \{\emptyset\}$.

Was ist nun $2^{2^{\{0,1\}}}$ und wieviele Elemente enthält diese Menge? \heartsuit

Seien A und B Mengen,

$A \setminus B := \{a \in A : a \notin B\}$ ist die Differenzmenge A ohne B,

$A \cap B := \{e : e \in A \wedge e \in B\}$ ist der Schnitt von A und B,

$A \cup B := \{e : e \in A \vee e \in B\}$ ist die Vereinigung von A und B,

$A \Delta B := (A \cup B) \setminus (A \cap B)$ ist die symmetrische Differenz von A und B.

Gilt $A \cap B = \emptyset$, so heißen A und B disjunkt (elementfremd).

Für $B \subseteq A$ heißt $A \setminus B$ das Komplement von B bezüglich A.

Beispiel: $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $(\mathbb{Q} \cap \mathbb{N}_0) \setminus \mathbb{N} = \{0\} = \mathbb{N} \Delta \mathbb{N}_0$. \heartsuit

Rechengesetze:

1. Kommutativität:

$$A \cap B = B \cap A, \quad A \cup B = B \cup A,$$

2. Assoziativität:

$$(A \cap B) \cap C = A \cap (B \cap C), \quad (A \cup B) \cup C = A \cup (B \cup C),$$

3. Distributivität:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

4. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$, $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.

Ü: Zeige die Richtigkeit der Rechengesetze! Als Beispiel erster Teil von 3:

Beweis: s. Vorlesung \square

\times

Für zwei Mengen A und B ist das Kartesische Produkt $A \times B = \{(a, b) : a \in A, b \in B\}$ die Menge der geordneten Paare von Elementen aus A und B.

Zwei geordnete Paare (a, b) und (c, d) sind gleich, wenn $a = c$ und $b = d$ gilt, formal $(a, b) = (c, d) :\Leftrightarrow (a = c) \wedge (b = d)$.

Beachte: I. A. ist $A \times B \neq B \times A$, weil die Ordnung wichtig ist.

1.3 Abbildungen (Funktionen)

Nun müssen die Mengen und ihre Elemente zueinander in Beziehung gebracht werden, damit auch etwas Dynamik entsteht. Abbildungen tun dies in sehr allgemeiner, aber dennoch kontrollierter Form.

Seien A und B Mengen.

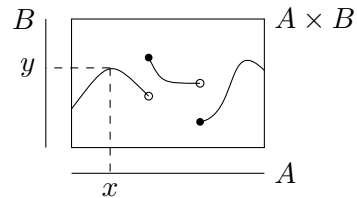
Eine Teilmenge $\Gamma \subseteq A \times B$ heißt Graph einer Abbildung, falls die folgenden beiden Eigenschaften erfüllt sind:

- (Graph a) Zu jedem $x \in A$ gibt es ein $y \in B$ mit $(x, y) \in \Gamma$,
- (Graph b) $(x, y), (x, y') \in \Gamma \Rightarrow y = y'$.

Die zugehörige Abbildung definiert man über

$$f: A \rightarrow B$$

$$x \mapsto y \text{ mit } (x, y) \in \Gamma.$$



Für $(x, y) \in \Gamma$ schreibt man dann $f(x) = y$.

Beispiel: $f: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 1$ ($f(x) = 1$), $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto 3x - y$.

Die Sprache ist *keine* Abbildung von Wörtern in Bedeutungen, z. B. „Chemnitz“. Auch mathematische Begriffe haben nicht immer nur eine Bedeutung („Graph“), aber im jeweiligen Zusammenhang muss die Bedeutung eindeutig sein. ♡

Für eine Abbildung $f: A \rightarrow B, x \mapsto y = f(x)$ nennt man

A Definitionsbereich (von f),

B Wertevorrat/Bildraum (von f),

für $x \in A$ $f(x)$ Bild von x (unter f),

für $A' \subseteq A$ $f(A') := \{f(x) : x \in A'\}$ Bild(menge) von A' ,

für $A' = A$ $f(A)$ Bild von f,

für $B' \subseteq B$ $f^{-1}(B') := \{x \in A : f(x) \in B'\}$ Urbildmenge von B',

für $y \in B$ $f^{-1}(\{y\})$ die Faser über y

oder das Urbild von y .

Vorsicht! Das Urbild ist keine Funktion von B nach A ! Man kann sie aber als Funktion $2^B \rightarrow 2^A$ auffassen. In der Praxis werden für einelementige Mengen dennoch oft die Mengenklammern weggelassen, also z. B. $f^{-1}(y)$ statt $f^{-1}(\{y\})$ geschrieben. Hier soll das sauber unterschieden werden. ▽

Von der Betrachtung aller Abbildungen $A \rightarrow B$ auf einmal hat man nicht viel. Deshalb untersucht man immer nur Teilmengen von Funktionen, für die man gewisse grundlegende Eigenschaften fordert, und versucht daraus

oder aus dem Zusammenspiel solcher Funktionen weitere Eigenschaften und strukturelle Gemeinsamkeiten abzuleiten.

Eine Abbildung $f: A \rightarrow B$ heißt

- surjektiv, falls $f(A) = B$,
[jedes Element aus B wird getroffen]
- injektiv, falls aus $f(x) = f(x')$ folgt, dass $x = x'$,
[jedes Element aus B wird höchstens einmal getroffen]
- bijektiv, falls f surjektiv und injektiv ist.

Beispiele:

- $f: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto 1$ ist weder surjektiv noch injektiv,
 $f: \mathbb{R} \rightarrow \{1\}$, $x \mapsto 1$ ist surjektiv, aber nicht injektiv,
 $f: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto e^x$ (e -Funktion) ist injektiv, aber nicht surjektiv,
 $f: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto -3x$ ist surjektiv und injektiv, also bijektiv,
 $f: \emptyset \rightarrow \emptyset$, \mapsto ist surjektiv und injektiv, also bijektiv.
- Für eine Grundmenge A und eine Teilmenge $S \subseteq A$ ist die charakteristische Funktion von S in A erklärt durch

$$\chi_S: A \rightarrow \{0, 1\}$$

$$a \mapsto \chi_S(a) := \begin{cases} 1 & a \in S, \\ 0 & a \in A \setminus S. \end{cases}$$

Konkret ergibt sich etwa

$$\text{für } A = \{1, 2, 3\}, S = \{2, 3\} \text{ die Funktion } \chi_S = \begin{pmatrix} \chi_S(1) \\ \chi_S(2) \\ \chi_S(3) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

$$\text{für } A = \mathbb{N}, S = 2\mathbb{N} \text{ die Funktion } \chi_S = \begin{pmatrix} \chi_S(1) \\ \chi_S(2) \\ \chi_S(3) \\ \chi_S(4) \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ \vdots \end{pmatrix}.$$

♡

Ist f bijektiv, ist die Urbildmenge für jedes $y \in B$ (die Faser) einelementig,

$$f^{-1}(\{y\}) = \{x\} \quad \text{mit } (x, y) \in \Gamma,$$

wir nutzen dann die gleiche Bezeichnung für die Umkehrabbildung

$$f^{-1}: B \rightarrow A, y \mapsto x \text{ mit } (x, y) \in \Gamma \quad (\text{nur für } f \text{ bijektiv}).$$

Es gilt dann $f(f^{-1}(y)) = y$ und $f^{-1}(f(x)) = x$ für alle $(x, y) \in \Gamma$.

Falls es die Umkehrabbildung gibt, ist sie eindeutig.

Beispiele:

- Die Identität $\text{id}_A: A \rightarrow A$, $x \mapsto x$ ist ihre eigene Umkehrabbildung.

- Seien A, B Mengen. Die Kardinalität/Mächtigkeit von A , in Zeichen $|A|$, dient zum Vergleich der „Anzahl“ der Elemente von Mengen. Ohne dies hier voll zu rechtfertigen, sagen wir:
 - Die Kardinalität von A ist kleiner oder gleich der von B , $|A| \leq |B|$, wenn es eine injektive Abb. $f: A \rightarrow B$ gibt.
 - Gibt es eine bijektive Abb. $f: A \rightarrow B$, sind beide gleichmächtig, $|A| = |B|$.
 - Die Mächtigkeit von A ist echt kleiner, $|A| < |B|$, falls $|A| \leq |B|$ und $|A| \neq |B|$.
- Gibt es insbesondere ein bijektives $f: A \rightarrow B$ für
- $B = \{1, \dots, n\}$ mit $n \in \mathbb{N}_0$, setzt man $|A| = n$ und nennt A endliche Menge oder endlich,
 - $B = \mathbb{N}$, nennt man A abzählbar unendlich,
- Gibt es *keine* surjektive Abb. $f: \mathbb{N} \rightarrow A$, nennt man A überabzählbare Menge (z. B. \mathbb{R} , s. Analysis).

♡

Satz 1.2 Sind A, B endliche Mengen mit $|A| = |B|$, so ist für $f: A \rightarrow B$ äquivalent:

- (a) f ist injektiv,
- (b) f ist surjektiv,
- (c) f ist bijektiv.

Beweis: s. Vorlesung □

Ü: Zeige: Unter den Abbildungen $\mathbb{N} \rightarrow \mathbb{N}$ gibt es sowohl Abbildungen, die injektiv aber nicht surjektiv, als auch Abbildungen, die surjektiv aber nicht injektiv sind (S 1.2 funktioniert also *nur* für endliche Mengen). ✗

Satz 1.3 Die Kardinalität einer Menge A ist echt kleiner als die Kardinalität ihrer Potenzmenge 2^A , in Zeichen $|A| < |2^A|$.

Beweis: s. Vorlesung □

Diese Beweistechnik zur Wahl von M geht auf Cantor zurück und wird Cantorsches Diagonalverfahren genannt.

Komposition (Hintereinanderschaltung) von Abbildungen:

Sind $f: A \rightarrow B$ und $g: B \rightarrow C$ Abbildungen, dann ist

$$g \circ f: A \rightarrow C \quad \text{die Komposition von } g \text{ und } f.$$

$$a \mapsto (g \circ f)(a) = g(f(a))$$

Vorsicht! Bei manchen Autoren ist $(g \circ f) = f(g(a))!$ ▽

Ist $h: C \rightarrow D$ eine weitere Abb., so gilt Assoziativität:

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

denn: für $a \in A$ gilt $((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = (h(g(f(a)))) = h((g \circ f)(a)) = (h \circ (g \circ f))(a)$. \square

Beobachtung 1.4 Für $f: A \rightarrow B$ und $g: B \rightarrow A$ sei $g \circ f = \text{id}_A$, dann ist f injektiv und g surjektiv.

Beweis: s. Vorlesung \square

Ü: Geben Sie ein Beispiel für $A = B, f, g$ mit $g \circ f = \text{id}_A$ aber f nicht surjektiv an. \times

Beobachtung 1.5 Für $f, g: A \rightarrow A$ sei $g \circ f = \text{id}_A$. Ist f bijektiv, dann ist $g = f^{-1}$ die Umkehrabbildung von f . Ist g bijektiv, dann ist $f = g^{-1}$ die Umkehrabbildung von g .

Beweis: s. Vorlesung \square

Ü: Sind $f: A \rightarrow B$ und $g: B \rightarrow C$ injektiv (surjektiv), ist auch $g \circ f$ injektiv (surjektiv). \times

Die Komposition ist also immer assoziativ,
ABER: die Komposition ist i. A. nicht kommutativ!

Beispiel: Schon für $f: A \rightarrow B$ bijektiv ist $f^{-1} \circ f = \text{id}_A \neq f \circ f^{-1} = \text{id}_B$.
 \heartsuit

Die Menge der Permutationen $S_n, n \in \mathbb{N}_0$:

Eine Permutation $\sigma \in S_n$ ist eine bijektive Abbildung

$$\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$$

$$i \quad \mapsto \sigma(i)$$

wobei $\{1, \dots, 0\} := \emptyset$ gesetzt wird und die konkreten Zuweisungen oft über

$$\left(\begin{array}{cccc} 1 & 2 & 3 & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \sigma(n) \end{array} \right) \quad \text{oder nur} \quad \sigma(1) \sigma(2) \sigma(3) \cdots \sigma(n)$$

angegeben werden.

$|S_n| = n! := n \cdot (n-1) \cdots \cdots 1$, „ n Fakultät“, „ n Faktorielle“, $0! := 1$.

Sind $\sigma, \tau \in S_n$, dann ist auch $\tau \circ \sigma \in S_n$ (nach obiger Ü, da beide bijektiv),
ABER i. A. $\tau \circ \sigma \neq \sigma \circ \tau$, etwa für S_3

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Natürlich kann man auch Funktionen auf Funktionenmengen erklären, z. B. das Signum einer Permutation

$$\text{sgn}: S_n \rightarrow \{+1, -1\}$$

$$\sigma \mapsto \text{sgn}(\sigma) = \begin{cases} +1, & \text{falls } \left[\begin{array}{l} \text{die Anzahl der Paare } (i, j) \\ \text{mit } i > j \text{ aber } \sigma(i) < \sigma(j) \end{array} \right] \text{ gerade} \\ -1, & \text{falls } \left[\begin{array}{l} \text{die Anzahl der Paare } (i, j) \\ \text{mit } i > j \text{ aber } \sigma(i) < \sigma(j) \end{array} \right] \text{ ungerade} \end{cases} \text{ ist.}$$

Im obigen Beispiel zu S_3 : $\text{sgn}(\sigma) = -1$, $\text{sgn}(\tau) = +1$, $\text{sgn}(\text{id}_{\{1, \dots, n\}}) = 1$.

Ü: Zeige $\text{sgn}(\sigma \circ \tau) = \text{sgn}(\tau \circ \sigma) = \text{sgn}(\tau) \cdot \text{sgn}(\sigma)$. x

1.4 Mengen-Familien und Kartesische Produkte

Eine Familie von Mengen A_i , $i \in \mathcal{J}$, besteht aus einer Indexmenge \mathcal{J} und für jedes $i \in \mathcal{J}$ einer Menge A_i , die man als Teilmengen einer oft nicht explizit erwähnten Grundgesamtheit Ω aller relevanten Elemente auffasst; $A_i = A_j$ für verschiedene Indices $i, j \in \mathcal{J}$ ist zulässig. Formal darf man sich eine Familie als Funktion $A: \mathcal{J} \rightarrow 2^\Omega$, $i \mapsto A(i) := A_i \subseteq \Omega$ vorstellen.

$$\bigcap_{i \in \mathcal{J}} A_i := \{e: \text{für alle } i \in \mathcal{J} \text{ ist } e \in A_i\}, \quad \text{für } \mathcal{J} = \emptyset \text{ ist } \bigcap_{\emptyset} = \Omega,$$

$$\bigcup_{i \in \mathcal{J}} A_i := \{e: \text{es gibt } i \in \mathcal{J} \text{ mit } e \in A_i\}, \quad \text{für } \mathcal{J} = \emptyset \text{ ist } \bigcup_{\emptyset} = \emptyset. [\text{Warum?}]$$

Das kartesische Produkt zweier Mengen bildet die Grundlage für Funktionen. Wir werden aber viel allgemeinere Konstrukte als nur Paare von Elementen aus zwei Mengen benötigen und führen diese nun in zwei Schritten ein.

Seien $n \in \mathbb{N}_0$ und A_1, \dots, A_n Mengen, dann ist

$$A_1 \times A_2 \times \dots \times A_n := \prod_{i=1}^n A_i := \{(a_1, \dots, a_n): a_i \in A_i \text{ für } i = 1, \dots, n\}.$$

Man nennt (a_1, \dots, a_n) ein n -Tupel, das man sich als geordnete Liste mit Zeilen 1 bis n vorstellen kann. Insbesondere ist $()$ das 0-Tupel (oder die „leere Liste“), also $\prod_{i=1}^0 A_i = \{()\}$ eine ein-elementige Menge!

[Diese Definition wird sich später noch als sehr hilfreich erweisen.]

Ganz allgemein werden nun beliebige Elemente als Listenindices zugelassen: Sei A_i , $i \in \mathcal{J}$, eine Familie von Mengen (Teilmengen von Ω), dann bezeichnet ein \mathcal{J} -Tupel eine Funktion $x: \mathcal{J} \rightarrow \Omega$, $i \mapsto x_i \in \Omega$ und

$$\prod_{i \in \mathcal{J}} A_i := \{\mathcal{J}\text{-Tupel } x: x_i \in A_i \text{ für } i \in \mathcal{J}\}.$$

Die vorhergehenden Definitionen entpuppen sich als Spezialfälle, wenn man $n \in \mathbb{N}_0$ als Kurzschreibweise der Indexmenge $\mathcal{J} = \{1, \dots, n\}$ interpretiert,

wobei $\{1, \dots, 0\} = \emptyset$ gesetzt wird. Wie zuvor ist $\times_{i \in \emptyset} A_i = \{()\}$ eine einelementige Menge.

Vorsicht! Formal unterscheidet sich das $\{0, 1, 2\}$ -Tupel vom 3-Tupel dadurch, dass beim ersten die Indexelemente 0, 1, und 2 sind, beim zweiten 1, 2 und 3! Wenn es nur um die Anzahl der Elemente geht und die Indexverwendung klar ist, wird oft nicht so streng unterschieden. ∇

Zwei \mathcal{J} -Tupel x und y sind gleich falls $x_i = y_i$ für $i \in \mathcal{J}$ gilt, formal

$$x = y \quad :\Leftrightarrow \quad \forall i \in \mathcal{J} (x_i = y_i).$$

Für allgemeine Indexmengen ist die Anordnung im \mathcal{J} -Tupel nicht mehr so einfach wie für n -Tupel, sie ist aber auch nicht wichtig. Wenn es eine sinnvolle Anordnung der Menge \mathcal{J} gibt, dann übernimmt man diese meist auch für die Darstellung der Tupel. Aber auch andere Anordnungen sind erlaubt, solange die Zuordnung der Tupelemente zu den Indexelementen klar bleibt.

Beispiel: Bei Einkaufslisten ist \mathcal{J} die Menge der einkaufbaren Produkte und für $i \in \mathcal{J}$ beschreibt A_i die möglichen Mengenangaben (z. B. $A_i = \mathbb{N}_0$ für Stück, \mathbb{Q} für Anteil an Litern/kg). Für eine Einkaufsliste $L \in \times_{i \in \mathcal{J}} A_i$ gibt man meist nur die Einträge $\neq 0$ an, für alle anderen wird 0 angenommen. Wie man Einkaufslisten aufschreibt, ist egal, solange die Produktbezeichnung (die Indexzuordnung) eindeutig bleibt. \heartsuit

Wichtige Spezialfälle:

- Wird in $A_1 \times A_2 \times \dots \times A_k$ eine Menge A_i durch eine Zahl $n \in \mathbb{N}_0$ ersetzt, steht n für die Menge $\{1, \dots, n\}$ (mit $\{1, \dots, 0\} = \emptyset$) wie zuvor.

Beispiel: Für $m, n \in \mathbb{N}_0$ ist

$$m \times n = \{(i, j) : i \in \{1, \dots, m\}, j \in \{1, \dots, n\}\}.$$

Statt (i, j) schreibt man kurz ij , wenn keine Verwechslungsgefahr besteht. \heartsuit

Ü: Wieviele und welche Elemente haben die Mengen $1 \times 2 \times 3$, 2×0 , 0×3 ? \boxtimes

- Gilt für alle $i \in \mathcal{J}$, dass $A_i = A$, sind also alle Mengen der Familie die gleiche Menge A , schreibt man kurz

$$A^{\mathcal{J}} := \times_{i \in \mathcal{J}} A = \{\mathcal{J}\text{-tupel } x : x_i \in A \text{ für } i \in \mathcal{J}\}.$$

Insbesondere ist etwa für $n \in \mathbb{N}_0$

$$A^n = A^{\{1, \dots, n\}} = \times_{i=1}^n A = \{(x_1, \dots, x_n) : x_i \in A \text{ für } i = 1, \dots, n\}.$$

[In solchen n -Tupeln legt die Reihenfolge die Indexzuordnung fest, daher ist hier die Reihenfolge wichtig.]

Beispiele:

- \mathbb{R}^2 lernt man in der Schule als die Ebene, deren Punkte als Zeilen-2-Tupel (x, y) mit kartesischen Koordination x und y beschrieben werden, kennen. Alternativ kann \mathbb{R}^2 auch als Raum der 2-dimensionalen Vektoren aufgefasst werden, dann schreibt man die Elemente meist als Spalten-2-Tupel $\begin{pmatrix} x \\ y \end{pmatrix}$. Genauso kann man \mathbb{R}^3 als 3-dimensionalen Raum von Punkten (x, y, z) oder Vektoren $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ auffassen. Wegen der besseren Zuordnung ist die mathematische Bezeichnung der Koordinaten aber meist durchnummeriert, also statt $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ eher $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, und $x \in \mathbb{R}^3$ bezeichnet dann gleich den ganzen Punkt/Vektor.
- Für $F = \{\diamond, \heartsuit, \clubsuit, \spadesuit\}$ und $W = \{A, K, D, B, 10, \dots, 2\}$ beschreibt $F \times W$ das Französische Blatt.
- Kartenspiel SET: Jede Karte hat 4 Eigenschaften in 3 Ausprägungen:
Anzahl = $\{1, 2, 3\}$, Form = $\{\text{Ellipse, Rechteck, Tilde}\}$,
Farbe = $\{\text{rot, blau, grün}\}$, Füllgrad = $\{\text{leer, schraffiert, gefüllt}\}$,
und die Menge der Karten ist Anzahl \times Form \times Farbe \times Füllgrad.
- Ist die Grundmenge A eine Menge von Zahlen mit Rechenregeln (also z. B. eine aus $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$), so nennt man ein Element $H \in M(m \times n; A) := A^{m \times n}$ eine $m \times n$ **Matrix** (über A) und $M(m \times n; A)$ die Menge der $m \times n$ Matrizen über A . Man ordnet die Einträge von H in m Zeilen und n Spalten in folgender Weise an:

$$H = \begin{bmatrix} H_{11} & \dots & H_{1n} \\ \vdots & \ddots & \vdots \\ H_{m1} & \dots & H_{mn} \end{bmatrix}$$

Die Einträge der Matrix heißen auch **Matrixelemente** und werden oft auch in Kleinschreibung angesprochen, etwa in der Form $H = (h_{ij}) \in \mathbb{R}^{m \times n}$, dann sind h_{ij} und H_{ij} bedeutungsgleich.

Konkret ist z. B. $H = \begin{bmatrix} 5 & -1 & 8 \\ -1 & 6 & 4 \end{bmatrix}$ eine 2×3 -Matrix (über \mathbb{Z}), und $H_{1,3} = 8$. Beachte, dass man je nach Lesbarkeit und Bedarf für das Indexpaar (i, j) nur ij oder i, j schreibt. Soeben war $H_{1,3}$ besser lesbar als H_{13} , aber weniger Arbeit als $H_{(1,3)}$. Man verwendet für Matrizen meistens Großbuchstaben, aber natürlich oft andere als H .

♥

Matrizen und Vektoren mit ihren Rechenregeln und Eigenschaften sind ein Kernthema dieser Vorlesung, denn durch sie lassen sich lineare Gleichungssysteme und die später noch einzuführenden linearen Abbildungen hervorragend beherrschen.

Beispiele:

- In der Schule wird die Lösung linearer Gleichungssysteme meist für drei Variablen und drei Gleichungen besprochen, also etwa ein Beispiel der Art

$$\begin{aligned}x_1 - x_2 + 3x_3 &= 6, \\2x_1 + 5x_2 - 4x_3 &= -5, \\-x_1 + 2x_2 &= 4.\end{aligned}$$

Die **Koeffizienten** der Gleichungen – das sind die Zahlen, mit denen die Variablen x_1, x_2, x_3 multipliziert werden – fassen wir in einer 3×3 Matrix H zusammen, ebenso die rechte Seite der Gleichungen in einem Vektor b und schreiben für das Gleichungssystem kurz

$$H \cdot x = b \quad \text{mit} \quad H = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 5 & -4 \\ -1 & 2 & 0 \end{bmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 6 \\ -5 \\ -4 \end{pmatrix}$$

(es ist ein wenig Geschmacksfrage, wann man eckige und wann man runde Klammern verwendet). Der Ausdruck $H \cdot x$ wird später noch sauber definiert werden; hier steht er einfach für die linke Seite des oberen Gleichungssystems mit dem Vektor der Variablen x , der, wenn möglich, geeignet aus \mathbb{R}^3 gewählt werden soll, um das Gleichungssystem zu erfüllen.

Allgemein hat für n Variablen x_1, \dots, x_n ein reelles lineares Gleichungssystem mit m Gleichungen die Form,

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

mit gegebenen Koeffizienten $h_{ij} \in \mathbb{R}$ für $i = 1, \dots, m, j = 1, \dots, n$ und gegebener rechter Seite $b_i \in \mathbb{R}$. Dieses lässt sich mit der $m \times n$ Matrix $H = (h_{ij}) \in \mathbb{R}^{m \times n}$, $x = (x_j) \in \mathbb{R}^n$ und $b = (b_i) \in \mathbb{R}^m$ genauso kurz wie das 3×3 System als $H \cdot x = b$ schreiben.

- Matrizen und Lineare Gleichungssysteme sind *das* Hauptwerkzeug in der angewandten Mathematik, wenn etwas gerechnet werden muss. Fast alles wird darauf zurückgeführt. Besonders große Beispiele findet man etwa in der Datenspeicherung und -analyse bei Netflix (Benutzer \times Filme $\{0, 1\}$ Matrix) und Google (Webseiten \times Webseiten Matrix).

♡

Die Menge der Abbildungen von A nach B kann mit B^A identifiziert werden und wird oft mit

$$\text{Abb}(A, B) := B^A \quad \text{bezeichnet,}$$

denn jedes Element $f \in B^A$ ist ein A -Tupel, das jedem $x \in A$ genau ein $y = (f)_x \in B$ zuordnet,

$$\Gamma := \{(x, (f)_x) : x \in A\} \subseteq A \times B \quad \text{erfüllt (Graph a) und (Graph b)}.$$

Beispiel: Für jedes $S \subseteq A$ ist die charakteristische Funktion $\chi_S \in \{0, 1\}^A$. Die folgende Funktion ordnet nun jeder Teilmenge ihre charakteristische Funktion zu,

$$\begin{aligned} \chi: 2^A &\rightarrow \{0, 1\}^A \\ S &\mapsto \chi_S \end{aligned}$$

Sie weist jeder Teilmenge $S \subseteq A$ das A -Tupel zu, das in Position a genau dann eine 1 hat, wenn $a \in S$. Jedes A -Tupel aus $\{0, 1\}^A$ passt zu genau einer Teilmenge, wird also einmal und nur einmal von χ zugeordnet, daher ist die Funktion χ bijektiv. \heartsuit

Ü: Zeige: Sind A, B Mengen mit $|B| \geq 2$, dann ist die Kardinalität von A echt kleiner als die der Abbildungen B^A von A nach B . \times

1.5 Relationen

Im Vergleich zu Abbildungen bieten Relationen noch umfassendere Möglichkeiten, Elemente von Mengen zueinander in Beziehung zu setzen.

Eine Relation ist eine Menge, deren Elemente geordnete Paare sind:

Seien A, B Mengen. $R \subseteq A \times B$ ist eine **Relation**.

Gilt $(a, b) \in R$, **stehen a und b in Relation**, formal schreibt man dafür

$$aRb \Leftrightarrow (a, b) \in R.$$

Beispiele:

1. Für eine Abb. $f: A \rightarrow B$ ist ihr Graph $\Gamma \subseteq A \times B$ eine Relation.
Auch die von f induzierte

Faserung $R_f = \{(a_1, a_2) \in A \times A : f(a_1) = f(a_2)\}$ ist eine Relation.

Konkret gilt für $\cos: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \cos(x)$ mit $k \in \mathbb{Z}$ und $x \in \mathbb{R}$ jeweils $xR_{\cos}(x + 2k\pi)$ und $xR_{\cos}(-x + 2k\pi)$.

2. Die Ordnungsrelation $R_{\leq} \subseteq \mathbb{R} \times \mathbb{R}$, $aR_{\leq}b \Leftrightarrow a \leq b$ für $a, b \in \mathbb{R}$,
die Teilmengenrelation $R_{\subseteq} \subseteq 2^A \times 2^A$, $A'R_{\subseteq}A'' \Leftrightarrow A' \subseteq A''$ für $A', A'' \in 2^A$.

3. Sei $Z = \mathbb{N}^2$, dann erklärt $R_Z \subseteq Z \times Z$

$$(n_1, m_1)R_Z(n_2, m_2) \Leftrightarrow n_1 + m_2 = n_2 + m_1 \quad \text{eine Relation auf } Z.$$

Konkret: $(1, 7)R_Z(9, 15)$, nicht aber $(1, 7)$ und $(7, 1)$.

Sei $Q = \mathbb{Z} \times \mathbb{N}$, dann erklärt $R_Q \subseteq Q \times Q$

$$(z_1, n_1)R_Q(z_2, n_2) :\Leftrightarrow z_1 \cdot n_2 = z_2 \cdot n_1 \quad \text{eine Relation auf } Q.$$

Konkret: $(-1, 3)R_Z(-4, 12)$, nicht aber $(1, 3)$ und $(-1, 3)$.

4. Kongruenz modulo m : Sei $m \in \mathbb{N}$, $R_m \subseteq \mathbb{Z} \times \mathbb{Z}$

$$aR_m b :\Leftrightarrow m|a - b \quad (m \text{ teilt } a - b).$$

Man schreibt dann auch $a \equiv b \pmod{m}$

und sagt: a ist kongruent b modulo m .

[a hat den gleichen Rest wie b bei Division durch m .]

Konkret: $15 \equiv 3 \pmod{12}$, $24 \equiv 0 \pmod{12}$, $-1 \equiv 11 \pmod{12}$.



Für eine gegebene Relation $R \subseteq A \times B$ sucht man oft für eine

Teilmenge $X \subseteq A$ alle Elemente $B_R(X) := \{b \in B : \forall a \in X (a, b) \in R\}$

(oder für $Y \subseteq B$ alle Elemente $A_R(Y) := \{a \in A : \forall b \in Y (a, b) \in R\}$),

die zu allen Elementen der Teilmenge in Relation stehen.

Diese Funktionen $B_R: 2^A \rightarrow 2^B$, $A_R: 2^B \rightarrow 2^A$ sind über R eng miteinander verbunden:

Lemma 1.6 (Dualitätslemma für Relationen)

Seien A, B Mengen, $R \subseteq A \times B$ eine Relation und A_R, B_R wie oben, dann gilt für $X, X' \subseteq A$, $Y, Y' \subseteq B$

$$(a) X \subseteq X' \Rightarrow B_R(X) \supseteq B_R(X'), \quad Y \subseteq Y' \Rightarrow A_R(Y) \supseteq A_R(Y'),$$

$$(b) X \subseteq A_R(B_R(X)), \quad Y \subseteq B_R(A_R(Y)),$$

$$(c) B_R(X) = B_R(A_R(B_R(X))), \quad A_R(Y) = A_R(B_R(A_R(Y))).$$

Insbesondere erklärt für $\mathcal{A}_R := \{A_R(Y) : Y \subseteq B\}$, $\mathcal{B}_R := \{B_R(X) : X \subseteq A\}$

$$\begin{array}{l} \bar{A}_R: \mathcal{A}_R \rightarrow \mathcal{B}_R \\ X \mapsto B_R(X) \end{array} \text{ eine bijektive Abbildung mit } \begin{array}{l} \bar{A}_R^{-1} = \bar{B}_R: \mathcal{B}_R \rightarrow \mathcal{A}_R \\ Y \mapsto A_R(Y). \end{array}$$

Beweis: s. Vorlesung □

Eine Relation $R \subseteq A \times A$ (zwischen Elementen der gleichen Menge) heißt

<u>reflexiv</u> ,	wenn aRa	(R)	für alle $a, b, c \in A$ gilt.
<u>symmetrisch</u> ,	wenn $aRb \Rightarrow bRa$	(S)	
<u>antisymmetrisch</u> ,	wenn $aRb \wedge bRa \Rightarrow a = b$	(A)	
<u>transitiv</u> ,	wenn $aRb \wedge bRc \Rightarrow aRc$	(T)	

Zu den vorigen Beispielen:

1. Faserung R, S, T , aber nicht A ,
2. R_{\leq} und R_{\subseteq} R, A, T , aber nicht S ,
3. $R_{\mathbb{Z}}$ und R_Q R, S, T , aber nicht A ,
4. $a \equiv b \pmod{m}$ R, S, T , aber nicht A .

Äquivalenzrelationen

Ein $R \subseteq A \times A$ heißt **Äquivalenzrelation**, falls R reflexiv, symmetrisch und transitiv (RST) ist, man verwendet dann gerne \sim statt R .

Für eine Äquivalenzrelation \sim gilt damit für alle $a, b, c \in A$

$$a \sim a, \quad a \sim b \Leftrightarrow b \sim a, \quad a \sim b \wedge b \sim c \Rightarrow a \sim c.$$

Auf einer Menge bringt jede Äquivalenzrelation automatisch eine Einteilung dieser Grundmenge in Äquivalenzklassen mit sich:

Eine *nichtleere* Teilmenge $K \subseteq A$ heißt **Äquivalenzklasse** (bzgl. \sim), falls

- (a) $x \in K \wedge y \in K \Rightarrow x \sim y$,
- (b) $x \in K \wedge y \in A \setminus K \Rightarrow x \not\sim y$.

Für $a \in A$ definiert $K(a) := \{b \in A : a \sim b\}$ eine Äquivalenzklasse.

[Das ist übrigens gerade $A_{\sim}(\{a\})$ aus dem Dualitätslemma.]

Beweis: s. Vorlesung □

Man schreibt oft $[a] := K(a)$ für die durch a **repräsentierte** Äquivalenzklasse. [Später werden wir oft noch kürzer \bar{a} verwenden.]

Beobachtung 1.7 Für $a, b \in A$ gilt

- (a) $a \sim b \Leftrightarrow K(a) = K(b)$,
- (b) $K(a) \neq K(b) \Rightarrow K(a) \cap K(b) = \emptyset$.

Beweis: s. Vorlesung □

Definition 1.8 Eine Familie $K_i \subseteq A, i \in \mathcal{J}$, von Teilmengen einer Grundmenge A heißt **Klasseneinteilung** oder **Partition** von A , falls folgende drei Eigenschaften erfüllt sind:

- (a) $K_i \neq \emptyset$ für alle $i \in \mathcal{J}$,
- (b) $K_i \cap K_j = \emptyset$ für alle $i, j \in \mathcal{J}$ mit $i \neq j$,
- (c) $A = \bigcup_{i \in \mathcal{J}} K_i$.

Die Beobachtungen oben beweisen:

Satz 1.9 Die Menge der Äquivalenzklassen $\mathcal{K} = \{K(a) : a \in A\}$ einer Äquivalenzrelation auf A ist stets eine Klasseneinteilung von A .

(Verwende \mathcal{K} selbst als Indexmenge \mathcal{J} , $A = \bigcup_{K \in \mathcal{K}} K$.)

Umgekehrt induziert jede Klasseneinteilung eine Äquivalenzrelation:

Satz 1.10 Ist $\mathcal{K} \subseteq 2^A$ eine Klasseneinteilung von A , dann ist die durch

$$a \sim_{\mathcal{K}} b :\Leftrightarrow \exists K \in \mathcal{K} (a \in K \wedge b \in K)$$

auf A definierte Relation $\sim_{\mathcal{K}}$ eine Äquivalenzrelation

Beweis: s. Vorlesung □

Ist K eine Äquivalenzklasse, heißt jedes $a \in K$ Repräsentant der Klasse und es gilt $K = K(a)$. Die Äquivalenzrelation/Klasseneinteilung wird meist verwendet um herauszuarbeiten, welche Elemente der Menge man *nicht mehr* unterscheiden will. Für diese soll es gleichgültig sein, mit welchem Repräsentanten man weiterarbeitet. Man unterscheidet nur noch die Äquivalenzklassen selbst, die „äquivalenten“ Elemente werden „herausfaktoriert“.

Definition 1.11 Sei \sim eine Äquivalenzrelation auf A . Die Menge der Äquivalenzklassen heißt Faktormenge oder Quotientenmenge von A bezüglich \sim und wird mit $A_{/\sim}$ oder $[A]$ bezeichnet.

Beispiele:

zu 1.: Für $f: A \rightarrow B$ induziert die Faserung die Äquivalenzrelation \sim_f .

Für $[a] \in A_{/\sim_f}$ und $a' \in A$ gilt $a' \in [a] \Leftrightarrow f(a') = f(a)$, d. h., die Abbildung

$$\begin{aligned} \bar{f}: A_{/\sim_f} &\rightarrow B \\ [a] &\mapsto f([a]) := f(a) \end{aligned}$$

ist wohldefiniert und nach Konstruktion immer injektiv. [„Wohldefiniertheit“ ist immer zu zeigen, wenn eine Definition gleichzeitig eine Behauptung enthält, hier die, dass durch \bar{f} wirklich eine Funktion definiert wird, also der Wert $f([a])$ unabhängig von der Wahl des Repräsentanten von $[a]$ ist.]

Für das Beispiel $\cos: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \cos(x)$ ist $[x] = \{\pm x + 2k\pi : k \in \mathbb{Z}\}$. Jedes $[x]$ hat genau einen Repräsentanten im Intervall $[0, \pi]$, und weil $\mathbb{R} = \bigcup_{x \in [0, \pi]} [x]$ können wir uns unter $\overline{\cos}: \mathbb{R}_{/\sim_{\cos}} \rightarrow \mathbb{R}$, $[x] \mapsto \cos(x)$ einfach die Einschränkung von \cos auf $[0, \pi]$ vorstellen, d. h., $\overline{\cos}$ verhält sich genau wie $\cos|_{[0, \pi]}: [0, \pi] \rightarrow \mathbb{R}$, $x \mapsto \cos(x)$.

zu 3.: Für $Z = \mathbb{N}^2$ ist R_Z die Äquivalenzrelation

$$(n_1, m_1) \sim_Z (n_2, m_2) :\Leftrightarrow n_1 + m_2 = n_2 + m_1$$

Ü: Zeige: $\bar{f}: \mathbb{N}^2_{/\sim_Z} \rightarrow \mathbb{Z}$, $[(n, m)] \mapsto n - m$ ist wohldefiniert und bijektiv. Wie muss man die Addition auf Z erklären, damit die Addition

beliebiger Repräsentanten immer in der zu dieser Bijektion passenden Äquivalenzklasse landet? [So führt man formal \mathbb{Z} ein.] \times

Für $Q = \mathbb{Z} \times \mathbb{N}$ ist R_Q die Äquivalenzrelation

$$(z_1, n_1) \sim_Z (z_2, n_2) :\Leftrightarrow z_1 \cdot n_2 = z_2 \cdot n_1$$

Ü: Zeige: $\bar{f}: (\mathbb{Z} \times \mathbb{N}) / \sim_Q \rightarrow \mathbb{Q}, [(z, n)] \mapsto \frac{z}{n}$ ist wohldefiniert und bijektiv. Wie muss man die Multiplikation auf Q erklären, damit die Multiplikation beliebiger Repräsentanten immer in der zu dieser Bijektion passenden Äquivalenzklasse landet? [So führt man formal \mathbb{Q} ein.] \times

zu 4.: Kongruenz modulo $m \in \mathbb{N}$: Statt $a \sim_m b$ schreibt man für $a, b \in \mathbb{Z}$

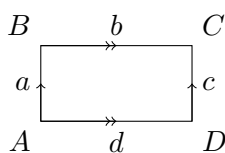
$$a \equiv b \pmod{m} :\Leftrightarrow m | a - b$$

Die Äquivalenzklassen nennt man Restklassen/Kongruenzklassen.

$$\begin{array}{ll} \text{Restklasse zu } 0 & : K(0) = \{x: x = m \cdot k, k \in \mathbb{Z}\} \\ \text{— " — } 1 & : K(1) = \{x: x = m \cdot k + 1, k \in \mathbb{Z}\} \\ & \vdots \\ & \vdots \\ m - 1 & : K(m - 1) = \{x: x = m \cdot k - 1, k \in \mathbb{Z}\} \\ m & : K(m) = K(0) \\ -1 & : K(-1) = K(m - 1) \end{array}$$

Ü: Zeige: $\mathbb{Z} / \equiv_{\text{mod } m} = \{[0], \dots, [m - 1]\}$. Wie überträgt sich die Addition aus \mathbb{Z} auf $\mathbb{Z} / \equiv_{\text{mod } m}$? \times

Topologie: Betrachte ein Rechteck $ABCD$ (etwa ein Blatt Papier).

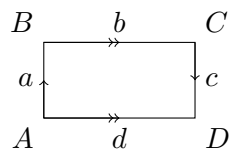


Es bezeichne a die gerichtete Strecke \overrightarrow{AB} von A nach B , b die Strecke \overrightarrow{BC} , c die Strecke \overrightarrow{DC} (!) und d die Strecke \overrightarrow{AD} (!).

In der Relation $\sim_{\substack{a=c \\ b=d}}$: seien die Punkte auf a äquivalent zu den entsprechenden Punkten auf c , die auf b äquivalent zu denen auf d , und alle Punkte im Rechteck nur zu sich selbst äquivalent.

Welche Gestalt hat $ABCD / \sim_{\substack{a=c \\ b=d}}$?

Was passiert, wenn man nun statt dessen c als \overrightarrow{CD} definiert, also AB mit CD gleichsetzt (notfalls erst ohne BC und AD zu identifizieren)?



Kapitel 2

Algebraische Grundbegriffe

Ziel der Algebra ist es, aus den einfachsten und „kleinsten“ gemeinsamen Eigenschaften von (Rechen-)Operationen auf Zahlen- und Funktionenmengen möglichst viele gemeinsame Struktureigenschaften herauszuarbeiten bzw. diese danach einzuordnen.

Da diese einfachen Eigenschaften sehr oft auftreten, gewinnt man viel:

- Die Resultate muss man nur einmal für viele Anwendungen zeigen.
- Die Beweise bleiben „einfach“, weil nur wenig zur Verfügung steht.

Erkennt man diese Eigenschaften/Strukturen in neuen Anwendungen,

- kann man alle Resultate sofort nutzen,
- Querverbindungen zu bekannten Anwendungen und Vorstellungswelten ziehen.

Algebra ist „die Reduktion auf das Allerwesentlichste“. Algebra ist eine eigene Vorlesung im 4. Semester. Wir beschränken uns auf die ersten Schritte, die für das Verständnis der Vorlesung wesentlich sind.

2.1 Halbgruppen, Monoide und Gruppen

Sei A eine Menge. Eine Operation oder Verknüpfung auf A ist eine Abbildung $*$: $A \times A \rightarrow A$, $(a, b) \mapsto *(a, b)$.

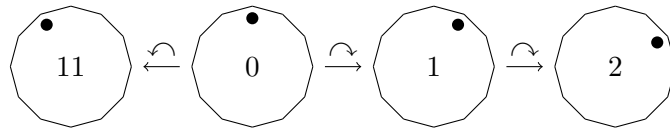
Statt $*(a, b)$ schreibt man meist $a * b$ oder nur ab .

Beispiele:

- Auf $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind die Addition $+(a, b) = a + b$ und die Multiplikation $\cdot(a, b) = a \cdot b$ solche Operationen, auf allen außer \mathbb{C} auch das Maximum $\max(a, b)$ oder Minimum $\min(a, b)$.

- Für eine Menge M sind für $P, Q \in 2^M$ der Schnitt $P \cap Q$ und die Vereinigung $P \cup Q$ Operationen auf der Potenzmenge von M , z. B. $\cap: 2^M \times 2^M \rightarrow 2^M, (P, Q) \mapsto P \cap Q$.
- Die Komposition von Funktionen $f, g: B \rightarrow B$ ist eine Operation $\circ(f, g) = f \circ g$ auf der Menge B^B der Funktionen von B in sich selbst, formal $\circ: B^B \times B^B \rightarrow B^B, (f, g) \mapsto f \circ g = f(g(\cdot))$.
- Auf der Menge von Drehungen/Spiegelungen/Manipulationen von Objekten bildet die Hintereinanderschaltung/Komposition eine Operation, denn zwei Manipulationen hintereinander ausgeführt ergeben wieder eine Manipulation.

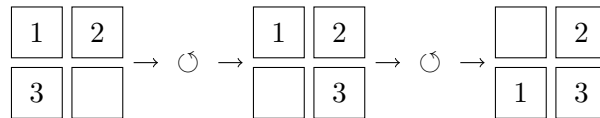
Konkret betrachten wir die Drehungen eines regelmäßigen Zwölfecks in der Ebene, die zu einer der zwölf deckungsgleichen Positionen führen:



Die Grundmenge, sie soll hier \mathcal{D} heißen, ist die der Drehungen um $n \in \mathbb{N}_0$ Positionen (um n mal 30 Grad) nach links oder rechts, also „ n mal \curvearrowright bzw. \curvearrowleft ausführen“. Zwei derartige Drehungen hintereinander ausgeführt ergeben wieder eine dieser Drehungen. Damit ist Menge (Drehungen) und Operation (Komposition) vollständig beschrieben.

Wenn man es hilfreich findet, darf man sich jede Manipulation als eine Funktion vorstellen, die die derzeitige Konfiguration der Objekte auf eine andere Konfiguration abbildet, also eine Funktion von den möglichen Positionen/Zuständen in die möglichen Positionen/Zustände. Für \mathcal{D} bilden die Elemente das Objekt jeweils aus einer seiner zwölf Positionen $\mathcal{P} = \{0, \dots, 11\}$ auf eine andere ab, z. B. $\curvearrowright: \mathcal{P} \rightarrow \mathcal{P}, p \mapsto p + 1 \pmod{12}$. Die Komposition zweier solcher Abbildungen ist wieder eine Abbildung aus \mathcal{D} .

- Betrachte das stark vereinfachte Schiebispiel mit den Basismanipulationen „ \circlearrowright , eins gegen den Uhrzeigersinn schieben“, „ \circlearrowleft , eins im Uhrzeigersinn schieben“



Die Grundmenge, wir nennen sie \mathcal{C} , besteht hier aus den Manipulationen/Abbildungen „ $n \in \mathbb{N}_0$ mal \circlearrowright bzw. \circlearrowleft ausführen“, die Operation ist die Komposition zweier Manipulationen aus \mathcal{C} .

♡

Eine Halbgruppe $(A, *)$ ist eine Menge A mit einer Operation $*$: $A \times A \rightarrow A$, für die das Assoziativgesetz gilt:

$$(\text{Ass}*) \quad (a * b) * c = a * (b * c) \quad \text{für alle } a, b, c \in A.$$

Ist $(A, *)$ eine Halbgruppe, so heißt $e \in A$ neutrales Element,

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

Beobachtung 2.1 *Besitzt eine Halbgruppe $(A, *)$ ein neutrales Element, so ist dieses eindeutig bestimmt.*

Beweis: s. Vorlesung □

Eine Halbgruppe mit neutralem Element heißt Monoid.

Eine Halbgruppe/ein Monoid $(A, *)$ heißt kommutativ oder abelsch, falls

$$a * b = b * a \quad \text{für alle } a, b \in A \text{ gilt (Kommutativgesetz).} \quad (\text{Komm}*)$$

Beispiele:

- Für $A = \mathbb{N}$ sind die Addition $+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, $(n, m) \mapsto n + m$ und die Multiplikation $\cdot: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, $(n, m) \mapsto n \cdot m$ abelsche Halbgruppen $(\mathbb{N}, +)$ und (\mathbb{N}, \cdot) . Neutrales Element für die Addition gibt es nicht, für die Multiplikation ist es die 1. Wie ist das für $A \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ mit $(A, +)$ und (A, \cdot) ?
- Für eine Menge M ist $(2^M, \cap)$ abelsches Monoid mit neutralem Element M . Wie ist das bei $(2^M, \cup)$?
- Für eine Menge B ist $(B^B = \text{Abb}(B, B), \circ)$ nicht-kommutatives (warum?) Monoid mit neutralem Element id_B .
- Im Fall der Zwölfeck-Drehungen bildet (\mathcal{D}, \circ) ein kommutatives Monoid mit neutralem Element „0 mal drehen“.
- Im Schiebispiel bilden die Manipulationen \mathcal{C} mit Hintereinanderschaltung ein kommutatives Monoid (\mathcal{C}, \circ) mit neutralem Element „0 mal schieben“.
- Für $A = \{0, 1\}$ seien über folgende Tabellen die Operationen „Addition“ $+: \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$, $(a, b) \mapsto a + b$ und „Multiplikation“ $\cdot: \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$, $(a, b) \mapsto a \cdot b$ erklärt:

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

Dann sind $(A, +)$ und (A, \cdot) abelsche Monoide.



In vielen Anwendungen steht $*$ für eine Form der Multiplikation \cdot , oder $-$ meist in Kombination mit abelsch $-$ für eine Form der Addition $+$.

Falls wir eine Halbgruppe/Monoid in additiver Form $(A, +)$ schreiben, nennen wir ein neutrales Element Nullelement (meist 0) und in multiplikativer Form (A, \cdot) Einselement (meist e oder 1).

Ist $(A, *)$ ein Monoid, so heißt $a \in A$ invertierbar, falls es ein

$$\text{inverses Element } a' \in A \text{ gibt mit } a * a' = a' * a = e.$$

Brauchen wir „links“ und „rechts“?

Beobachtung 2.2 *Ist $(A, *)$ ein Monoid und gibt es für ein $a \in A$ ein Linksinverses a' mit $a' * a = e$ sowie ein Rechtsinverses a'' mit $a * a'' = e$, so ist $a' = a''$. Insbesondere ist ein inverses Element eindeutig bestimmt.*

Beweis: s. Vorlesung □

Beispiele:

- In (\mathbb{N}, \cdot) hat nur 1 ein Inverses, in $(\mathbb{Z}, +)$ ist für $a \in \mathbb{Z}$ das Inverse $-a$ (mit $0 = -0$), in (\mathbb{Z}, \cdot) haben nur zwei Zahlen ein Inverses. Wie ist das für $A \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $(A, +)$, (A, \cdot) und $(A \setminus \{0\}, \cdot)$?
- Welche Elemente von $(2^M, \cap)$ und $(2^M, \cup)$ haben ein Inverses?
- Welche Elemente von $(B^B = \text{Abb}(B, B), \circ)$ haben ein Inverses?
- Drehungen des Zwölfecks: Was ist das Inverse zu n -mal \curvearrowright ?
- Schiebepfeil: Was ist das Inverse zu n -mal \circlearrowright ?
- Wie ist das für $(\{0, 1\}, +)$ und $(\{0, 1\}, \cdot)$?



Multiplikativ (A, \cdot) schreibt man a^{-1} für das Inverse (falls es existiert) sowie $a^n := \underbrace{a \cdot \dots \cdot a}_{n\text{-mal}}$, $a^{-n} := \underbrace{a^{-1} \cdot \dots \cdot a^{-1}}_{n\text{-mal}}$, $a^0 := e$.

Multipliziert man in einem Monoid (A, \cdot) die Elemente $a_1, \dots, a_n \in A$ mit $n \in \mathbb{N}_0$, schreibt man kurz

$$\prod_{i=1}^n a_i := a_1 \cdot \dots \cdot a_n \quad \text{mit} \quad \prod_{i=1}^0 := 1.$$

Ist das Monoid kommutativ und bildet man das Produkt über eine Familie $(a_i)_{i \in \mathcal{J}}$, in der die Menge der Indices $\{i \in \mathcal{J} : a_i \neq 1\}$ endlich ist, schreibt man ebenso $\prod_{i \in \mathcal{J}} a_i$ und setzt $\prod_{\emptyset} := 1$.

Beispiele:

- $n! = \prod_{i=1}^n i$ und $0! = \prod_{i=1}^0 i = 1$; das ist sinnvoll und praktisch, man denke z. B. an die Möglichkeiten k Elemente aus n auszuwählen: $\frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$ ist nun auch für $k = n$ einsetzbar.
- Die Hintereinanderausführung ist i. A. nicht kommutativ und wird meist Multiplikativ notiert, selbst wenn Kommutativität in Spezialfällen wie im Schiebepiel gelten sollte. Als Beispiel kennen wir bereits $f \circ f^{-1} = f^0 = \text{id}$ und $f \circ f \circ f = f^3$.

Vorsicht! Im Allgemeinen ist $f^3(x) \neq f(x)^3$, z. B. für $f: \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$ ist $f^3(x) = ((x^2)^2)^2 = x^8 \neq f(x)^3 = (x^2)^3 = x^6$. [Algebraisch gesehen bezieht sich das erste 3 auf die Operation \circ aus $(\text{Abb}(\mathbb{R}, \mathbb{R}), \circ)$, das zweite 3 auf die reelle Multiplikation (\mathbb{R}, \cdot) .] ∇

Bei den Drehungen des Zwölfecks steht \curvearrowright^n , $n \in \mathbb{Z}$, für n -mal \curvearrowright ausführen. Dabei ist $\curvearrowright = \curvearrowright^{-1}$ und $e = \curvearrowright^0 = \curvearrowright \circ \curvearrowright$. Was ist $\curvearrowright^5 \circ \curvearrowright^3 \circ \curvearrowright^{-2}$? Brauchen wir unendlich großes n ? Nein, denn $\curvearrowright^{12} = e$, also ist $\curvearrowright^{13} = \curvearrowright$, und $\curvearrowright^{11} = \curvearrowright^{-1} = \curvearrowleft$, insgesamt also $\mathcal{D} = \{\curvearrowright^n : n \in \{0, \dots, 11\}\}$. Ebenso sieht man im Schiebepiel $\mathcal{C} = \{\cup^n : n \in \{0, \dots, 11\}\}$.

♡

Additiv $(A, +)$ schreibt man $-a$ für das Inverse (falls es existiert)

$$\text{sowie } na := \underbrace{a + \cdots + a}_{n\text{-mal}}, \quad -na := \underbrace{(-a) + \cdots + (-a)}_{n\text{-mal}}, \quad 0a := 0,$$

und statt $a + (-b)$ einfach $a - b$.

Summiert man in einem Monoid $(A, +)$ die Elemente $a_1, \dots, a_n \in A$ mit $n \in \mathbb{N}_0$, schreibt man kurz

$$\sum_{i=1}^n a_i := a_1 + \cdots + a_n \quad \text{mit} \quad \sum_{i=1}^0 := 0.$$

Ist das Monoid kommutativ und bildet man die Summe über eine Familie $(a_i)_{i \in \mathcal{J}}$, in der die Menge der Indices $\{i \in \mathcal{J} : a_i \neq 0\}$ endlich ist, schreibt man ebenso $\sum_{i \in \mathcal{J}} a_i$ und setzt $\sum_{\emptyset} := 0$.

Ein Monoid $(A, *)$, in dem jedes Element invertierbar ist, heißt Gruppe (abelsche/kommutative Gruppe, falls $*$ kommutativ ist), genauer reicht schon folgendes aus:

Definition 2.3 Eine Menge A mit einer Verknüpfung $*$: $A \times A \rightarrow A$ heißt Gruppe, falls

(G1) $(a * b) * c = a * (b * c)$ „Assoziativgesetz“

(G2) $\exists e \in A$ mit

(G2a) $e * a = a$ für alle $a \in A$, „e ist linksneutrales Element“

(G2b) zu jedem $a \in A$ gibt es ein $a' \in A$ mit $a' * a = e$

„linksinverses Element“

Beobachtung 2.4 Eine Gruppe $(A, *)$ ist auch ein Monoid und es gilt:

- (a) Ein linksneutrales Element e erfüllt auch $a * e = a$ für alle $a \in A$ und ist somit eindeutiges neutrales Element.
- (b) Für jedes $a \in A$ erfüllt ein linksinverses Element a' auch $a * a' = e$ und ist somit eindeutiges inverses Element, wir nennen es a^{-1} .
- (c) Für $a, b \in A$ gilt: $a^{-1} * a = a * a^{-1} = e$, $(a^{-1})^{-1} = a$, $(a * b)^{-1} = b^{-1} * a^{-1}$.
- (d) Kürzungsregeln: $(a * x = a * \hat{x}) \Rightarrow (x = \hat{x})$ und $(x * a = \hat{x} * a) \Rightarrow (x = \hat{x})$.

Beweis (1. Variante): s. Vorlesung □

Folgende zweite Beweisvariante zeigt besonders schön, dass für jedes feste Element die Gruppenoperation von links (von rechts) mit diesem festem Element die Gruppenelemente bijektiv auf sich selbst abbildet (im nichtkommutativen Fall können sich diese Bijektionen für links und rechts unterscheiden). Für $a \in A$ bezeichne die

$$\text{Linkstranslation } {}_a\tau: A \rightarrow A, b \mapsto a * b \quad [= *(a, \cdot)]$$

die „Multiplikation“ mit a von links.

Beweis (2. Variante für (a),(b)): s. Vorlesung □

Analog zur Linkstranslation definiert man für $a \in A$ die

$$\text{Rechtstranslation } \tau_a: A \rightarrow A, b \mapsto b * a \quad [= *(\cdot, a)].$$

Auch diese ist für Gruppen $(A, *)$ bijektiv (nutze die Kürzungsregel).

Beispiele:

- abelsche/kommutative Gruppen: $(\mathbb{Z}, +)$ und für $A \in \{\{0, 1\}, \mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ jeweils $(A, +)$ und $(A \setminus \{0\}, \cdot)$. Für kommutative Gruppen ist ${}_a\tau = \tau_a$, konkret für $(\mathbb{Z}, +)$ und $2 \in \mathbb{Z}$ etwa ${}_2\tau = \tau_2: \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto 2 + z = z + 2$.
- $(B^B = \text{Abb}(B, B), \circ)$ ist i. A. keine Gruppe, denn es gibt keineswegs zu jedem $f: B \rightarrow B$ eine Umkehrabb. mit $f^{-1} \circ f = f \circ f^{-1} = \text{id}_B$. Das geht nur für bijektive Abbildungen. Sei $S(B) := \{f \in \text{Abb}(B, B): f \text{ bijektiv}\}$. $(S(B), \circ)$ ist die symmetrische Gruppe (Ü!). Gilt $|B| = n \in \mathbb{N}$, kann man B bijektiv mit $\{1, \dots, n\}$ identifizieren, $\rightarrow (S(B), \circ)$ verhält sich wie (S_n, \circ) , die Permutationsgruppe aus Abschnitt 1.3. Sie ist nur für $n = 1, 2$ kommutativ, für $n \geq 3$ nicht mehr.
- Die Drehungen des Zwölfecks bilden eine kommutative Gruppe (\mathcal{D}, \circ) .
- Im Schiebepiel ist (\mathcal{C}, \circ) eine kommutative Gruppe.
- Rubic's Cube: Die Folgen der Seitendrehungen bilden eine nicht kommutative Gruppe:
Inverses Element: rückwärts drehen; neutrales Element: keine Drehung; Assoziativgesetz: wie man die Drehungen zu Teilfolgen zusammenfasst, ist egal.
- Ähnlich: Rotationen und Spiegelungen geometrischer Objekte.



Definition 2.5 Ist $(A, *)$ Gruppe und erfüllt eine Teilmenge U von A

(U1) $\emptyset \neq U \subseteq A$,

(U2) mit $a, b \in U$ ist $a * b \in U$ und $a^{-1} \in U$,

so heißt $(U, *)$ Untergruppe von $(A, *)$.

Ü: Für jede Untergruppe U von $(A, *)$ ist $e \in U$, $(U, *)$ selbst eine Gruppe und für $a \in A$ ist

$$a \in U \iff {}_a\tau(U) = U \iff \tau_a(U) = U.$$

Das heißt, für die Elemente $a \in U$, und nur für diese, bilden Links- und Rechstranslation U bijektiv auf U ab. x

Beispiel: Sei $(A, *)$ ein Gruppe und $a \in A$. Dann ist

$$\langle a \rangle := \{a^n : n \in \mathbb{Z}\}, *$$

die von a erzeugte zyklische Untergruppe von $(A, *)$, denn $a^0 = e \in \langle a \rangle$ und $a^n * a^m = a^{n+m} \in \langle a \rangle$ für $n, m \in \mathbb{Z}$.

Man nennt a den Erzeuger der Gruppe $\langle a \rangle$. Eine Gruppe $(A, *)$ heißt zyklisch, wenn sie von einem $a \in A$ erzeugt wird, also falls es ein $a \in A$ gibt mit $\langle a \rangle = A$.

Konkret:

- Für das neutrale Element gilt $\langle e \rangle = \{e\}$, es ist die kleinste Untergruppe von $(A, *)$.
- Für $(\mathbb{Z}, +)$ ist $\langle 2 \rangle = \{0, \pm 2, \pm 4, \dots\}$, $\langle 0 \rangle = \{0\}$ und allgemein, für $m \in \mathbb{Z}$, $\langle m \rangle = m\mathbb{Z} = \{0, \pm m, \pm 2m, \dots\}$.
- Für $(\mathbb{Q}, +)$ und $a = 1$ ist $\langle 1 \rangle = \mathbb{Z}$.
- Drehungen: $(\mathcal{D} = \{\curvearrowright^n : n \in \{0, \dots, 11\}\}, \circ)$ ist selbst eine zyklische Gruppe, denn $\mathcal{D} = \{\curvearrowright^n : n \in \mathbb{Z}\}$. Eine Untergruppe davon wäre etwa $(\{\curvearrowright^{3n} : n \in \mathbb{Z}\}, \circ)$, die Anzahl ihrer Elemente ist 4.
- Schiebepispiel: das gleiche gilt für $(\mathcal{C} = \{\circlearrowleft^n : n \in \{0, \dots, 11\}\}, \circ)$.



Sei $(A, *)$ eine Gruppe und $(U, *)$ eine Untergruppe. Für $a \in A$ heißt

$a * U := \{a * u : u \in U\} = {}_a\tau(U)$ Linksnebenklasse von U in A ,

$U * a := \{u * a : u \in U\} = \tau_a(U)$ Rechtsnebenklasse von U in A .

Vorsicht! Hier wird $*$ nicht strikt als die definierte Operation, sondern nur als intuitive Schreibweise genutzt! ▽

Beispiel: Bei den Drehung mit $U = (\{\curvearrowright^{3n} : n \in \mathbb{Z}\}, \circ)$ ist $\curvearrowright \circ U = \{\curvearrowright^{3n+1} : n \in \mathbb{Z}\} = U \circ \curvearrowright$ wegen der Kommutativität. Allgemein ist hier $a \in A$ von der

Form $a = \curvearrowright^k$ für ein $k \in \mathbb{Z}$ und $a \circ U = \{\curvearrowright^{3n+k} : n \in \mathbb{Z}\} = U \circ a$. Wieviele unterschiedliche Nebenklassen gibt es? Wie ist das im Schiebespiel? ♡

Wegen der Bijektivität von ${}_a\tau$ und τ_a auf A folgt

Satz 2.6 *Ist $(U, *)$ eine Untergruppe der Gruppe $(A, *)$, so sind alle Linksnebenklassen gleichmächtig und bilden eine Klasseneinteilung von A . insbesondere sind für $a, b \in A$ äquivalent:*

- (a) $a * U = b * U$,
- (b) $a * U \cap b * U \neq \emptyset$,
- (c) $a \in b * U$,
- (d) $b^{-1} * a \in U$.

Beweis: s. Vorlesung □

Bemerkung: Weil es zu jedem Gruppenelement ein eindeutiges Inverses gibt, ist die Abbildung $\cdot^{-1} : A \rightarrow A, a \mapsto a^{-1}$ bijektiv (auch für jede Untergruppen) und bildet für $a \in A$ Linksnebenklasse $a * U$ bijektiv auf $U * a^{-1}$ ab ($a * u \mapsto u^{-1} * a^{-1}$, B.2.4(c)), daher gilt ein analoger Satz für Rechtsnebenklassen. ◇

Die Menge der Linksnebenklassen von U in A bezeichnet man mit A/U .

Die Menge der Rechtsnebenklassen von U in A bezeichnet man mit $U \setminus A$.

A/U und $U \setminus A$ sind wegen der Bijektion \cdot^{-1} gleichmächtig und ihre Mächtigkeit (=Anzahl der Klassen) heißt der Index, in Zeichen $(A : U)$, von U in A , also $(A : U) = |A/U| = |U \setminus A|$.

Die Mächtigkeit von A selbst wird die Ordnung $\text{ord } A$ genannt, $\text{ord } A := |A|$.

Korollar 2.7 (Satz von Lagrange) *Ist $(A, *)$ eine endliche Gruppe und $(U, *)$ eine Untergruppe, so gilt $\text{ord } A = \text{ord } U \cdot (A : U)$. [Multipl. in \mathbb{N}]*

Beispiel: Drehungen: $\text{ord } \mathcal{D} = |\mathcal{D}| = 12$, die Untergruppe U hat $\text{ord } U = |U| = 4$, und der Index von U in \mathcal{D} ist $(\mathcal{D} : U) = 3$. ♡

Definition 2.8 *Eine Untergruppe $(U, *)$ von $(A, *)$ heißt Normalteiler oder normale Untergruppe, falls $a * U = U * a$ für alle $a \in A$ gilt, also falls Links- und Rechtsnebenklassen übereinstimmen (immer bei abelsch!).*

*Die zu $a \in A$ gehörende Nebenklasse $a * U$ heißt dann Restklasse von a modulo U .*

Beispiel: Für $(\mathbb{Z}, +)$ und $m \in \mathbb{Z}, U = \langle m \rangle = m\mathbb{Z} = \{m \cdot k : k \in \mathbb{Z}\}$

ist für $a \in \mathbb{Z}$ $a + U = \{a + m \cdot k : k \in \mathbb{Z}\} = U + a$.

Ebenso gilt für $a, b \in \mathbb{Z}$ mit $a \equiv b \pmod{m}$, dass $a + U = b + U$, die Bezeichnungen stimmen also überein. ♡

Für Normalteiler lässt sich $*$ als Operation auf die Nebenklassen übertragen.

Für $X \subseteq A, Y \subseteq A$ setze $X \hat{*} Y := \{x * y : x \in X, y \in Y\}$ (das erklärt eine Operation $\hat{*} : 2^A \times 2^A \rightarrow 2^A, (X, Y) \mapsto X \hat{*} Y$), dann ist für $u' \in U$ stets $\{u'\} \hat{*} U = u' * U = U = e * U$ und für $a, b \in A$

$$(a * U) \hat{*} (b * U) = \underbrace{\{a\} \hat{*} U \hat{*} \{b\}}_{=U * b \stackrel{\text{Normalt.}}{=} b * U} \hat{*} U = \underbrace{\{a\} \hat{*} \{b\}}_{\{a * b\}} \hat{*} \underbrace{U \hat{*} U}_U = (a * b) * U.$$

Vorsicht! Statt $\hat{*}$ nutzt man meistens das gleiche Symbol $*$ weiter, da sich die neue Verwendung direkt aus der alten ergibt. Man sagt, dass sie von der ursprünglichen induziert wird. ∇

$(A/U, \hat{*})$ heißt Faktor-, Quotienten- oder Restklassengruppe von A modulo U und ist wirklich ein Gruppe:

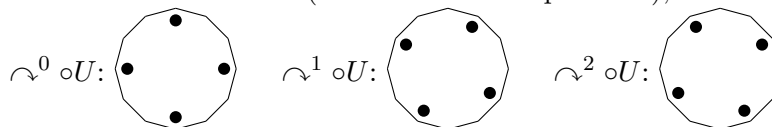
- (G1) gilt, weil schon $*$ assoziativ in $(A, *)$ ist,
- (G2) gilt, weil $e * U$ als neutrales Element fungiert und für $a * U$ das Inverse $a^{-1} * U \in A/U$ ist.

Die Wahl des Repräsentanten $a' \in a * U$ und $b' \in b * U$ ist für das Ergebnis egal („wohldefiniert“), denn aus S 2.6(d) folgt $a^{-1} * a' \in U \wedge b^{-1} * b' \in U$, also $(a^{-1} * a') * U = U = (b^{-1} * b') * U$ und

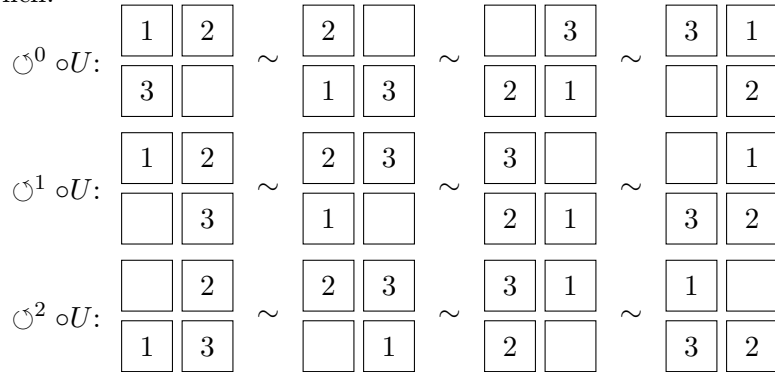
$$\begin{aligned} (a * U) \hat{*} (b * U) &= \underbrace{(a * ((a^{-1} * a') * U))}_e \hat{*} \underbrace{(b * ((b^{-1} * b') * U))}_e \\ &= (a' * U) \hat{*} (b' * U) = (a' * b') * U. \end{aligned}$$

Beispiele:

- Für $(A, *)$ mit neutralem Element e ist für $U = \langle e \rangle = \{e\}$ die Quotientenmenge $A/\langle e \rangle = \{a * \langle e \rangle : a \in A\} = \{\{a\} : a \in A\}$ und für $a, b \in A$ ist $(a * U) \hat{*} (b * U) = \{a * b\} \in A/\langle e \rangle$, also identifiziert man gleich $A/\langle e \rangle$ mit A sowie $\hat{*}$ mit $*$ und schreibt salopp $(A/\langle e \rangle, \hat{*}) = (A, *)$.
- $(\mathbb{Z}/m\mathbb{Z}, +)$ ist für jedes $m \in \mathbb{N}_0$ wieder eine Gruppe, für $m > 0$ nennt man sie zyklische Gruppe der Ordnung m (Resklassen $[0], \dots, [m - 1]$)
für $m = 0$ ist $(\mathbb{Z}/0\mathbb{Z}, +) = (\mathbb{Z}, +)$ die unendliche zyklische Gruppe.
Konkret: $(\mathbb{Z}/9\mathbb{Z}, +): [7] + [3] = [1], -[4] = [5]$.
Wo findet man $(\mathbb{Z}/12\mathbb{Z}, +)$ im täglichen Leben?
- Drehungen: $(\mathcal{D} = \{\curvearrowright^n : n \in \{0, \dots, 11\}\}, \circ)$ ist eine zyklische Gruppe der Ordnung 12. Für $U = \{\curvearrowright^{3n} : n \in \mathbb{Z}\}$ ist $(\mathcal{D}/U = \{\curvearrowright^n \circ U : n \in \{0, 1, 2\}\}, \circ)$ die Quotientengruppe von \mathcal{D} modulo U . Anschaulich bedeutet die Quotientenbildung, dass wir die Manipulationen aus U für ununterscheidbar halten (sie sind für uns äquivalent), etwa so:



- Schiebepiel: $(\mathcal{C} = \{\circ^n : n \in \{0, \dots, 11\}\}, \circ)$ ist offenbar eine zyklische Gruppe der Ordnung 12. Für $U = \{\circ^{3n} : n \in \mathbb{Z}\}$ ist $(\mathcal{C}/U = \{\circ^n \circ U : n \in \{0, 1, 2\}\}, \circ)$ die Quotientengruppe von \mathcal{C} modulo U , anschaulich:



Es kommt also nicht mehr darauf an, an welcher Stelle das leere Kästchen steht, sondern nur noch ob es um 0, 1 oder 2 Positionen relativ zu drei unterscheidbaren Objekten mit vorgegebener Reihenfolge verschoben wurde. Das wäre etwa der Fall, wenn wir statt des Schiebepiels eine Kette mit vier Perlen in den Farben rot, grün, blau und weiß betrachten, für die die Reihenfolge von rot, grün und blau vorgegeben ist, aber die Position von weiß jeweils mit einem Nachbarn vertauscht werden darf.

♡

2.2 Homomorphismen

Homomorph kommt von $\acute{o}\mu\acute{o}\iota\omicron\varsigma$ (ähnlich) und $\mu\omicron\rho\varphi\acute{\eta}$ (Gestalt).

Definition 2.9 Sind $(A, *)$ und (B, \odot) Gruppen, dann heißt eine Abbildung $\varphi: A \rightarrow B$ (Gruppen-)Homomorphismus, wenn

$$\varphi(a * b) = \varphi(a) \odot \varphi(b) \quad \text{für alle } a, b \in A \text{ gilt.}$$

D.h., φ ist mit den Operationen auf A und B verträglich (es ist egal, ob die Operation vor Anwendung oder erst nach Anwendung der Abbildung ausgeführt wird).

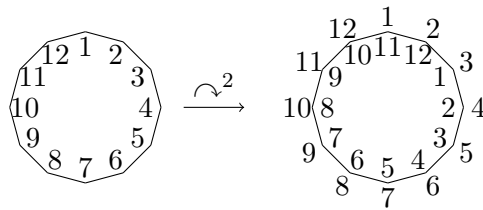
Vorsicht! Homöomorph ist ein ganz anderer Begriff, der für stetige Abb. mit stetiger Inversen verwendet wird! ▽

Beispiele:

- Die natürlichen Einbettungen $\varphi_{\mathbb{Z}}^{\mathbb{Q}}: \mathbb{Z} \rightarrow \mathbb{Q}, z \mapsto z, \varphi_{\mathbb{Q}}^{\mathbb{R}}: \mathbb{Q} \rightarrow \mathbb{R}, q \mapsto q$ sind Homomorphismen (erstere nur für $+$, letztere für $+$ und \cdot), aber

auch $\varphi: \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto 2z$ ist wegen $\varphi(z_1 + z_2) = 2(z_1 + z_2) = 2z_1 + 2z_2 = \varphi(z_1) + \varphi(z_2)$ ein Homomorphismus.

- Weitere Beispiele sind $\varphi_{\mathcal{D}}^{\mathcal{C}}: \mathcal{D} \rightarrow \mathcal{C}, \curvearrowright^k \mapsto \cup^k$ von (\mathcal{D}, \circ) nach (\mathcal{C}, \circ) oder $\varphi_{\mathcal{D}}^{\mathbb{Z}/12\mathbb{Z}}: \mathcal{D} \rightarrow \mathbb{Z}/12\mathbb{Z}, \curvearrowright^k \mapsto [k]$ von (\mathcal{D}, \circ) nach $(\mathbb{Z}/12\mathbb{Z}, +)$ (Ü).
- In der Geometrie studiert man die Symmetrien eines Objekts über Homomorphismen von der Gruppe der Drehungen und Spiegelungen, die das Objekt äußerlich unverändert lassen, in die symmetrische Gruppe. Als Beispiel betrachte die Drehungen \mathcal{D} . Nach Anwendung von $a \in \mathcal{D}$ liegt jede Ecke des Zwölfecks nun auf der ehemaligen Position einer anderen, z. B. für $a = \curvearrowright^2$:



Durch \curvearrowright^2 wird Ecke 1 zu Ecke 3, Ecke 2 zu Ecke 4 etc. Dies ist eine bijektive Abbildung von den Ecken auf die Ecken, also ein Element der Symmetrischen Gruppe über den Ecken. Wenn wir die Ecken mit $\{1, \dots, 12\}$ benennen, ist es eine Permutation $\sigma \in S_{12}$. Für jedes $a \in \mathcal{D}$ beschreibt eine eigene Permutation die Veränderung der Lage, und zwar für \curvearrowright^k die Permutation $\sigma_k: \{1, \dots, 12\} \rightarrow \{1, \dots, 12\}, i \mapsto (i - 1 + k \text{ mod } 12) + 1$. Das definiert eine Funktion

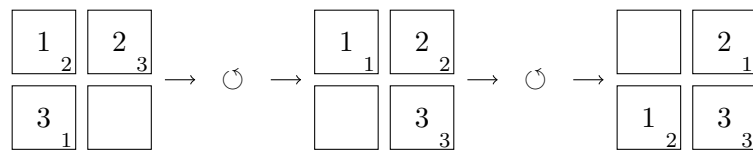
$$\rho: \mathcal{D} \rightarrow S_{12}, \curvearrowright^k \mapsto \sigma_k.$$

Nun ist $\sigma_k = (\sigma_1)^k$ (Ü), daher gilt für $\curvearrowright^h, \curvearrowright^k \in \mathcal{D}$,

$$\rho(\curvearrowright^h \circ \curvearrowright^k) = \rho(\curvearrowright^{h+k}) = \sigma_{h+k} = (\sigma_1)^{h+k} = (\sigma_1)^h \circ (\sigma_1)^k = \rho(\curvearrowright^h) \circ \rho(\curvearrowright^k).$$

Also ist ρ ein Gruppenhomomorphismus von (\mathcal{D}, \circ) nach (S_{12}, \circ) . Einen Homomorphismus von einer Gruppe in die symmetrische Gruppe nennt man auch eine Darstellung der Gruppe. Je nach Wahl des Objekts, auf das man die Gruppe wirken lässt, kann es sehr unterschiedliche Darstellungen geben, als Beispiel untersuche man die Wirkung von \mathcal{D} auf ein regelmäßiges 24-Eck.

- Um im Schiebespiel mitverfolgen zu können, welches Teil als nächstes durch \cup den Platz verändern wird, weisen wir in jeder Konstellation den Teilen im Uhrzeigersinn ab dem leeren Platz die Nummern 1, 2, und 3 zu:



Für jede Konfiguration erhalten wir jeweils genau eine Permutation der Elemente $\{1, 2, 3\}$, also ein $\sigma \in S_3$. Wie verändert sich σ bei Anwendung von \circlearrowleft ? Der alte Platz 2 wird zu Platz 1, Platz 3 zu Platz 2 und Platz 1 zu Platz 3, also mit $\tilde{\sigma} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ wird aus σ bei Anwendung von \circlearrowleft nun $\tilde{\sigma} \circ \sigma$. Die entsprechende Veränderung bei Anwendung von $\circlearrowleft^k \in \mathcal{C}$ beschreibt die Abbildung

$$\gamma: \mathcal{C} \rightarrow S_3, a = \circlearrowleft^k \mapsto \varphi(a) := \tilde{\sigma}^k.$$

Dieses γ ist ein Gruppenhomomorphismus ($\ddot{\mathbf{U}}$).



Ist ein Homomorphismus
 injektiv, nennt man ihn Monomorphismus,
 surjektiv, — " — Epimorphismus,
 bijektiv, — " — Isomorphismus.

Ein Homomorphismus $\varphi: A \rightarrow A$ heißt auch Endomorphismus,
 ein Isomorphismus $\varphi: A \rightarrow A$ heißt auch Automorphismus.

Beispiele:

- Für $(\mathbb{Z}, +)$ ist $\varphi: \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto 2z$ injektiv, also ein Monomorphismus oder ein injektiver Endomorphismus, aber nicht surjektiv.
 Für $(\mathbb{Q}, +)$ ist $\varphi': \mathbb{Q} \rightarrow \mathbb{Q}, q \mapsto 2q$ bijektiv, also ein Isomorphismus oder Automorphismus
- $\varphi_{\mathcal{D}}^{\mathcal{C}}$ und $\varphi_{\mathcal{D}}^{\mathbb{Z}/12\mathbb{Z}}$ sind Isomorphismen.
- Bei den Drehungen ist ρ ein Monomorphismus, aber nicht surjektiv.
- Im Schiebepiel ist γ weder surjektiv noch injektiv.



Homomorphismen übertragen die algebraische Struktur und erlauben auch, diese Strukturen selbst zu analysieren.

Von besonderer Bedeutung sind dabei für $\varphi: A \rightarrow B$

$$\begin{aligned} \text{Im } \varphi &:= \varphi(A) = \{b = \varphi(a) : a \in A\} \subseteq B && \text{das Bild von } \varphi, \\ \text{Ker } \varphi &:= \varphi^{-1}(\{e_B\}) = \{a \in A : \varphi(a) = e_B\} \subseteq A && \text{der Kern von } \varphi. \end{aligned}$$

$\ddot{\mathbf{U}}$: Für γ aus dem Schiebepiel gilt:

$$\text{Im } \gamma = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$

$$\text{Ker } \gamma = \{\circlearrowleft^{3n} : n \in \mathbb{Z}\}$$



Das Bild $\text{Im } \varphi$ ist immer abgeschlossen bzgl. der Operation \odot :

$$\text{Mit } \varphi(a), \varphi(b) \in \text{Im } \varphi \text{ ist auch } \varphi(a) \odot \varphi(b) = \varphi(a * b) \in \text{Im } \varphi.$$

Wir zeigen nun noch, dass $\varphi(e_A) = e_B \in \text{Im } \varphi$ (das zeigt (U1)) und dass mit $\varphi(a)$ auch $\varphi(a)^{-1} = \varphi(a^{-1}) \in \text{Im } \varphi$ (das zeigt (U2)), also ist $(\text{Im } \varphi, \odot)$ immer eine Untergruppe von B .

Beobachtung 2.10 Sei $\varphi: A \rightarrow B$ ein Gruppenhomomorphismus. Es gilt:

- (a) $\varphi(e_A) = e_B$, wobei e_A und e_B die neutralen Elemente von A und B bezeichnen.
- (b) $\varphi(a^{-1}) = \varphi(a)^{-1}$ für alle $a \in A$.
- (c) Ist φ ein Isomorphismus, dann ist auch $\varphi^{-1}: B \rightarrow A$ ein Isomorphismus.

Beweis: s. Vorlesung □

Beispiele:

- Für $(\mathbb{Z}, +)$ und $\varphi: \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto 2z$ ist $\varphi(0) = 0$ und $\varphi(-z) = 2(-z) = -2z = -\varphi(z)$, aber φ ist nur injektiv, nicht surjektiv.
Für $(\mathbb{Q}, +)$ ist $\varphi': \mathbb{Q} \rightarrow \mathbb{Q}, q \mapsto 2q$ ein Isomorphismus. Wie sieht der inverse Isomorphismus aus?
- Im Schiebepfeil ist $\gamma(e) = \gamma(\cup^0) = \tilde{\sigma}^0 = \text{id}$, etc.

♡

Für Homomorphismen $\varphi: A \rightarrow B, \psi: B \rightarrow C$ zu Gruppen $(A, *), (B, \odot), (C, \bullet)$ ergibt die Komposition $\psi \circ \varphi: A \rightarrow C, a \mapsto \psi(\varphi(a))$ immer einen Homomorphismus:

$$\begin{aligned} (\psi \circ \varphi)(a * b) &= \psi(\varphi(a * b)) = \psi(\varphi(a) \odot \varphi(b)) = \psi(\varphi(a)) \bullet \psi(\varphi(b)) \\ &= (\psi \circ \varphi)(a) \bullet (\psi \circ \varphi)(b). \end{aligned}$$

Die Hintereinanderausführung/Komposition von Homomorphismen (allg. Funktionen) lässt sich gut über ein kommutierendes Diagramm darstellen:

$$\begin{array}{ccc} (A, *) & \xrightarrow{\psi \circ \varphi} & (C, \bullet) \\ & \searrow \varphi & \nearrow \psi \\ & (B, \odot) & \end{array}$$

Man sagt, „das Diagramm kommutiert“, wenn für alle unterschiedlichen gerichteten Wege von einer Menge/Gruppe zur anderen die Komposition der entsprechenden Funktionen immer die gleiche Abbildung ergibt.

Für $A = B = C$ sind φ, ψ und $\psi \circ \varphi$ Endomorphismen in $\text{Abb}(A, A)$ und mit id_A bilden die Endomorphismen ein Untermonoid von $(\text{Abb}(A, A), \circ)$.

Beschränken wir uns auf bijektive Funktionen, also auf die symmetrische Gruppe $(S(A), \circ)$, bilden die Automorphismen davon ein Untergruppe:

Beobachtung 2.11 Sei $(A, *)$ eine Gruppe, dann ist

$$\text{Aut}(A) := \{\varphi: A \rightarrow A \text{ Automorphismus}\}$$

mit der Komposition \circ eine Gruppe, die Automorphismengruppe zu A .

Beweis: s. Vorlesung □

Die Automorphismengruppe $\text{Aut}(A)$ ist also eine Untergruppe der Symmetrischen Gruppe $(S(A), \circ)$.

Definition 2.12 Zwei Gruppen $(A, *)$, (B, \odot) heißen isomorph, in Zeichen $(A, *) \simeq (B, \odot)$, wenn es einen Gruppenisomorphismus $\varphi: A \rightarrow B$ gibt.

In Hinblick auf die algebraische Gruppenstruktur sind zwei isomorphe Gruppen einfach identisch, sie tun das gleiche.

Beispiele:

- $(\mathcal{D}, \circ) \simeq (\mathcal{C}, \circ) \simeq (\mathbb{Z}/_{12}\mathbb{Z}, +)$. Insbesondere sind also ρ und γ unterschiedliche Darstellungen der, bis auf Isomorphie, gleichen Gruppe $\mathbb{Z}/_{12}\mathbb{Z}$.
- Für Mengen A und B gleicher Kardinalität gilt $(S(A), \circ) \simeq (S(B), \circ)$, denn ist $f: A \rightarrow B$ bijektiv, dann ist folgendes θ so ein Isomorphismus,

$$\theta: S(A) \rightarrow S(B)$$

$$g \mapsto h = f \circ g \circ f^{-1}$$

$$\left[\begin{array}{l} h: B \rightarrow B \\ b \mapsto f(g(f^{-1}(b))) \end{array} \right]$$

$$\begin{array}{ccc} A & \xrightarrow{g} & A \\ f^{-1} \uparrow & & \downarrow f \\ B & \xrightarrow{h=f \circ g \circ f^{-1}} & B \end{array}$$

Ü: Zeige $\theta(g \circ g') = \theta(g) \circ \theta(g')$ und θ ist bijektiv. ⊗

Insbesondere gilt: Für jede endliche Menge A mit $|A| = n \in \mathbb{N}$ ist

$$(S(A), \circ) \simeq (S_n = S(\{1, \dots, n\}), \circ) \text{ Permutationsgruppe.}$$

[Auf endlichen Mengen darf man sich jede bijektive Abbildung als Permutation vorstellen, und die Komposition bijektiver Abbildungen als Hintereinanderausführung von Permutationen.]

- Jede Gruppe $(A, *)$ ist isomorph zu einer Untergruppe der symmetrischen Gruppe $(S(A), \circ)$ über den Homomorphismus $\varphi: A \rightarrow S(A)$, $a \mapsto {}_a\tau$ (**Ü**). Insbesondere ist damit jede endliche Gruppe der Ordnung n isomorph zu einer Untergruppe der Permutationsgruppe S_n .

♡

Beobachtung 2.13 Für einen Gruppenhomomorphismus $\varphi: A \rightarrow B$ ist $\text{Ker } \varphi$ stets ein Normalteiler von A .

Beweis: s. Vorlesung □

$a * \text{Ker } \varphi$ ist die Restklasse von $a \pmod{\text{Ker } \varphi}$ und

$a, b \in A$ haben das gleiche Bild $\varphi(a) = \varphi(b)$ (gehören zur gleichen Faser)

genau dann, wenn sie zur gleichen Restklasse $a * \text{Ker } \varphi = b * \text{Ker } \varphi$ gehören:

$$\varphi(a) = \varphi(b) \Leftrightarrow \varphi(b)^{-1} \odot \varphi(a) = e_B \Leftrightarrow \varphi(b^{-1} * a) = e_B$$

$$\Leftrightarrow (b^{-1} * a) \in \text{Ker } \varphi \stackrel{S_{2,6}}{\Leftrightarrow} b * \text{Ker } \varphi = a * \text{Ker } \varphi$$

Also: Für Gruppenhomomorphismen φ hat jede Faser die Gestalt $a * \text{Ker } \varphi$ für ein beliebiges a der Faser und die Mächtigkeit $|\text{Ker } \varphi|$. Insbesondere ist φ genau dann injektiv, wenn $\text{Ker } \varphi = \{e_A\}$.

Man kann φ als „aufgeblasene“ Version eines auf der Faktorgruppe $A/\text{Ker } \varphi$ erklärten Monomorphismus (injektiver Hom.)

$$\bar{\varphi}: A/\text{Ker } \varphi \rightarrow B, a * \text{Ker } \varphi \mapsto \varphi(a)$$

auffassen, der jeder Nebenklasse ihren gemeinsamen Wert zuordnet. Anders ausgedrückt, kennt man $\text{Ker } \varphi$ und für jede Restklasse $a * \text{Ker } \varphi$ den Wert $\varphi(b)$ für einen Repräsentanten $b \in a * \text{Ker } \varphi$, kann man φ vollständig rekonstruieren.

Ü: Für das Schiebispiel ist $\text{Ker } \gamma = \{\cup^{3n}: n \in \mathbb{Z}\} = U$ wie im Beispiel zur Quotientengruppe, also $\mathcal{C}/\text{Ker } \varphi = \mathcal{C}/U$ wie dort. Für $a \in \mathcal{C}$ könne wir aus $\gamma(a)$ nur noch ablesen, zu welcher Nebenklasse aus \mathcal{C}/U das a gehört hat. Überprüfe, dass $\hat{\gamma}: \mathcal{C}/U \rightarrow S_3, (a \circ U) \mapsto \gamma(a)$ ein Monomorphismus ist. \times

Die Permutationsgruppe ($S_n = S(\{1, \dots, n\}), \circ$)

Eine Permutation $\sigma \in S_n$ heißt Zyklus der Länge $r \geq 1$, falls es paarweise verschiedene Zahlen $a_1, \dots, a_r \in \{1, \dots, n\}$ gibt mit $\sigma(a_1) = a_2, \sigma(a_2) = a_3, \dots, \sigma(a_r) = a_1$ und $\sigma(i) = i$ sonst.

Man schreibt σ dann auch kurz $(a_1 a_2 \dots a_r)$.

Zwei Zyklen $\sigma = (a_1 \dots a_r), \sigma' = (a'_1 \dots a'_k) \in S_n$ heißen disjunkt, falls $\{a_1, \dots, a_r\} \cap \{a'_1, \dots, a'_k\} = \emptyset$.

Die Komposition disjunkter Zyklen ist kommutativ: $\sigma \circ \sigma' = \sigma' \circ \sigma$.

Jeder Zyklus der Länge 1 (etwa (a_1)) ist gleich der Identität $\text{id}_{\{1, \dots, n\}}$.

Satz 2.14 *Jede Permutation $f \in S_n$ lässt sich als Komposition disjunkter Zyklen schreiben. Bis auf die Reihenfolge der Zyklen und die „erste Zahl“ in jedem Zyklus ist die Darstellung eindeutig.*

Beispiel: $f: \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 1 & 4 & 3 & 2 \end{pmatrix} \longleftrightarrow (1 \ 5 \ 3)(2 \ 6)(4)$ ♡

Beweis: s. Vorlesung □

Jedes i mit $i = \sigma(i)$ heißt Fixpunkt.

Ein Zyklus der Länge 2 heißt Transposition.

Eine Transposition $\tau = (a \ b) \in S_n$ vertauscht nur a und b ,
 $\tau \circ \tau = (a \ b) \circ (a \ b) = \text{id}_{\{1, \dots, n\}}$.

Ü: $\text{sgn}(\tau) = -1$ (ungerade Anzahl an Paaren $i < j$ mit $\tau(i) > \tau(j)$)
 \times

Jeder Zyklus $(a_1 \dots a_r)$ mit $r \geq 2$ lässt sich als Komposition von $r - 1$ (i. A. nicht disjunkten) Transpositionen darstellen,

$$(a_1 \dots a_r) = (a_1 a_2) \circ (a_2 a_3) \circ \dots \circ (a_{r-1} a_r).$$

Wir haben schon beobachtet (bei den Funktionen), dass

$$\text{für } f, g \in S_n \quad \text{sgn}(f \circ g) = \text{sgn}(f) \cdot \text{sgn}(g) \quad \begin{array}{c|c|c} \cdot & +1 & -1 \\ \hline +1 & +1 & -1 \\ \hline -1 & -1 & +1 \end{array},$$

also ist $\text{sgn}: S_n \rightarrow \{+1, -1\}$ ein Gruppenhomomorphismus, und es gilt $f \mapsto \text{sgn}(f)$

$$\text{sgn}(f) = \begin{cases} +1 & f \text{ ist Komposition geradzahlig vieler} \\ -1 & f \text{ ist Komposition ungeradzahlig vieler} \end{cases} \text{ Transpositionen.}$$

$\text{Ker sgn} = \{f \in S_n : \text{sgn}(f) = +1\}$ ist ein Normalteiler von S_n und heißt alternierende Gruppe $A_n = \{f \in S_n : \text{sgn}(f) = +1\}$

Ü:

- Für das Schiebspiel gilt $A_3 = \text{Im } \gamma$.
- Zeige: A_4 ist die Gruppe der Permutationen der Ecken eines regelmäßigen Tetraeders bei Drehung um
 - eine der 4 Drehachsen durch eine Ecke und den Mittelpunkt der gegenüberliegenden Seite,
 - eine der 3 Drehachsen durch die Mittelpunkte gegenüberliegender Kanten.

⊠

2.3 Ringe

Ab nun geht es um Mengen mit zwei Verknüpfungen $+$ und \cdot , wobei die „Addition“ $+$ eine abelsche Gruppe, die „Multiplikation“ \cdot mindestens eine Halbgruppe induziert und beide „sich vertragen“ müssen.

Definition 2.15 $(R, +, \cdot)$ bestehend aus einer Menge R und Verknüpfungen

$$\begin{array}{ll} +: R \times R \rightarrow R & \cdot: R \times R \rightarrow R \\ (a, b) \mapsto a + b & (a, b) \mapsto a \cdot b \end{array}$$

heißt Ring, falls gilt:

- (R1) $(R, +)$ ist abelsche Gruppe,
- (R2) (R, \cdot) ist eine Halbgruppe,

(R3) $+$ und \cdot erfüllen die Distributivgesetze: für alle $a, b, c \in R$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Für die Klammerung bindet \cdot stärker als $+$.

In einem Ring wird das neutrale Element von $(R, +)$ Nullelement genannt und meist mit 0 bezeichnet.

Ein Ring heißt kommutativ, falls $a \cdot b = b \cdot a$ für alle $a, b \in R$ gilt.

Ist (R, \cdot) ein Monoid, nennt man das neutrale Element zu \cdot das Einselement, in Zeichen $1 \in R$, und spricht von einem Ring mit Eins. (Für $R = \{0\}$ ist $0 = 1!$)

Beobachtung 2.16 Ist $(R, +, \cdot)$ ein Ring und 0 das Nullelement, dann gilt $0 \cdot a = a \cdot 0 = 0$ für alle $a \in R$.

Beweis: s. Vorlesung □

Beispiele:

- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$
- Für Restklassen $\bar{a}, \bar{b} \in \mathbb{Z}/m\mathbb{Z}$ (modulo $m \in \mathbb{N}$) haben wir schon gezeigt, dass $\bar{a} + \bar{b} = \overline{a + b}$ eine abelsche Gruppe, die „zyklische Gruppe der Ordnung m “, definiert.

Multiplikation: $\cdot : \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z}$
 $(\bar{a}, \bar{b}) \mapsto \bar{a} \cdot \bar{b} := \overline{a \cdot b}$

Wohldefiniertheit: Zu zeigen ist nur $a' \cdot b' \in \overline{a \cdot b}$, denn daraus folgt sofort $\overline{a \cdot b} = \overline{a' \cdot b'}$, weil $\mathbb{Z}/m\mathbb{Z}$ eine Klasseneinteilung ist.

Für $a' \in \bar{a}$ und $b' \in \bar{b}$ gibt es $k, h \in \mathbb{Z}$ mit

$$a' = a + mk \quad \text{und} \quad b' = b + mh,$$

$$a' \cdot b' = (a + mh)(b + mk) = ab + m(bk + ah + mkh) \in \overline{a \cdot b}.$$

Z. B. für $(\mathbb{Z}/4\mathbb{Z}, +, \cdot)$ ergeben sich die (kommutativen) Rechentabellen

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

\cdot	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Beachte: Das Produkt von zwei Elementen $\neq 0$ kann 0 werden.

- Seien $(R, +, \cdot)$ ein Ring mit Eins und $m, n, k \in \mathbb{N}_0$. Für $A, B \in M(m \times n; R) = R^{m \times n}$ (Matrizen mit m Zeilen und n Spalten) erklären wir eine Addition

$$\begin{aligned} +: R^{m \times n} \times R^{m \times n} &\rightarrow R^{m \times n} \\ (A, B) &\mapsto (A + B)_{ij} := A_{ij} + B_{ij} \text{ für } (i, j) \in m \times n \end{aligned}$$

und für $A \in R^{m \times k}$, $B \in R^{k \times n}$ (Spaltenzahl der ersten und Zeilenzahl der zweiten stimmen überein!) eine Multiplikation

$$\begin{aligned} \cdot: R^{m \times k} \times R^{k \times n} &\rightarrow R^{m \times n} \\ (A, B) &\mapsto (A \cdot B)_{ij} := \sum_{h=1}^k A_{ih} \cdot B_{hj} \text{ für } (i, j) \in m \times n \end{aligned}$$

Vorsicht! Weil (R, \cdot) i. A. nicht kommutativ ist, ist i. A. auch $A_{ih} \cdot B_{hj} \neq B_{hj} \cdot A_{ih}$ und die Reihenfolge ist selbst in der Summe wichtig! ∇

Ein konkretes Beispiel für die Addition bei $R = \mathbb{Z}$, $m = 2$, $n = 3$ wäre

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 0 & -3 & 2 \end{bmatrix}, B = \begin{bmatrix} 7 & -1 & 4 \\ -2 & 0 & 6 \end{bmatrix}, A + B = \begin{bmatrix} 8 & 4 & 4 \\ -2 & -3 & 8 \end{bmatrix},$$

und für die Multiplikation mit $m = 2$, $n = 2$, $k = 3$ (mit gleichem A)

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 0 & -3 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & 6 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, A \cdot B = \begin{bmatrix} -3 & 6 \\ 3 & -2 \end{bmatrix}.$$

Um das Element $(A \cdot B)_{ij}$ zu berechnen, nimmt man also die „ i -te Zeile“ $A_{i,\bullet}$ von A und die „ j -te Spalte“ $B_{\bullet,j}$ von B (beide sind gleich „lang“, \bullet steht dabei jeweils für alle möglichen Werte), multipliziert die Einträge elementweise und summiert,

$$(A \cdot B)_{11} = A_{1,\bullet} \cdot B_{\bullet,1} = \begin{bmatrix} 1 & 5 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = 1 \cdot 2 + 5 \cdot (-1) + 0 \cdot 0 = -3,$$

$$(A \cdot B)_{21} = A_{2,\bullet} \cdot B_{\bullet,1} = \begin{bmatrix} 0 & -3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = 0 \cdot 2 + (-3) \cdot (-1) + 2 \cdot 0 = 3,$$

etc.

Man darf die Berechnung gerne auch anders organisieren (und das geschieht auch in Computerprogrammen), etwa indem man gleich die gesamte erste Spalte so bestimmt

$$\begin{aligned} (A \cdot B)_{\bullet,1} &= A_{\bullet,1} \cdot B_{11} + A_{\bullet,2} \cdot B_{2,1} + A_{\bullet,3} \cdot B_{3,1} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot 2 + \begin{bmatrix} 5 \\ -3 \end{bmatrix} \cdot (-1) + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot 0 = \begin{bmatrix} -3 \\ 3 \end{bmatrix}. \end{aligned}$$

Oft ist es nützlich, sich die Matrixmultiplikation gleich spalten- oder zeilenweise vorzustellen. Für allgemeine $m, n, k \in \mathbb{N}$ ist die

$$\text{spaltenweise Sicht: } (A \cdot B)_{\bullet,j} = \sum_{h=1}^k A_{\bullet,h} B_{h,j} \quad (j = 1, \dots, n).$$

Ist etwa in der Spalte $B_{\bullet,j}$ nur ein einziger Eintrag B_{ij} ungleich Null, bedeutet dies für die j -te Spalte $(A \cdot B)_{\bullet,j} = A_{\bullet,i} B_{ij}$, und ist zusätzlich $B_{ij} = 1$, kann man die i -te Spalte von A einfach als j -te Spalte des Ergebnisses übernehmen.

Ü: Zeige die Distributivgesetze: Für $A \in R^{m \times k}, B, C \in R^{k \times n}$ ist $A \cdot (B + C) = A \cdot B + A \cdot C$, und für $A, B \in R^{m \times k}, C \in R^{k \times n}$ ist $(A + B) \cdot C = A \cdot C + B \cdot C$. x

Wir überlegen uns nun, dass für $n \times n$, also sogenannte quadratische Matrizen, mit dieser Addition und Multiplikation $(R^{n \times n}, +, \cdot)$ selbst ein Ring mit Eins ist.

$(R^{n \times n}, +)$ ist abelsche Gruppe: Da $(R, +)$ eine abelsche Gruppe ist, ist auch die Matrixaddition assoziativ und kommutativ, hat die Nullmatrix (alle Einträge sind Null) als Nullelement und zu $A \in R^{n \times n}$ ist die Matrix $-A = (-A_{ij})$ das Inverse bezüglich der Addition.

$(R^{n \times n}, \cdot)$ ist ein (nicht-kommutatives) Monoid: Assoziativität, also $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ gilt sogar allgemein für die Matrixmultiplikation, ist aber mit derzeitigen Mitteln eine mühselige Indexschlacht (\rightarrow Ü). Das Einselement ist nach den obigen Überlegungen die Matrix

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}, \text{ sie wird } \underline{\text{Identität}} \text{ genannt.}$$

Mathematisch prägnant formuliert man sie mit dem

$$\underline{\text{Kroneckersymbol}} \quad \delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst.} \end{cases}$$

Die Identität I hat die Einträge $I_{ij} = \delta_{ij}$ für $i, j \in \{1, \dots, n\}$. Will man die Abmessungen von I betonen, schreibt man auch I_n .

Schließlich gelten die Distributivgesetze auf Grund der vorigen Übung. Beachte: Der Ring $(R^{n \times n}, +, \cdot)$ ist i. A. nicht kommutativ und das Produkt von zwei Matrizen kann die Nullmatrix sein:

$$A = \begin{bmatrix} 4 & -2 \\ 6 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}, \quad A \cdot B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B \cdot A = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}$$



Ein Ring heißt nullteilerfrei, wenn für alle $a, b \in R$

$$a \cdot b = 0 \quad \Rightarrow \quad a = 0 \vee b = 0.$$

Ein nullteilerfreier, kommutativer Ring mit Eins heißt Integritätsbereich.

Beobachtung 2.17 $\mathbb{Z}/m\mathbb{Z}$ ist genau dann Integritätsbereich, wenn m eine Primzahl ist.

Beweis: s. Vorlesung □

Eine Teilmenge $R' \subseteq R$ heißt Unterring eines Rings $(R, +, \cdot)$, falls $(R', +)$ Untergruppe von $(R, +)$ und (R', \cdot) Unterhalbgruppe von (R, \cdot) ist.

Ü: Ein Unterring ist ein Ring. ⊗

Für Ringe $(R, +, \cdot)$, (S, \oplus, \odot) heißt eine Abbildung $\varphi: R \rightarrow S$ (Ring-)Homomorphismus, wenn für alle $a, b \in R$ gilt:

$$\varphi(a + b) = \varphi(a) \oplus \varphi(b) \quad \text{und} \quad \varphi(a \cdot b) = \varphi(a) \odot \varphi(b).$$

Beispiele:

- $(m\mathbb{Z}, +, \cdot)$ ist Unterring von $(\mathbb{Z}, +, \cdot)$,
 $(\mathbb{Q}, +, \cdot)$ ist Unterring von $(\mathbb{R}, +, \cdot)$,
 $\varphi: \mathbb{Z} \rightarrow m\mathbb{Z}$, $z \mapsto m \cdot z$ ist ein Ringhomomorphismus.
- Ist M eine Menge und (R, \oplus, \odot) ein Ring, so ist die Menge der Funktionen $f \in R^M = \text{Abb}(M, R)$ mit
 punktwiser Addition $+$: $(f + g)(a) := f(a) \oplus g(a)$, $a \in M$
 punktwiser Multiplikation \cdot : $(f \cdot g)(a) := f(a) \odot g(a)$, $a \in M$
 ein Ring. Beweis: **Ü**.

Z.B. ist $(\text{Abb}(\mathbb{R}, \mathbb{R}), +, \cdot)$ ein Ring und die Menge der stetigen Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$, „ $f \in C(\mathbb{R})$ “, ein Unterring. Konkret ist etwa $(\cos \cdot \cos + \sin \cdot \sin)(x) = \cos(x)^2 + \sin(x)^2$. Ist $(C(\mathbb{R}), +, \cdot)$ nullteilerfrei?



2.4 Körper

In einem Integritätsbereich $(R, +, \cdot)$ bleibt zwar die Multiplikation auf $R \setminus \{0\}$ abgeschlossen, aber es muss noch kein multiplikatives Inverses geben, z. B. $(\mathbb{Z}, +, \cdot)$. Für einen Körper muss nun auch $(R \setminus \{0\}, \cdot)$ eine abelsche Gruppe sein. [Wegen der Bedeutung, explizit:]

Definition 2.18 $(K, +, \cdot)$ bestehend aus einer Menge K und zwei Verknüpfungen

$$\begin{aligned} +: K \times K &\rightarrow K & \cdot: K \times K &\rightarrow K \\ (a, b) &\mapsto a + b & (a, b) &\mapsto a \cdot b \end{aligned}$$

heißt Körper [engl. field], falls gilt:

- (K1) $(K, +)$ ist abelsche Gruppe, ihr neutrales Element wird mit 0 , das zu $a \in K$ inverse Element mit $-a$ bezeichnet.
- (K2) $(K \setminus \{0\}, \cdot)$ ist abelsche Gruppe, ihr neutrales Element wird mit 1 , das zu $a \in K \setminus \{0\}$ inverse Element mit a^{-1} oder $\frac{1}{a}$ bezeichnet.
- (K3) $+$ und \cdot erfüllen die Distributivgesetze: für alle $a, b, c \in R$

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Für die Klammerung bindet \cdot stärker als $+$.

Rechenregeln 2.19 Für einen Körper K gilt für $a, b, x, \hat{x} \in K$:

- (a) $1 \neq 0$ (also ist $|K| \geq 2$),
- (b) $0 \cdot a = a \cdot 0 = 0$,
- (c) $a \cdot b = 0 \Rightarrow a = 0 \vee b = 0$,
- (d) $(-a)b = -(ab)$ und $(-a)(-b) = ab$,
- (e) $xa = \hat{x}a \wedge a \neq 0 \Rightarrow x = \hat{x}$.

Beweis: s. Vorlesung □

Beobachtung 2.20 Ein Integritätsbereich $(K, +, \cdot)$ (nullteilerfreier kommutativer Ring mit Eins) mit endlich vielen Elementen ist ein Körper.

Beweis: s. Vorlesung □

Beispiel: Für p Primzahl ist $\mathbb{F}_p := (\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ ein endlicher Integritätsbereich, also ein Körper. Übrigens ist $\mathbb{F}_2 \simeq (\{0, 1\}, +, \cdot)$ mit der bereits definierten Addition und Multiplikation. ♡

Die Ringe \mathbb{Z} und $\mathbb{Z}/m\mathbb{Z}$ unterscheiden sich beim Aufsummieren der 1:

Für \mathbb{Z} und $n \in \mathbb{N}$ gilt $n \cdot 1 := 1 + \dots + 1 = n \neq 0$.

Für $\mathbb{Z}/m\mathbb{Z}$ ist $m \cdot \bar{1} := \bar{1} + \dots + \bar{1} = \bar{m} = 0$.

Ist $(R, +, \cdot)$ ein Ring mit Eins 1 , definiert man die Charakteristik des Rings

$$\text{char}(R) := \begin{cases} 0 & \text{falls } n \cdot 1 \neq 0 \text{ für alle } n \in \mathbb{N}, \\ \min\{n \in \mathbb{N} : n \cdot 1 = 0\} & \text{sonst.} \end{cases}$$

Lemma 2.21 Ist K ein Körper, so ist $\text{char}(K)$ entweder 0 oder eine Primzahl.

Beweis: s. Vorlesung □

Für Primzahl p nennt man \mathbb{F}_p auch den Primkörper der Charakteristik p .

Beispiele:

- $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper.
- $\mathbb{C} = (\mathbb{R}^2, +, \cdot)$ mit $(a, b) + (c, d) := (a + c, b + d)$
 $(a, b) \cdot (c, d) := (ac - bd, ad + bc)$

ist Körper: Nullelement $(0, 0)$, additives Inverses $-(a, b) = (-a, -b)$,
 Einselement $(1, 0)$, für $(a, b) \neq 0$ ist $(a, b)^{-1} = (\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2})$ (Ü).

Beachte: $(\{(a, 0) : a \in \mathbb{R}\}, +, \cdot)$ ist ein zu $(\mathbb{R}, +, \cdot)$ isomorpher Unterkörper.

Statt $z = (a, b)$ schreibt man auch $z = a + ib = (a, 0) + (0, b)$

imaginäre Einheit $i := (0, 1)$, also ist $i^2 = -1 = (-1, 0)$

Realteil $\operatorname{re} z = a$

Imaginärteil $\operatorname{im} z = b$

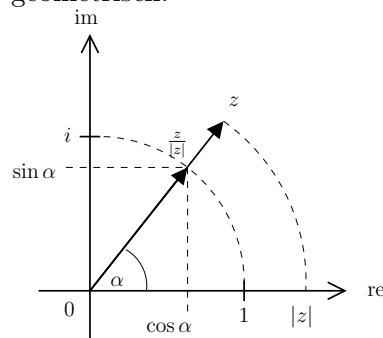
konjugiert komplexe Zahl $\bar{z} := a - ib = (a, -b)$

Insbesondere ist $z + \bar{z} \in \mathbb{R}$ und $0 \leq z \cdot \bar{z} = (a^2 + b^2, 0)$.

Der Absolutbetrag $|z| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$ erfüllt

für alle $z, z' \in \mathbb{C}$ $|z + z'| \leq |z| + |z'|$, die Dreiecksungleichung,
 und $|z \cdot z'| = |z| \cdot |z'|$.

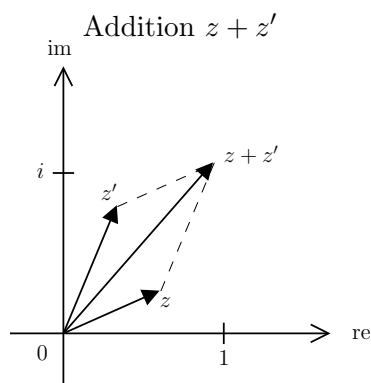
geometrisch:



aus der Analysis:

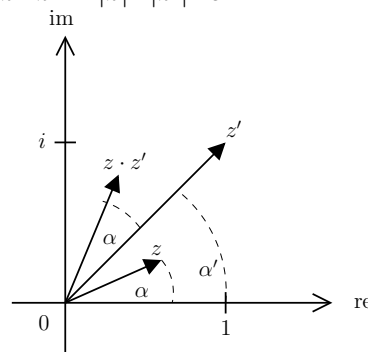
$$\frac{z}{|z|} = \cos \alpha + i \sin \alpha = e^{i\alpha},$$

wobei $\arg z = \alpha$ Argument von z .



Multiplikation: Drehstreckung

$$z \cdot z' = |z| \cdot |z'| \cdot e^{i(\arg z + \arg z')}$$



 Lineare Gleichungssysteme über Körpern und deren Lösung

Ein lineares Gleichungssystem über einem Körper K (oder nur Ring) mit $m \in \mathbb{N}_0$ Gleichungen in $n \in \mathbb{N}_0$ Variablen $x_1, \dots, x_n \in K$ hat die Form,

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m, \end{array}$$

wobei $a_{ij} \in K$ für $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$ gegebene Koeffizienten und $b_i \in K$ für $i = 1, \dots, m$ die gegebene rechte Seite sind; mit $A = (a_{ij})$, $x = (x_i)$ und $b = (b_i)$ schreibt man kurz $Ax = b$. Man sucht die Lösungsmenge $\text{Lös}(A, b) := \{x \in K^n : Ax = b\}$, also alle x , für die die Gleichung $Ax = b$ erfüllt wird.

Beispiel: $K = \mathbb{Q}$, $n = 6$, $m = 3$,

$$\begin{array}{rcccccc} & & & - & x_4 & + & \frac{5}{7}x_6 & = & -4 \\ 2x_2 & - & \frac{2}{3}x_3 & + & \frac{1}{5}x_4 & & & = & 6 \\ 4x_2 & - & \frac{4}{3}x_3 & + & \frac{12}{5}x_4 & - & x_6 & = & \frac{88}{5} \end{array}$$

$$A = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & \frac{5}{7} \\ 0 & 2 & -\frac{2}{3} & \frac{1}{5} & 0 & 0 \\ 0 & 4 & -\frac{4}{3} & \frac{12}{5} & 0 & -1 \end{bmatrix}, \quad b = \begin{pmatrix} -4 \\ 6 \\ \frac{88}{5} \end{pmatrix}$$

♥

Die Lösungsmenge bestimmt ein sehr einfacher, sehr alter Algorithmus (= Rechenrezept, "al-Chwarizmi"), der historisch schon im 2. Jh.v.Chr. in China schriftlich erschien, belegt durch „Neun Kapitel der Mathematischen Künste“, (Zhan Cang, 256–125 v.Chr.), die von Liu Hui (225–295 n.Chr.) neu aufgelegt wurden, bei uns bekannt als Gauß-Algorithmus.

Grundlage: Folgende Operationen ändern nicht die Lösungsmenge eines Gleichungssystems:

- (O1) Addieren eines Vielfachen einer Zeile zu einer anderen.
- (O2) Vertauschen von zwei Zeilen.

Idee: Nutze (O1) und (O2), um das System in „Zeilenstufenform“ zu bringen, um es leichter lösen zu können:

$$\left[\begin{array}{cccc|cc} 0 & \dots & 0 & & * & * \\ & & & \dots & 0 & * & * \\ & & & & & \ddots & \\ & & & & \dots & 0 & * & * \\ & & & & & & \dots & 0 \\ & & & & & & & \vdots \\ & & & & & & & \dots & 0 \end{array} \right]$$

Definition 2.22 Eine $m \times n$ Matrix A ist in Zeilenstufenform, wenn es Spaltenindices $j_1 < j_2 < \dots < j_r \leq n$ mit $r \leq m$ gibt, sodass

(ZS1) in Zeile i mit $i \leq r$ $A_{ij} = 0$ für $j = 1, \dots, j_i - 1$ und $A_{ij_i} \neq 0$,

(ZS2) in Zeile i mit $i > r$ $A_{ij} = 0$ für $j = 1, \dots, n$.

r heißt Zeilenrang der Matrix.

Beispiel: Bringt man das obige Beispiel in Zeilenstufenform (wie, s. unten), ist die Lösung sofort klar:

$$\begin{bmatrix} 0 & 2 & -\frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & \frac{5}{7} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_6 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \\ 0 \end{pmatrix} \rightarrow \begin{array}{l} \text{ZSF: } j_1 = 2, j_2 = 4, r = 2 \\ \text{Zeile}_2: x_4 = 4 + \frac{5}{7}x_6 \\ \text{in Zeile}_1: x_2 = 2 - \frac{5}{28}x_6 + \frac{2}{6}x_3 \end{array}$$

Nur $r = 2$ Gleichungen schränken $x \in \mathbb{Q}^6$ ein. Die $n - r = 4$ Variablen x_1, x_3, x_5, x_6 dürfen frei gewählt werden, es sind freie Variable. Jede Wahl der freien Variablen führt zu einer Lösung, in der x_2 und x_4 dadurch bestimmt sind. x_2 und x_4 sind gebundene Variable. Für dieses \hat{A} und rechte Seite \hat{b} gilt

$$\text{Lös}(\hat{A}, \hat{b}) = \left\{ \begin{pmatrix} x_1 \\ 2 - \frac{5}{28}x_6 + \frac{2}{6}x_3 \\ x_3 \\ 4 + \frac{5}{7}x_6 \\ x_5 \\ x_6 \end{pmatrix} : x_1, x_3, x_5, x_6 \in \mathbb{Q} \right\}.$$

Ist auf der rechten Seite in einer Zeile $> r$ ein Eintrag $\neq 0$, ist die Lösung auch sofort klar, wie etwa für

$$\begin{bmatrix} 0 & | 2 & -\frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & | -1 & 0 & \frac{5}{7} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_6 \end{pmatrix} = \begin{pmatrix} 6 \\ -4 \\ 5 \end{pmatrix} \leftarrow \sum_{j=1}^6 0 \cdot x_j = 5 \text{ nicht lösbar!}$$

Für dieses \hat{A} und \tilde{b} gilt $\text{Lös}(\hat{A}, \tilde{b}) = \emptyset$. Die Lösbarkeit (=Existenz von Lösungen) erfordert $b_i = 0$ für $r < i \leq m$ ♥

Ist A erst in Zeilenstufenform, bestimmt ganz allgemein der folgende Algorithmus ob es eine Lösung gibt, und in diesem Fall diejenige, in der alle freien Variablen den Wert 0 haben:

Algorithmus 2.23 (zur Lösung von $Ax = b$ in Zeilenstufenform)

Eingabe(Input): Körper K , Matrix $A \in K^{m \times n}$ in Zeilenstufenform mit $j_1 < j_2 < \dots < j_r < n + 1 =: j_{r+1}$ und rechte Seite $b \in K^m$.

Ausgabe(Output): ein $x \in \text{Lös}(A, b)$ oder der Nachweis $\text{Lös}(A, b) = \emptyset$.

1. Falls $b_i \neq 0$ für ein $r < i \leq m$, STOP, es gibt keine Lösung.
2. Für $k = r, r - 1, \dots, 1$ führe folgendes aus:
 - (a) Für $j = j_k + 1, j_k + 2, \dots, j_{k+1} - 1$ setze $x_j \leftarrow 0$.
 - (b) Setze $x_{j_k} \leftarrow (A_{kj_k})^{-1} \cdot (b_k - \sum_{j=j_k+1}^n A_{kj} x_j)$.

Beachte: Hier wird das multiplikative Inverse benötigt!

Wie bringt man ein Gleichungssystem $Ax = b$ durch (O1) and (O2) in Zeilenstufenform?

Beispiel: Zur Illustration des Vorgehens wandeln wir die Matrix und rechte Seite des Anfangsbeispiels gemeinsam schrittweise in Zeilenstufenform um und vergeben bei jeder Anwendung von (O1) eine neue Nummer (das Vertauschen der Zeilen (O2) ändert ja nicht wirklich etwas am System).

$$A = \begin{bmatrix} \overline{0} & 0 & 0 & -1 & 0 & \frac{5}{7} \\ 0 & \boxed{2} & -\frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 4 & -\frac{4}{3} & \frac{12}{5} & 0 & -1 \end{bmatrix}, \quad b = \begin{pmatrix} -4 \\ 6 \\ \frac{88}{5} \end{pmatrix} \quad \begin{array}{l} \text{finde erste } \neq 0 \text{ Spalte,} \\ \text{tausche Zeile}_2 \text{ mit Zeile}_1. \end{array}$$

In der noch interessierenden Restmatrix (anfangs noch die ganze), suche die erste Spalte $\neq 0$ und vertausche mit (O2) die erste $\neq 0$ Zeile mit der ersten der Restmatrix (auch in b):

$$A^{(0)} = \begin{bmatrix} 0 & \overline{2} & -\frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & \frac{5}{7} \\ 0 & 4 & -\frac{4}{3} & \frac{12}{5} & 0 & -1 \end{bmatrix}, \quad b = \begin{pmatrix} 6 \\ -4 \\ \frac{88}{5} \end{pmatrix} \quad \begin{array}{l} \text{Zeile}_2: \text{ nichts zu tun,} \\ \text{Zeile}_3 - 4 \cdot 2^{-1} \cdot \text{Zeile}_1. \end{array}$$

Nutze (O1) mit diesem ersten $\neq 0$ -Element, um alle Einträge darunter zu Null zu machen:

$$A^{(1)} = \begin{bmatrix} 0 & 2 & -\frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \overline{0} & \boxed{-1} & 0 & \frac{5}{7} \\ 0 & 0 & 0 & \frac{7}{5} & 0 & -1 \end{bmatrix}, \quad b = \begin{pmatrix} 6 \\ -4 \\ \frac{28}{5} \end{pmatrix} \quad \begin{array}{l} \text{kein (O2) nötig,} \\ \text{Zeile}_3 - \frac{7}{5} \cdot (-1)^{-1} \cdot \text{Zeile}_2. \end{array}$$

Rekursives Fortsetzen auf der Untermatrix liefert:

$$A^{(2)} = \begin{bmatrix} 0 & 2 & -\frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \boxed{-1} & 0 & \frac{5}{7} \\ 0 & 0 & 0 & 0 & \overline{0} & 0 \end{bmatrix}, \quad b = \begin{pmatrix} 6 \\ -4 \\ 0 \end{pmatrix} \quad \text{nichts zu tun.}$$

♡

Allgemein:

Algorithmus 2.24 (bringt $Ax = b$ in Zeilenstufenform)

Eingabe(Input): Körper K , Matrix $A \in K^{m \times n}$ und rechte Seite $b \in K^m$.

Ausgabe(Output): Äquivalentes Gleichungssystem $A^{(r)}x = b^{(r)}$ in Zeilenstufenform mit Stufenindices $j_1 < j_2 < \dots < j_r \leq n$, $r \leq m$.

1. Setze $A^{(0)} \leftarrow A$, $b^{(0)} \leftarrow b$, $r \leftarrow 0$, $j_0 \leftarrow 0$.
2. Sei \hat{j} der kleinste Index mit $j_r < \hat{j} \leq n$, so dass

$$A_{\hat{i}\hat{j}} \neq 0 \text{ f\u00fcr ein } \hat{i} \text{ mit } r < \hat{i} \leq m.$$
 Gibt es kein solches Paar (\hat{i}, \hat{j}) , STOP,
 sonst tausche mit (O2) in $A^{(r)}$ und $b^{(r)}$ Zeile $_{r+1}$ gegen Zeile $_{\hat{i}}$,
 setze dann $r \leftarrow r + 1$ und $j_r \leftarrow \hat{j}$.
3. F\u00fcr $i = 1, \dots, r$ setze $A_{i,\bullet}^{(r)} \leftarrow A_{i,\bullet}^{(r-1)}$, $b_i^{(r)} \leftarrow b_i^{(r-1)}$
4. F\u00fcr $i = r + 1, \dots, m$ addiere α_i mal Zeile $_r$ zu Zeile $_i$ mit (O1):
 setze $\alpha_i \leftarrow A_{i,j_r}^{(r-1)} \cdot (A_{r,j_r}^{(r-1)})^{-1}$, $b_i^{(r)} \leftarrow b_i^{(r-1)} - \alpha_i \cdot b_r^{(r-1)}$,
 (a) f\u00fcr $j = j_r + 1, \dots, n$ setze $A_{i,j}^{(r)} \leftarrow A_{i,j}^{(r-1)} - \alpha_i A_{r,j}^{(r-1)}$,
 (b) f\u00fcr $j = 1, \dots, j_r$ setze $A_{i,j}^{(r)} \leftarrow 0$.
5. Gehe zu 2.

Auch dieser Algorithmus ben\u00f6tigt das multiplikative Inverse. \u00dcber K\u00f6rpern k\u00f6nnen (endliche) lineare Gleichungssysteme damit immer exakt gel\u00f6st werden oder es kann gezeigt werden, dass es keine gibt. Die strukturellen Eigenschaften der L\u00f6sungsmenge werden uns jedoch noch lange besch\u00e4ftigen.

Vorsicht! Da der Computer f\u00fcr $K = \mathbb{R}$ meist nicht exakt rechnet, wird in numerisch stabilen Umsetzungen das Paar (\hat{i}, \hat{j}) so in der Restmatrix gew\u00e4hlt, dass es das betragsgr\u00f6\u00dfteste ist, s. Vorlesung Numerik. Dazu werden dann jeweils Zeilen und Spalten (also auch die Variablen) umsortiert. ∇

\u00dcr: Es seien 4 Lampen $\{L_1, L_2, L_3, L_4\}$ mit 3 Schaltern $\{S_1, S_2, S_3\}$ schaltbar. Die Schalter \u00e4ndern jeweils auf einer Teilmenge der Lampen den Zustand jeder dieser Lampen. Seien die Mengen durch $S_1 = \{L_1, L_2\}$, $S_2 = \{L_2, L_3\}$ und $S_3 = \{L_1, L_4\}$ gegeben. Kann man die Schalter so bet\u00e4tigen, dass sich nachher nur der Zustand von Lampe L_4 ge\u00e4ndert hat?

Beschreibe ein allgemeines L\u00f6sungsverfahren f\u00fcr derartige Fragestellungen mit m Lampen und n Schaltern \u00fcber die L\u00f6sung von linearen Gleichungssystemen \u00fcber einem geeigneten K\u00f6rper. \times

Die Fragestellung, ob ein lineares Gleichungssystem \u00fcber einem Ring eine L\u00f6sung besitzt, l\u00e4sst sich \u00fcbrigens ebenfalls algorithmisch beantworten, wenn es in dem Ring ein geeignetes Konzept einer Division mit Rest gibt. Mit dieser kann man dann das multiplikative Inverse umgehen. Insbesondere gibt es solche Algorithmen f\u00fcr die Ringe $(\mathbb{Z}, +, \cdot)$ („lineare diophantische Gleichungen“) und den Ring der Polynome.

2.5 Polynome

Im Folgenden sei R ein kommutativer Ring mit Eins und x eine Unbestimmte.

x sollte man sich als formales Symbol – also als eine Art Platzhalter – vorstellen, für das man später je nach Bedarf Zahlen oder auch Objekte wie quadratische Matrizen einsetzen darf.

Ein Polynom mit Koeffizienten in R (ein Polynom über R) ist ein formaler Ausdruck

$$f(x) = a_0(x^0) + a_1x^{(1)} + a_2x^2 + \dots + a_nx^n = \sum_{i=0}^n a_ix^i \quad \text{mit } a_0, \dots, a_n \in R.$$

Die Menge aller solchen Polynome bezeichnen wir mit

$$R[x] := \{f: f(x) = a_0 + a_1x + \dots + a_nx^n \text{ mit } n \in \mathbb{N}_0 \text{ und } a_0, \dots, a_n \in R\}.$$

Sind alle Koeffizienten $a_i = 0$, schreibt man $f = 0$ und nennt es das Nullpolynom.

Der Grad von f ist $\deg f := \begin{cases} -\infty, & \text{falls } f = 0, \\ \max\{i \in \mathbb{N}_0: a_i \neq 0\} & \text{sonst.} \end{cases}$

f heißt konstant, falls $\deg f \leq 0$.

$f \neq 0$ heißt normiert, falls der führende Koeffizient $a_{\deg f} = 1$.

Wenn man das Polynom für konkrete Werte von x auswerten will — für diese müssen die auftretenden Operationen natürlich erklärt sein, wie hier z. B. für Werte aus R selbst —, macht man formal aus dem Polynom $f \in R[x]$ (z. B.) eine Abbildung $\hat{f} \in \text{Abb}(R, R)$ durch

$$\begin{aligned} \hat{\cdot}: R[x] &\rightarrow \text{Abb}(R, R) \\ f &\mapsto \hat{f}: R \rightarrow R, \quad a \mapsto f(a) \end{aligned}$$

Beispiele:

- Für $R = \mathbb{F}_2$ und $f = x + x^2 \in \mathbb{F}_2[x]$ ist $\hat{f}(0) = 0 + 0^2, \hat{f}(1) = 1^2 + 1 = 0$, also \hat{f} die Nullabbildung, aber $f \neq 0$. Die Abbildung $\hat{\cdot}$ ist i. A. nicht injektiv!
- Für $R = \mathbb{Z}$ wollen wir nun eine 2×2 Matrix $A \in M(2 \times 2, \mathbb{Q})$ für x einsetzen, also mit $\hat{\cdot}: \mathbb{Z}[x] \rightarrow \text{Abb}(\mathbb{Q}^{2 \times 2}, \mathbb{Q}^{2 \times 2})$ arbeiten, mit konkreten Werten etwa

$$f(x) = 3 + x^2, \quad \hat{f}(A) = 3A^0 + A^2.$$

Für ein konkretes A ergibt sich

$$A = \begin{bmatrix} -2 & \frac{1}{2} \\ 1 & 0 \end{bmatrix}, \quad A^0 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A^2 = \begin{bmatrix} \frac{9}{2} & -1 \\ -2 & \frac{1}{2} \end{bmatrix},$$

$$\text{insgesamt } \hat{f}(A) = \begin{bmatrix} \frac{15}{2} & -1 \\ -2 & \frac{7}{2} \end{bmatrix}.$$



Wenn klar ist, was gemeint ist, wird $\hat{}$ oft weggelassen.

$R[x]$ als Polynomring über R

Wir erklären Verknüpfungen $+$, \cdot für $f = a_0 + a_1x + \dots + a_nx^n$ und $g = b_0 + b_1x + \dots + b_mx^m$.

O. B. d. A. sei $n = m$ (sonst setze etwa für $n > m$ $b_i = 0$ für $m < i \leq n$).

$$f + g := \underbrace{(a_0 + b_0)}_{c_0} + \underbrace{(a_1 + b_1)}_{c_1}x + \dots + \underbrace{(a_n + b_n)}_{c_n}x^n$$

(kommutativ, 0 ist neutrales, $b_i = -a_i$ inverses Element),

$$f \cdot g := \underbrace{(a_0b_0)}_{c_0} + \underbrace{(a_0b_1 + a_1b_0)}_{c_1}x + \dots + \underbrace{\sum_{i+j=k} (a_i b_j)}_{c_k} x^k + \dots + \underbrace{a_n b_n}_{c_{2n}} x^{2n}$$

(kommutativ, neutrales Element $a_0 = 1$, $a_i = 0$ für $i \in \mathbb{N}$).

Vorsicht! Wird für x irgendein konkretes Objekt A als Wert eingesetzt, muss je nach Rechengesetzen für A nicht immer $\hat{f}(A)\hat{g}(A) = \widehat{(f \cdot g)}(A)$ oder $\hat{f}(A) + \hat{g}(A) = \widehat{(f + g)}(A)$ gelten, z. B.

$$\begin{aligned} \hat{f}(A) \cdot \hat{g}(A) &= (a_0A^0 + a_1A^1)(b_0A^0 + b_1A^1) \\ &\stackrel{?}{=} a_0A^0b_0A^0 + a_0A^0b_1A^1 + a_1A^1b_0A^0 + a_1A^1b_1A^1 \\ &\quad \text{(benötigt Distributivgesetze)} \\ &\stackrel{?}{=} (a_0b_0)A^0 + (a_0b_1 + a_1b_0)A^1 + (a_1b_1)A^2 = \widehat{(f \cdot g)}(A) \end{aligned}$$

(benötigt Distr. + Kommutativität, $A^i \cdot b = b \cdot A^i$ für $b \in R$, $i \in \mathbb{N}_0$).

Für $A \in R$ ist das wegen der Kommutativität und Distributivität von R kein Problem. Will man andere Objekte einsetzen, ist das jeweils vorher zu überprüfen. \(\nabla\)

Beobachtung 2.25 Sei R ein kommutativer Ring mit Eins, dann ist auch $(R[x], +, \cdot)$ ein kommutativer Ring mit Eins. Ist R ein Integritätsbereich (also nullteilerfrei, etwa ein Körper), so ist auch $R[x]$ nullteilerfrei und es gilt

$$\deg(f \cdot g) = \deg f + \deg g,$$

wobei formal $n - \infty = -\infty + n = (-\infty) + (-\infty) = -\infty$ gesetzt wird.

Beweis: s. Vorlesung \(\square\)

Ist R Integritätsbereich oder Körper, gilt daher die Kürzungsregel

$$\text{für } f \neq 0: \quad fg = fh \Leftrightarrow f(g - h) = 0 \Leftrightarrow g = h.$$

Beispiel: Weil $R[x]$ ein Ring mit Eins ist, ist auch die Matrizenrechnung für Polynome sinnvoll einsetzbar, und $(M(n \times n, R[x]), +, \cdot)$ ist selbst wieder ein

Ring; konkret etwa für 2×2 Matrizen über $\mathbb{Z}[x]$

$$A = \begin{bmatrix} 3 & x \\ 1-x & 0 \end{bmatrix}, B = \begin{bmatrix} x^2 & 4-x^5 \\ 0 & 1 \end{bmatrix}, A \cdot B = \begin{bmatrix} 3x^2 & 12+x-3x^5 \\ x^2-x^3 & 4-4x-x^5+x^6 \end{bmatrix}.$$

♡

Kapitel 3

Vektorräume

3.1 Vektorraum

Definition 3.1 Sei $(K, +, \cdot)$ ein Körper. Eine Menge V mit einer inneren Verknüpfung $\oplus: V \times V \rightarrow V$, $(v, w) \mapsto v \oplus w$ „Addition“, äußerer Verknüpfung $\odot: K \times V \rightarrow V$, $(\alpha, v) \mapsto \alpha \odot v$ „skalare Multiplikation/Multiplikation mit Skalaren“ heißt K -Vektorraum/Vektorraum über K , falls gilt:

- (V1) (V, \oplus) ist abelsche Gruppe
(neutrales Element: „Nullvektor“ $\mathbb{0}$ oder 0 , negativer Vektor $-v$),
- (V2) \odot verträgt sich mit $\oplus, +, \cdot$ wie folgt: Für alle $\alpha, \beta \in K, v, w \in V$
 - (V2a) $1 \odot v = v$,
 - (V2b) $(\alpha\beta) \odot v = \alpha \odot (\beta \odot v)$,
 - (V2c) $(\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v)$, \odot bindet stärker als \oplus ,
 - (V2d) $\alpha \odot (v \oplus w) = (\alpha \odot v) \oplus (\alpha \odot w)$.

Die Elemente von V heißen Vektoren (wegen (V1) ist immer $\mathbb{0} \in V \neq \emptyset$).
Die Elemente von K heißen Skalare.

Bemerkung: Die Definition in (V2) funktioniert auch, wenn $(K, +, \cdot)$ durch einen Ring mit Eins $(R, +, \cdot)$ ersetzt wird. Man nennt das dann einen R -Modul, aber das Fehlen des multiplikativen Inversen schränkt die Resultate deutlich ein. Wir gehen nicht näher auf Moduln ein. \diamond

Rechenregeln 3.2

- (a) $0 \odot v = \mathbb{0}$
- (b) $\alpha \odot \mathbb{0} = \mathbb{0}$
- (c) $\alpha \odot v = \mathbb{0} \Rightarrow \alpha = 0 \vee v = \mathbb{0}$
- (d) $(-1) \odot v = -v$

Beweis: s. Vorlesung □

$+$, \cdot , \oplus , \odot vertragen sich so gut, dass man gleich nur $+$ und \cdot verwendet und 0 für $\mathbb{0}$ und \mathbb{O} schreibt. Welche Variante gemeint ist, ergibt sich aus dem Kontext, konkret aus den Argumenten der Verknüpfung.

Beispiele:

- Der Standardraum über dem Körper K ist für $n \in \mathbb{N}_0$

$$(V =) \quad K^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : x_i \in K, i = 1, \dots, n \right\}$$

mit Addition

$$+ : K^n \times K^n \rightarrow K^n, \quad \left(x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) \mapsto x + y := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

und skalarer Multiplikation

$$\cdot : K \times K^n \rightarrow K^n, \quad \left(\alpha, x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) \mapsto \alpha \cdot x := \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

Insbesondere ist $K^0 = \{\mathbb{O}\} = \{[]\}$, er enthält die „leere Liste“.

Konkret wäre also etwa für $K = \mathbb{R}$ und $n = 3$

$$x = \begin{pmatrix} -1 \\ 3 \\ \frac{1}{2} \end{pmatrix} \in \mathbb{R}^3, \quad y = \begin{pmatrix} -\frac{2}{3} \\ 5 \\ 0 \end{pmatrix} \in \mathbb{R}^3, \quad 2x - 3y = \begin{pmatrix} 0 \\ -9 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

- Die Menge $M(m \times n; K) = K^{m \times n}$ der $m \times n$ Matrizen über einem Körper K ist ein Vektorraum über K mit der bereits bekannten Matrix-Addition

$$+ : K^{m \times n} \times K^{m \times n} \rightarrow K^{m \times n} \\ (A, B) \quad \mapsto (A + B)_{ij} = A_{ij} + B_{ij} \text{ für } ij \in m \times n$$

und der skalaren Multiplikation

$$\cdot : K \times K^{m \times n} \rightarrow K^{m \times n} \\ (\alpha, A) \quad \mapsto (\alpha A)_{ij} = \alpha A_{ij} \text{ für } ij \in m \times n.$$

Die Matrix-Multiplikation ist dafür vorerst nicht von Bedeutung. Konkret wäre also etwa für $K = \mathbb{R}$, $m = 2$, $n = 3$,

$$A = \begin{bmatrix} 5 & -3 & 6 \\ -1 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -8 & 4 \\ 0 & 7 & 0 \end{bmatrix}, \quad A - \frac{1}{2}B = \begin{bmatrix} 4 & 1 & 4 \\ -1 & -\frac{7}{2} & 0 \end{bmatrix}.$$

- Die Polynome $K[x]$ über K lassen sich als Vektorraum über K interpretieren, mit der bekannten Addition und der skalaren Multiplikation

$$\begin{aligned} \cdot: \quad & K \times K[x] \quad \rightarrow K[x] \\ & (\alpha, f = a_0 + \dots + a_n x^n) \mapsto \alpha \cdot f := \alpha a_0 + \dots + \alpha a_n x^n, \end{aligned}$$

konkret für $K = \mathbb{Q}$ etwa

$$f = \frac{2}{3}x^2, \quad g = 8x^2 - \frac{1}{4}x^5, \quad h = 7, \quad -6f + \frac{1}{2}g - h = -7 - \frac{1}{8}x^5.$$

- Die Menge der Abb. einer Menge M in einen K -Vektorraum W ($V =$) $W^M = \{f: M \rightarrow W\}$ mit (für $f, g \in W^M, \alpha \in K$)
 punktwiser Addition $(f + g)(m) := f(m) + g(m)$ für alle $m \in M$,
 skalarer Multiplikation $(\alpha \cdot f)(m) := \alpha \cdot f(m)$ für alle $m \in M$
 ist ein K -Vektorraum.

Konkret für das Intervall $M = [0, 1]$ und Bildvektorraum $W = \mathbb{R}^3$,

$$f(t) = \begin{pmatrix} 0 \\ \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix}, \quad g(t) = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix}, \quad (5(f + g))(t) = \begin{pmatrix} 5t \\ 5 \cos(2\pi t) \\ 5 \sin(2\pi t) \end{pmatrix}.$$

- Folgende Variante könnte als weiteres Beispiel dieser Art aufgefasst werden. Sei M eine Menge, und erkläre auf der Potenzmenge 2^M die symmetrische Differenz Δ als Addition ($A + B := A \Delta B = (A \cup B) \setminus (A \cap B)$) und für den Körper \mathbb{F}_2 eine skalare Multiplikation

$$\begin{aligned} \cdot: \quad & \{0, 1\} \times 2^M \rightarrow 2^M \\ & (\alpha, A) \mapsto \begin{cases} A & \alpha = 1, \\ \emptyset & \alpha = 0. \end{cases} \end{aligned}$$

Dann ist 2^M ein Vektorraum über \mathbb{F}_2 .

Konkret für $M = \{1, 2, 3, 4, 5\}$ ist

$$1 \cdot \{2, 3\} + 0 \cdot \{3, 4\} + 1 \cdot \{1, 2, 4\} = \{1, 3, 4\}.$$

♡

Definition 3.3 Ist V ein K -Vektorraum, heißt $W \subseteq V$ Untervektorraum, falls

(UV1) $W \neq \emptyset$,

(UV2) $v, w \in W \Rightarrow v + w \in W$ (W ist abgeschlossen bzgl. Addition),

(UV3) $v \in W, \alpha \in K \Rightarrow \alpha v \in W$ (W ist abgeschlossen bzgl. skal. Mult.).

Wegen (UV3) ist mit $v \in W$ auch $-v \in W$, also $(W, +)$ eine Untergruppe von $(V, +)$.

Beispiele:

- $V = \mathbb{R}^3$. Sind folgende Teilmengen Unterräume?
 $W = \{0\}$, $W = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$,
 $W = \left\{ \alpha \begin{pmatrix} a \\ b \\ c \end{pmatrix} : \alpha \in \mathbb{R} \right\}$ für festes $a, b, c \in \mathbb{R}$ [geometrisch: Gerade durch 0]
 $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_1 \geq 0 \right\}$ [„Halbraum“],
 $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : a_1 x_1 + a_2 x_2 + a_3 x_3 = b \right\}$ mit $a_1, a_2, a_3, b \in \mathbb{R}$ [„Ebene“ mit „Normalvektor“ $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$], enthält 0 nur, falls $b = 0$.
- Für eine gegebene $1 \times n$ Matrix $a \in M(1 \times n; K)$ betrachte die Menge aller $x \in K^n (= K^{n \times 1})$, die $a \cdot x = 0$ (die **homogene Gleichung** einer Gleichung $ax = b$ mit $b \in K$) erfüllen. Etwa konkret für \mathbb{R} ,

$$a = [1 \quad -1 \quad 3], \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad a \cdot x = [x_1 - x_2 + 3x_3] = [0].$$

In der letzten Gleichung kann auf die Klammern gerne verzichtet werden.

Die Menge $W := \text{Lös}(a, 0) = \{x \in K^n : a \cdot x = 0\}$ ist stets ein Untervektorraum von K^n , **denn:** (UV1): $x = 0 \in W$,

(UV2): für $x, y \in W$ ist $a(x + y) = a \cdot x + a \cdot y = 0$,

(UV3): für $x \in W, \alpha \in K$ ist $a(\alpha x) = \sum_{i=1}^n a_{1i}(\alpha x_i) \stackrel{\text{Komm.}}{=} \alpha(a \cdot x) = 0$.
 \square

Wir werden später sehen, dass für $a \neq 0$ die Lösungsmenge W einer sogenannten Hyperebene durch den Ursprung entspricht.

- $V = \text{Abb}(\mathbb{R}, \mathbb{R})$ hat jede Menge Unterräume: bezeichnet einfachheitshalber $\mathbb{R}[x]$ die Funktionen zu den Polynomen, $\mathbb{R}[x]_d$ diejenigen mit $\text{Grad} \leq d$, $C^1(\mathbb{R}, \mathbb{R})$ die einmal stetig differenzierbaren Funktionen, $C^0(\mathbb{R}, \mathbb{R})$ die stetigen, dann ist (s. Analysis)

$$\mathbb{R}[x]_d \subset \mathbb{R}[x] \subset C^1(\mathbb{R}, \mathbb{R}) \subset C^0(\mathbb{R}, \mathbb{R}) \subset \text{Abb}(\mathbb{R}, \mathbb{R}) = \mathbb{R}^{\mathbb{R}}.$$

♡

Satz 3.4 Sei V ein K -Vektorraum. Ein Untervektorraum $W \subseteq V$ bildet mit der induzierten Addition und skalaren Multiplikation einen K -Vektorraum.

Beweis: s. Vorlesung \square

Der Schnitt von Untervektorräumen ist wieder ein Untervektorraum.

Lemma 3.5 Sei $W_i \subseteq V$, $i \in \mathcal{J}$, eine Familie von Untervektorräumen von V , dann ist $W = \bigcap_{i \in \mathcal{J}} W_i$ wieder ein Untervektorraum von V ($\bigcap_{\emptyset} := V$).

Beweis: s. Vorlesung \square

Beispiele:

- Für eine gegebene $m \times n$ Matrix $A \in M(m \times n; K)$ betrachte die Menge aller $x \in K^n$, die $A \cdot x = 0$ (das homogene Gleichungssystem zu einem Gleichungssystem $Ax = b$ mit $b \in K^m$) erfüllen. Etwa konkret für \mathbb{R} ,

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 5 & -4 \\ -1 & 2 & 0 \end{bmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad A \cdot x = \begin{bmatrix} x_1 - x_2 + 3x_3 \\ 2x_1 + 5x_2 - 4x_3 \\ -x_1 + 2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Die Lösungsmenge $W = \{x \in K^n : Ax = 0\}$ ergibt sich als Schnitt der Lösungsmengen $W_i = \{x \in K^n : A_{i,\bullet}x = 0\}$ der einzelnen Zeilen $i \in \mathcal{J} := \{1, \dots, m\}$, im konkreten Beispiel aus den drei Gleichungszeilen

$$\begin{aligned} A_{1,\bullet} &= [1 \quad -1 \quad 3], & A_{1,\bullet} \cdot x &= x_1 - x_2 + 3x_3 = 0, \\ A_{2,\bullet} &= [2 \quad 5 \quad -4], & A_{2,\bullet} \cdot x &= 2x_1 + 5x_2 - 4x_3 = 0, \\ A_{3,\bullet} &= [-1 \quad 2 \quad 0], & A_{3,\bullet} \cdot x &= -x_1 + 2x_2 = 0. \end{aligned}$$

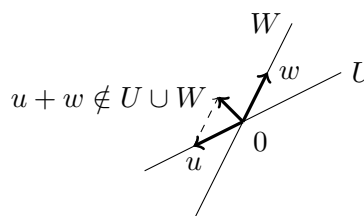
Jedes W_i ist ein Untervektorraum (s. oben), also auch $W = \bigcap_{i \in \mathcal{J}} W_i$. (Man könnte das für W aber auch wie oben direkt zeigen.)

Geometrisch darf man sich die Lösungsmenge eines homogenen Gleichungssystems als Schnitt sogenannter Hyperebenen durch den Ursprung vorstellen.

- Ist $W_d := \mathbb{R}[x]_d \subseteq \mathbb{R}[x]$ der Untervektorraum der Polynome vom Grad $\leq d$, dann ist $\bigcap_{i \in \mathbb{N}_0} W_i = W_0 = \mathbb{R}$.

♡

Die Vereinigung von zwei Untervektorräumen ist i. A. kein Untervektorraum.



Es gilt sogar:

Beobachtung 3.6 Sind $U, W \subseteq V$ Untervektorräume mit $U \cup W$ Untervektorraum, so gilt $U \subseteq W$ oder $W \subseteq U$.

Beweis: s. Vorlesung

□

3.2 Lineare Hülle, Linearkombinationen und lineare Unabhängigkeit

Oft sucht man für eine Teilmenge $S \subseteq V$ den kleinsten Untervektorraum von V , der alle Elemente aus S enthält. Man nennt dies die lineare Hülle von S

in V und erhält sie als Schnitt aller Untervektorräume die S enthalten,

$$\text{lin } S := \bigcap_{W \in \mathcal{W}_S} W \text{ mit } \mathcal{W}_S := \{W : S \subseteq W \wedge W \text{ ist Untervektorraum von } V\},$$

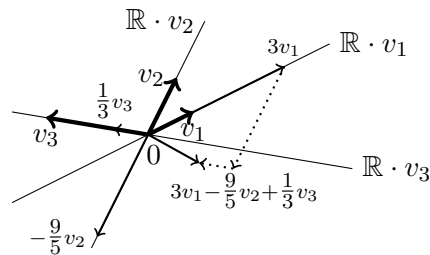
denn wegen L 3.5 ist $\text{lin } S$ wirklich ein Untervektorraum von V . Insbesondere ist $\text{lin } \emptyset = \{0\}$.

Ü: Zeige: Für $L = \{U \subseteq V : U \text{ ist UVR von } V\}$ und Relation $R = \{(v, U) \in V \times L : v \in U\}$ mit Funktionen V_R und L_R wie im Dualitätslemma 1.6 ist $\text{lin } S = V_R(L_R(S))$. Daher gilt $\text{lin}(\text{lin}(S)) = \text{lin}(S)$. Wie ist die bijektive Abbildung \bar{V}_R aus L 1.6 in diesem Zusammenhang zu interpretieren? \times

Jeder Untervektorraum W erfüllt (UV2) und (UV3), daher enthält er für beliebige $v_1, \dots, v_r \in W$, $r \in \mathbb{N}$ und $\alpha_1, \dots, \alpha_r \in K$ auch die Linearkombination

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_r v_r,$$

wobei „linear“ sich darauf bezieht, dass auf den durch die v_i vorgegebenen „geraden Linien“ durch den Ursprung über die α_i Vektoren gewählt werden, die dann durch die Summe „kombiniert“ werden.



Ist ein $v \in V$ für ein $r \in \mathbb{N}$ als Linearkombination von v_1, \dots, v_r darstellbar, sagt man, es wird von v_1, \dots, v_r aufgespannt.

Für eine Familie $(v_i)_{i \in \mathcal{J}}$ von Vektoren aus V nennt man

$$\text{span}_K(v_i)_{i \in \mathcal{J}} := \left\{ \sum_{i \in E} \alpha_i v_i : \alpha_i \in K, i \in E \subseteq \mathcal{J}, E \text{ endlich} \right\}$$

den von der Familie aufgespannten/erzeugten Raum. Die leere Summe ist stets als neutrales Element der Addition definiert, also $\sum_{\emptyset} = 0$, daher ist $\text{span}_K(v_i)_{i \in \emptyset} = \{0\}$.

„Endlich“ ist erforderlich, damit die Summe ein definiertes Ergebnis liefert. Ist K klar, schreibt man nur span statt span_K .

Für eine Menge $S \subseteq V$ setzt man $\text{span}_K(S) := \text{span}_K(v)_{v \in S}$. Der von S erzeugte Raum $\text{span}_K(S)$ enthält also alle endlichen Linearkombinationen von Vektoren aus S .

Beobachtung 3.7 Sei V ein K -Vektorraum und $(v_i)_{i \in \mathcal{J}}$ eine Familie von Vektoren aus V .

- (a) $\text{span}(v_i)_{i \in \mathcal{J}}$ ist ein Untervektorraum von V .
- (b) Ist $W \subseteq V$ Untervektorraum mit $v_i \in W$ für alle $i \in \mathcal{J}$, so ist $\text{span}(v_i)_{i \in \mathcal{J}} \subseteq W$.

Das bedeutet, $\text{span}(S) = \text{lin } S$ sind „innere“ und „äußere“ Beschreibung des kleinsten Untervektorraums, der S enthält.

Beweis: s. Vorlesung □

Beispiele:

- $V = \mathbb{R}^3$ mit $v_1, v_2, v_3 \in \mathbb{R}^3$:
 $\text{span } \emptyset = \{0\}$ ist der Ursprung und Untervektorraum,
 $\text{span}(v_1)$ mit $v_1 \neq 0$ ist geometrisch eine Gerade durch den Ursprung,
 $\text{span}(v_1, v_2)$ mit $v_2 \notin \text{span}(v_1)$: Ebene durch den Ursprung, die die (nicht identischen) Geraden $\text{span}(v_1)$ und $\text{span}(v_2)$ enthält,
 $\text{span}(v_1, v_2, v_3)$ mit $v_2 \notin \text{span}(v_1)$ und $v_3 \notin \text{span}(v_2, v_3)$: \mathbb{R}^3 .
- $V = K^n$, für $i \in \mathcal{J} := \{1, \dots, n\}$ sei $e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$, (oder mit
 Kroneckersymbol $(e_i)_j := \delta_{ij}$, $j \in \mathcal{J}$), dann ist $\text{span}_K(e_i)_{i \in \mathcal{J}} = K^n$.
- Für $v_1, \dots, v_n \in K^m$ und $A = [v_1, \dots, v_n] \in K^{m \times n}$ ist

$$\text{span}_k(v_1, \dots, v_n) = \{A \cdot x : x \in K^n\},$$

passend zur spaltenweise Interpretation der Matrixmultiplikation. Konkret für $K = \mathbb{R}$, $m = 3$, $n = 3$,

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}, A = [v_1, v_2, v_3] = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 5 & -4 \\ -1 & 2 & 0 \end{bmatrix},$$

$$A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} x_1 + \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix} x_2 + \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix} x_3 \stackrel{\text{Komm}}{=} x_1 v_1 + x_2 v_2 + x_3 v_3.$$

Ein Gleichungssystem $Ax = b$ fragt also danach, ob b im Untervektorraum enthalten ist, der von den Spalten $A_{\bullet, j}$ von A aufgespannt wird.

- Für $V = K[x]$ und $e_i := x^i$, $i \in \mathbb{N}_0$, ist $\text{span}_K(e_i)_{i \in \mathbb{N}_0} = K[x]$.
 Für $N = \{0, \dots, d\}$ ist $\text{span}_K(e_i)_{i \in N} = K[x]_d$.
- Für M eine beliebige Menge und $V = K^M$ bezeichne, für $m \in M$, e_m das M -Tupel mit 1 in Position m , 0 sonst.

Für $\mathcal{J} = M$ ist $\text{span}_K(e_i)_{i \in \mathcal{J}} \begin{cases} = K^M & \text{für } M \text{ endlich,} \\ \neq K^M & \text{für } M \text{ unendlich, weil} \end{cases}$

im unendlichen Fall $\mathbb{1} \notin \text{span}(e_i)_{i \in \mathcal{J}}$, wobei $\mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ (formal ist $\mathbb{1}: M \rightarrow K$, $m \mapsto 1$).

♡

Lässt sich in einer Familie $(v_i)_{i \in \mathcal{J}}$ ein v_i als Linearkombination anderer v_i darstellen, z. B.

$$v_i = \sum_{i \in \hat{E}} \alpha_i v_i \quad \text{für ein endliches } \hat{E} \subseteq \mathcal{J} \setminus \{i\},$$

dann ist $\text{span}(v_i)_{i \in \mathcal{J}} = \text{span}(v_i)_{i \in \mathcal{J} \setminus \{i\}}$, denn v_i kann immer durch die Vektoren aus \hat{E} ersetzt werden, und jedes $v \in \text{span}(v_i)_{i \in \mathcal{J}}$ kann dann auf viele Arten aus den v_i linear kombiniert werden.

Um Eindeutigkeit der Beschreibung als Linearkombination garantieren zu können, müssen solche Abhängigkeiten ausgeschlossen werden.

Definition 3.8 Sei V ein K -Vektorraum. Eine endliche Familie $v_1, \dots, v_r \in V$ heißt linear unabhängig (l.u.) falls für $\alpha_1, \dots, \alpha_r \in K$

$$\alpha_1 v_1 + \dots + \alpha_r v_r = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \dots = \alpha_r = 0.$$

Man sagt auch: Der Nullvektor lässt sich nur trivial aus v_1, \dots, v_r linear kombinieren. Insbesondere ist die leere Familie linear unabhängig.

Eine beliebige Familie $(v_i)_{i \in \mathcal{J}}$ von Vektoren aus V heißt linear unabhängig (l.u.), falls jede endliche Teilfamilie l.u. ist.

Ein Familie $(v_i)_{i \in \mathcal{J}}$, die nicht l.u. ist, heißt linear abhängig (l.a.).

Lemma 3.9 (Charakterisierung der linearen Unabhängigkeit)

Für eine Familie $(v_i)_{i \in \mathcal{J}}$ von Vektoren aus V ist äquivalent:

- (a) $(v_i)_{i \in \mathcal{J}}$ ist l.u.
- (b) Für jede echte Teilmenge $\mathcal{J}' \subsetneq \mathcal{J}$ ist $\text{span}(v_i)_{i \in \mathcal{J}'} \subsetneq \text{span}(v_i)_{i \in \mathcal{J}}$.
- (c) Jedes $v \in \text{span}(v_i)_{i \in \mathcal{J}}$ hat eine eindeutige Darstellung als Linearkombination der v_i .

Beweis: s. Vorlesung □

Beispiel: Die zuvor angegebenen (e_i) für K^n , $K[x]$ und K^M sind l.u., **denn:** für endliches $E \subseteq M$ folgt aus $\sum_{i \in E} \alpha_i e_i = 0$, dass für $j \in E$ jeweils $0_j = \sum_{i \in E} \alpha_i (e_i)_j = \alpha_j$. □ ♥

Beobachtung 3.10 (Regeln zur linearen Unabhängigkeit)

In jedem K -Vektorraum gilt:

- (a) Ist $v \in \text{span}(v_i)_{i \in \mathcal{J}}$, dann ist $(v, v_i)_{i \in \mathcal{J}}$ l.a.
- (b) Ist $(v_i)_{i \in \mathcal{J}}$ l.u. und $v \notin \text{span}(v_i)_{i \in \mathcal{J}}$, so ist $(v, v_i)_{i \in \mathcal{J}}$ l.u. (insbesondere ist ein einzelnes $v \in V$ l.u. $\Leftrightarrow v \neq 0$).
- (c) $(v_i)_{i \in \mathcal{J}}$ ist l.a. \Leftrightarrow Ein v_i ist Linearkombination der anderen.

Beweis: s. Vorlesung □

Unabhängig davon, ob man weiß, ob $(v_i)_{i \in \mathcal{J}}$ l.u. ist oder nicht, sagt man für einen einzelnen Vektor $v \in V$ mit

$$\begin{aligned} v \in \text{span}(v_i)_{i \in \mathcal{J}} & \text{ „}v \text{ ist l.a. von } (v_i)_{i \in \mathcal{J}}\text{“,} \\ v \notin \text{span}(v_i)_{i \in \mathcal{J}} & \text{ „}v \text{ ist l.u. von } (v_i)_{i \in \mathcal{J}}\text{.“} \end{aligned}$$

Beispiel 3.11

- (a) Betrachte für gegebenes $A \in M(m \times n, K)$ den von den Spalten von A aufgespannten Untervektorraum $S = \text{span}(A_{\bullet, j})_{j \in \{1, \dots, n\}} = \{\sum_{j=1}^n x_j A_{\bullet, j} : x \in K^n\}$. Das Gleichungssystem $Ax = b$ hat keine Lösung, falls $b \notin S$. Sonst hat es mindestens eine und genau eine, wenn der Lösungsuntervektorraum $\text{Lös}(A, 0) := \{y \in K^n : Ay = 0\}$ des homogenen Gleichungssystems $Ay = 0$ nur die triviale Lösung 0 enthält, **denn:** $\text{Lös}(A, 0) = \{0\}$ bedeutet $\sum_{j=1}^n A_{\bullet, j} y_j = 0 \Leftrightarrow y = 0$, also sind alle Spalten von A l.u. und die Linearkombination durch x in $Ax = b$ ist eindeutig (L3.9(c)). □

In der Tat verrät $\text{Lös}(A, 0)$ exakt, welche Spalten linear abhängig sind: Für $\mathcal{J} \subseteq \{1, \dots, n\}$ ist $(A_{\bullet, j})_{j \in \mathcal{J}}$ genau dann l.a., wenn es ein $0 \neq \bar{y} \in \text{Lös}(A, 0)$ gibt mit $\{j : \bar{y}_j \neq 0\} \subseteq \mathcal{J}$ (**Ü**).

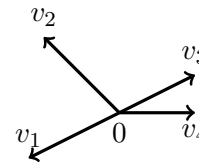
Ü: Der Algorithmus Alg 2.24 gebe für Eingabe $A \in K^{m \times n}$ und $b = 0 \in K^m$ die Stufenindices $\mathcal{J} = \{j_1, \dots, j_r\}$ aus. Man zeige, dass die Spaltenvektoren $(A_{\bullet, j})_{j \in \mathcal{J}}$ linear unabhängig sind und dass \mathcal{J} (inklusions-)maximal bezüglich dieser Eigenschaft ist, d. h., \mathcal{J} kann nicht mehr vergrößert werden, ohne die lineare Unabhängigkeit der zugehörigen Spalten zu verlieren. x

Algorithmus Alg 2.24 erlaubt also, für n gegebene Vektoren aus dem K^m eine maximale Familie linear unabhängiger Vektoren zu bestimmen. Das Ergebnis hängt entscheidend von der Reihenfolge ab, mit der die Vektoren in die Spalten der Eingabematrix geschrieben werden.

Erprobe die Rückgabemengen für die vier Vektoren

$$v_1 = \begin{pmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

und einige mögliche Anordnungen in Matrizen, etwa $A_1 = [v_1, v_2, v_3, v_4]$, $A_2 = [v_3, v_1, v_4, v_2]$ und $A_3 = [v_4, v_2, v_3, v_1]$.



Zu klären bleibt auch, wie stark $|\mathcal{J}|$ von der Reihenfolge abhängt.

- (b) Seien für $x \in K^n$ m Gleichungen der Form $A_i x = 0$ mit $A_i \in K^{1 \times n}$ gegeben, konkret etwa für $n = 3$ und $K = \mathbb{R}$ $m = 4$ Gleichungen,

$$\begin{aligned} A_1 &= [1 & -1 & 3], & A_1 \cdot x &= x_1 - x_2 + 3x_3 = 0, \\ A_2 &= [2 & 5 & -4], & A_2 \cdot x &= 2x_1 + 5x_2 - 4x_3 = 0, \\ A_3 &= [-1 & 2 & 0], & A_3 \cdot x &= -x_1 + 2x_2 = 0, \\ A_4 &= [5 & 0 & -1], & A_4 \cdot x &= 5x_1 - x_3 = 0. \end{aligned}$$

Da ja in jeder Gleichung auf beiden Seiten der gleiche Wert steht, erhält man wieder eine gültige Gleichung, wenn man Gleichung i mit einem Wert $\alpha_i \in K$ multipliziert und diese aufsummiert, also die neue Gleichung

$$A_{m+1} := \left(\sum_{i=1}^m \alpha_i A_i \right) x = 0$$

als Linearkombination der alten bildet. Die Lösungsmenge ändert sich dadurch nicht,

$$\{x \in K^n : A_i x = 0, i = 1, \dots, m\} = \{x \in K^n : A_i x = 0, i = 1, \dots, m+1\}.$$

Lässt sich nun eine der Gleichungen als eine Linearkombination der anderen ausdrücken (sie liegt im span der anderen, sie ist von den anderen l.a.), im Beispiel etwa

$$A_4 = 1 \cdot A_1 + 1 \cdot A_2 - 2 \cdot A_3,$$

gibt diese keine neue Information und kann ohne Veränderung der Lösungsmenge weggelassen werden.

In einem homogenen Gleichungssystem $Ax = 0$ mit $A \in K^{m \times n}$ hängt also der Lösungsuntervektorraum $\text{Lös}(A, 0)$ nur von dem durch die Zeilen von A aufgespannten anderen Untervektorraum $\text{span}(A_{i,\bullet})_{i=1,\dots,m}$ ab, der für die algorithmische Lösung am besten durch möglichst wenige l.u. Zeilenvektoren aufgespannt werden sollte.

♡

3.3 Basis und Dimension

Definition 3.12 Eine Familie $\mathcal{B} = (v_i)_{i \in \mathcal{J}}$ von Vektoren in einem Vektorraum V heißt Erzeugendensystem von V , wenn $V = \text{span}(v_i)_{i \in \mathcal{J}}$, also wenn sich jedes $v \in V$ als Linearkombination endlich vieler Vektoren aus \mathcal{B} darstellen lässt.

Ein linear unabhängiges Erzeugendensystem $\mathcal{B} = (v_i)_{i \in \mathcal{J}}$ heißt Basis.

V heißt endlich erzeugt, falls es ein endliches Erzeugendensystem $\mathcal{B} = (v_1, \dots, v_n)$ besitzt. Ist \mathcal{B} Basis, nennt man n die Länge der Basis.

Beispiele:

- Für $V = \{0\}$ ist die leere Familie ($\mathcal{J} = \emptyset$) eine Basis.
- $\mathcal{B} = (e_1, \dots, e_n)$ ist eine Basis des K^n , die kanonische/Standardbasis.
 $\mathcal{B}' = (e_1, e_2 + e_1, \dots, e_n + e_1)$ ist ebenfalls eine Basis.

- Für $m \times n$ Matrizen $M(m \times n; K) = K^{m \times n}$ ist $\mathcal{B} = (E_{ij}^{(m,n)})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ eine Basis, wobei $(E_{ij}^{(m,n)})_{hk} = \delta_{(i,j)(h,k)}$ für $1 \leq h \leq m$, $1 \leq k \leq n$, also eine 1 in Element (i, j) hat und sonst 0 ist. [Das entspricht e_{jm+i} in K^{mn} .]
- $\mathcal{B} = (1, x, x^2, \dots)$ ist die kanonische Basis von $K[x]$.

♡

Die entscheidende Frage ist: Gibt es immer eine Basis, und, wenn ja, haben alle Basen gleiche Kardinalität?

Wir werden das hier nur für Vektorräume mit endlichen Erzeugendensystemen beantworten, für unendliche benötigt man das Zornsche Lemma/das Auswahlaxiom als Voraussetzung.

Satz 3.13 (Charakterisierung von Basen)

Sei $\mathcal{B} = (v_i)_{i \in \mathcal{J}}$ eine Familie von Vektoren in V , dann sind äquivalent:

- \mathcal{B} ist Basis von V .
- \mathcal{B} ist minimales/„unverkürzbares“ Erzeugendensystem (lässt man einen beliebigen Vektor weg, ist es kein Erzeugendensystem mehr).
- Für jedes $v \in V$ gibt es eindeutig bestimmte $\alpha_i \in K$, $i \in \mathcal{J}$, mit $E = \{i \in \mathcal{J} : \alpha_i \neq 0\}$ endlich, sodass $v = \sum_{i \in \mathcal{J}} \alpha_i v_i$. Die α_i heißen Koordinaten von v bzgl. \mathcal{B} .
- \mathcal{B} ist maximal bzgl. linearer Unabhängigkeit (d.h., \mathcal{B} ist l.u. und man kann kein $v \in V$ zu \mathcal{B} hinzufügen, ohne die lin. Unabh. zu verlieren).

Beweis: s. Vorlesung □

Aus dem Beweis zu (d) \Rightarrow (a) folgt sofort:

Korollar 3.14 Ist V nicht endlich erzeugt, gibt es zu jeder endlichen linear unabhängigen Familie ein linear unabhängiges $v \in V$, insbesondere also eine unendliche linear unabhängige Familie.

Die Frage nach der Existenz einer Basis läuft darauf hinaus, ob es in jedem Raum eine bzgl. linearer Unabhängigkeit maximale Menge gibt. Nimmt man das Zornsche Lemma/Auswahlaxiom an, kann man zeigen:

Satz 3.15 Jeder Vektorraum besitzt eine Basis.

Aber das geht über den Rahmen dieser Vorlesung hinaus.

Für endlich erzeugte Vektorräume braucht es keine solche Annahme:

Satz 3.16 (Basisauswahlsatz)

Aus jedem endlichen Erzeugendensystem eines Vektorraums kann man eine Basis auswählen. Insbesondere hat jeder endlich erzeugte Vektorraum eine endliche Basis.

Beweis: s. Vorlesung □

Um zu zeigen, dass alle Basen eines endlich erzeugten Vektorraums gleich lang sind, werden wir schrittweise die Basiselemente gegeneinander austauschen:

Lemma 3.17 (Austauschlemma von Steinitz)

Sei $\mathcal{B} = (v_1, \dots, v_r)$ eine Basis eines K -Vektorraums V und $w = \alpha_1 v_1 + \dots + \alpha_r v_r \in V$, dann ist für jedes $k \in \{1, \dots, r\}$ mit $\alpha_k \neq 0$ $\mathcal{B}' = (v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r)$ wieder eine Basis von V .

Beweis: s. Vorlesung □

Satz 3.18 (Austauschsatz)

Sei $\mathcal{B} = (v_1, \dots, v_r)$ eine Basis eines K -Vektorraums V und (w_1, \dots, w_n) eine linear unabhängige Familie in V . Dann ist $n \leq r$ (Folgerung!) und es gibt Indices $i_1, \dots, i_n \in \{1, \dots, r\}$, sodass durch Ersetzen von v_{i_j} durch w_j , $j = 1, \dots, n$, wieder eine Basis entsteht.

Bei geeigneter Umnummerierung ist $\bar{\mathcal{B}} = (w_1, \dots, w_n, v_{n+1}, \dots, v_r)$ wieder eine Basis.

Beweis: s. Vorlesung □

Korollar 3.19 Hat ein Vektorraum V eine Basis der Länge $r \in \mathbb{N}_0$, so hat jede Basis von V Länge r .

Beweis: s. Vorlesung □

Definition 3.20 Für einen K -Vektorraum V heißt

$$\dim_K V := \begin{cases} r, & \text{falls } V \text{ eine Basis der Länge } r \in \mathbb{N}_0 \text{ besitzt,} \\ \infty, & \text{falls } V \text{ keine endliche Basis besitzt,} \end{cases}$$

die *Dimension* von V über K . Man schreibt nur $\dim V$, wenn klar ist, welcher Körper gemeint ist. Es ist günstig, $\dim \emptyset = -1$ zu setzen.

Korollar 3.21 Ist $W \subseteq V$ Untervektorraum eines endlich erzeugten Vektorraums V , so ist auch W endlich erzeugt mit $\dim W \leq \dim V$. Falls $\dim V = \dim W$, gilt $V = W$.

Beweis: s. Vorlesung □

Satz 3.22 (Basisergänzungssatz)

Sei V ein endlich erzeugter Vektorraum mit $\dim V = r$ und $w_1, \dots, w_n \in V$ $n \leq r$ gegebene l.u. Vektoren. Dann gibt es w_{n+1}, \dots, w_r , sodass $\mathcal{B} = (w_1, \dots, w_n, w_{n+1}, \dots, w_r)$ eine Basis von V ist.

Beweis: s. Vorlesung □

Beispiele:

- $\dim K^n = n$, weil $e_i, i = 1, \dots, n$, Basis.
- $\dim K[x] = \infty$.
- $\dim M(m \times n; K) = m \cdot n$.
- Sei $A \in (m \times n; K)$ gegeben:
 - $Ax = b$ in Spaltensicht: $S := \{Ax = \sum_{j=0}^n x_j A_{\bullet,j} : x \in K^n\} = \text{span}(A_{\bullet,j})_{j=1, \dots, n}$ ist der von den Spalten von A aufgespannte Unterraum des K^m . Egal, welche maximale (=nicht vergrößbare) Menge an l.u. Spalten $\mathcal{J} \subseteq \{1, \dots, n\}$ der Matrix man auswählt, jede solche Wahl ist eine Basis von S , sie enthält gleich viele Vektoren und spannt den gleichen Raum S auf. Für jedes gewählte \mathcal{J} lässt sich ein $b \in S$ auf eindeutige Weise als Linearkombination der $(A_{\bullet,j})_{j \in \mathcal{J}}$ ausdrücken (und ein $\hat{b} \notin S$ gar nicht). Für jedes \mathcal{J} gibt es dann pro $j \in \mathcal{J}$ genau einen Wert für Koeffizienten x_j , sodass $b = \sum_{j \in \mathcal{J}} x_j A_{\bullet,j}$ (dieser Wert hängt i. A. von der Wahl von \mathcal{J} ab), die restlichen x_j setzt man dann auf 0. Wählt man für ein (oder mehrere) x_j mit $\hat{j} \in \{1, \dots, n\} \setminus \mathcal{J}$ einen anderen Wert als 0, dann ist wegen $A_{\bullet,\hat{j}} \in \text{span}(A_{\bullet,j})_{j \in \mathcal{J}}$ auch $b - x_{\hat{j}} A_{\bullet,\hat{j}} \in \text{span}(A_{\bullet,j})_{j \in \mathcal{J}}$ und der Effekt von $A_{\bullet,\hat{j}} x_{\hat{j}}$ durch geeignete (eindeutige) Anpassung der $x_j, j \in \mathcal{J}$ kompensierbar. In der Sprache der linearen Gleichungssysteme legt die gewählte Basis die gebundenen Variablen fest, die restlichen sind die freien Variablen.
 - $Ax = 0$ in Zeilensicht: Der Untervektorraum $\text{Lös}(A, 0)$ ist für jede Auswahl einer maximalen linear unabhängigen Menge an Zeilen (= Basis des von den Zeilen der Matrix aufgespannten anderen Untervektorraums) der gleiche, und jede solche Auswahl hat gleich viele Zeilen.

♡

3.4 Summen von Vektorräumen

Der Schnitt von Untervektorräumen ist ein Vektorraum, die Vereinigung aber i. A. nicht, man muss zum „aufgespannten“ Unterraum übergehen.

Definition 3.23 Für Untervektorräume W_1, \dots, W_n eines K -Vektorraums V heißt

$$W_1 + \dots + W_n := \{v \in V : \text{es gibt } v_j \in W_j \text{ mit } v = v_1 + \dots + v_n\}$$

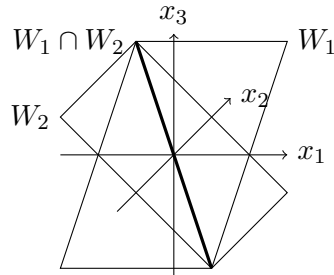
die Summe von W_1, \dots, W_n .

Beobachtung 3.24

- (a) $W_1 + \dots + W_n = \text{span}(W_1 \cup \dots \cup W_n)$,
- (b) $W_1 + \dots + W_n \subseteq V$ ist ein Untervektorraum,
- (c) $\dim(W_1 + \dots + W_n) \leq \dim W_1 + \dots + \dim W_n$.

Beweis: s. Vorlesung □

In (c) wird die linke Seite echt kleiner als die rechte Seite, wenn sich Teilsummen der Unterräume nicht nur in der Null schneiden (= nicht trivial schneiden). Im Fall von zwei Unterräumen ist das leicht exakt beschreibbar.



Satz 3.25 (Dimensionsformel für Summen)

Für endlichdimensionale Untervektorräume $W_1, W_2 \subseteq V$ gilt:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

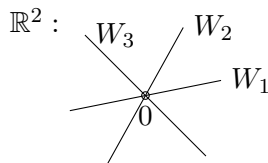
Beweis: s. Vorlesung □

Lemma 3.26 Ist $V = W_1 + W_2$ so sind äquivalent:

- (a) $W_1 \cap W_2 = \{0\}$.
- (b) Jedes $v \in V$ hat eine eindeutige Darstellung $v = w_1 + w_2$ mit $w_1 \in W_1, w_2 \in W_2$.
- (c) Je zwei Vektoren $v_1 \in W_1, v_2 \in W_2, v_1 \neq 0 \neq v_2$ sind linear unabhängig.

Beweis: s. Vorlesung □

Bei mehr als zwei Unterräumen reicht es *nicht* $W_1 \cap \dots \cap W_k = \{0\}$ zu fordern, um eine eindeutige Darstellung garantieren zu können:



$$\dim W_1 = \dim W_2 = \dim W_3 = 1,$$

$$\dim(W_1 + W_2 + W_3) = 2.$$

Die Verallgemeinerung von (c) führt aber zum Ziel:

Definition 3.27 Ein Vektorraum V heißt (*innere*) *direkte Summe* von Untervektorräumen W_1, \dots, W_k , in Zeichen $V = W_1 \oplus \dots \oplus W_k = \bigoplus_{i=1}^k W_i$, wenn gilt:

- (DS1) $V = W_1 + \dots + W_k$
- (DS2) Für $w_i \in W_i, i = 1, \dots, k$ mit $w_1 + \dots + w_k = 0$ folgt $w_1 = \dots = w_k = 0$.
[D.h., je k Vektoren aus den k Vektorräumen sind l.u.]

Beispiel: Ist (v_1, \dots, v_n) eine Basis von V über K , so ist

$$V = K \cdot v_1 \oplus \dots \oplus K \cdot v_n.$$

♡

Satz 3.28 Für endlichdimensionale Vektorräume sind äquivalent:

- (a) $V = W_1 \oplus \dots \oplus W_k$.
- (b) Sei $(v_1^{(j)}, \dots, v_{r_1}^{(j)})$ eine Basis von W_j , $j = 1, \dots, k$, dann ist $\mathcal{B} = (v_1^{(1)}, \dots, v_{r_1}^{(1)}, v_1^{(2)}, \dots, v_{r_k}^{(k)})$ eine Basis von V .
- (c) $V = W_1 + \dots + W_k$ und $\dim V = \dim W_1 + \dots + \dim W_k$.

Beweis: s. Vorlesung

□

3.5 Affine Unterräume, Quotientenvektorräume

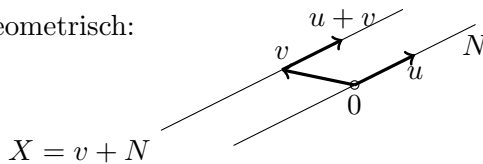
Untervektorräume müssen immer den Ursprung enthalten. „Verschiebt“ man den Untervektorraum um ein $v \in V$ (auch $v = 0$ ist erlaubt), spricht man von einem affinen Unterraum (Untervektorräume sind also Spezialfälle affiner Unterräume).

Definition 3.29

Eine Teilmenge X eines K -Vektorraums V heißt affiner Unterraum, falls $X = \emptyset$ ist oder es ein $v \in V$ und einen Untervektorraum $N \subseteq V$ gibt mit $X = v + N = \{w + v : w \in N\}$.

Die Dimension von $X = v + N$ ist $\dim X = \dim N$ und $\dim \emptyset = -1$.

Geometrisch:



Beispiele:

- Im Raum \mathbb{R}^2 sind \emptyset , jeder Punkt $p \in \mathbb{R}^2$ ($v = p$, $N = \{0\}$), jede Gerade ($v + \text{span}\{u\}$) und der ganze \mathbb{R}^2 alle affinen Unterräume.
- \mathbb{R}^3 : \emptyset , jeder Punkt, jede Gerade, jede Ebene, \mathbb{R}^3 selbst.
- \mathbb{F}_2^2 : \emptyset , jeder Punkt in $\{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})\}$, \mathbb{F}_2^2 und jede Gerade $v + N$: $(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}) + 1\text{-dim Untervektorraum: } \{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix})\}, \{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 0 \end{smallmatrix})\}, \{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})\}$, in \mathbb{F}_2^2 hat jede Gerade 2 Punkte, höchstens gibt es $\binom{4}{2} = \frac{4 \cdot 3}{2} = 6$. $(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}) + \text{Untervektorräume: } \{(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})\}, [(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 0 \end{smallmatrix})], \{(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix})\}$, $(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}) + \text{Untervektorräume: } [(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 0 \end{smallmatrix})], \{(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})\}$, fertig, 6 gefunden.
- Für $A \in M(m \times n; K)$ ist $N = \{x \in K^n : Ax = 0\} = \text{Lös}(A, 0)$ ein Untervektorraum. Für jedes $\hat{x} \in K$ ist $\hat{x} + N$ ein affiner Unterraum. Ist $b = A\hat{x}$, gilt für $x \in \hat{x} + N$, dass $x = \hat{x} + v$ mit $Av = 0$, also ebenso

$Ax = A\hat{x} + Av = b$ und damit $\hat{x} + N \subseteq \text{Lös}(A, b)$. Wir werden später sehen, dass sogar Gleichheit gilt und $\text{Lös}(A, b)$ für jedes $b \in K^m$ ein affiner Unterraum dieser Form ist.

♡

Definition 3.30 Sei X ein affiner Unterraum eines K -Vektorraums V .

Falls $\dim(X) = 0$, heißt X Punkt,
 $= 1$, Gerade,
 $= 2$, Ebene.

Falls $\dim V < \infty$ und $\dim X = \dim V - 1$, heißt X Hyperebene.

Beobachtung 3.31 Sei V ein K -Vektorraum und $X = v + N$ ein affiner Unterraum mit $v \in V$ und N Untervektorraum von V . Für jedes $w \in X$ ist $X = w + N$ bzw. $N = X - w$.

Beweis: s. Vorlesung □

Es ist somit egal, welchen Punkt aus X man als „Träger“ verwendet. Anders gesagt, für $u, v \in V$ und einen Untervektorraum N ist $u - v \in N \Leftrightarrow u + N = v + N$; jeder Untervektorraum N induziert also über $u \sim_N v :\Leftrightarrow u - v \in N$ eine Äquivalenzklasseneinteilung des Vektorraumes in affine Unterräume.

Beobachtung 3.32 Der Schnitt affiner Unterräume eines Vektorraums V ist wieder ein affiner Unterraum von V .

Beweis: s. Vorlesung □

Ähnlich zur linearen Hülle definiert man für $S \subseteq V$ die affine Hülle von S ,

$$\text{aff } S := \bigcap_{X \in \mathcal{X}_S} X, \quad \mathcal{X}_S := \{X \subseteq V : X \text{ affiner Unterraum, } S \subseteq X\}, \quad \text{aff } \emptyset = \emptyset.$$

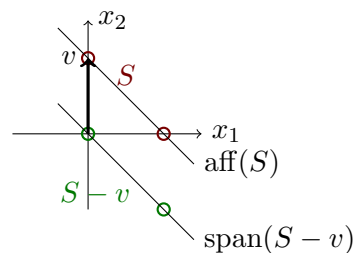
Ü: Beschreibe $\text{aff } S$ über eine Relation wie in der Übung zu $\text{lin } S$. ✕

Beobachtung 3.33 Für $v \in S \subseteq V$ gilt $\text{aff } S = v + \text{span}(S - v)$.

Beweis: s. Vorlesung □

Beispiel: Für $S = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^2$ und $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ist

$$\begin{aligned} \text{aff}(S) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \text{span} \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \\ &= \left\{ x = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}. \end{aligned}$$



♡

Ü: Konstruiere aff S für $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$. ⌘

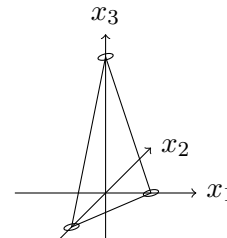
Beobachtung 3.34 Für eine Teilmenge S eines K -Vektorraumes ist $v \in \text{aff } S$ genau dann, wenn es $r \in \mathbb{N}$, $s_1, \dots, s_r \in S$ und $\alpha_1, \dots, \alpha_r \in K$ mit $\alpha_1 + \dots + \alpha_r = 1$ gibt, sodass $v = \alpha_1 s_1 + \dots + \alpha_r s_r$ (v ist Affinkombination der s_i).

Beweis: s. Vorlesung □

Beispiele:

- Für $S = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^2$ ist $\text{aff } S = \left\{ \alpha_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \alpha_1 + \alpha_2 = 1, \alpha_1, \alpha_2 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \alpha_2 \\ \alpha_1 \end{pmatrix} \in \mathbb{R}^2 : \alpha_1 + \alpha_2 = 1 \right\} = \left\{ \begin{pmatrix} 1-\alpha \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$.
- Im \mathbb{R}^3 sei die Gleichung $3x_1 - 2x_2 + x_3 = 6$ gegeben. Sie wird von den Punkten

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in S := \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix} \right\}$$



erfüllt. Jedes $x \in \text{aff } S$ ist Lösung (warum?). Die Frage ob $\text{aff } S$ auch schon die Menge aller Lösungen ist, wird später geklärt. ♡

Eine Familie $(v_i)_{i \in \mathcal{J}}$ von Vektoren in V heißt affin unabhängig, falls für endliches $E \subseteq \mathcal{J}$ stets $\dim(\text{aff}\{v_i : i \in E\}) = |E| - 1$.

Ü: Eine Familie $(v_i)_{i \in \mathcal{J}}$ ist affin unabhängig, wenn für jeden Vektor $v_i, i \in \mathcal{J}$, die Vektoren $(v_i - v_i)_{i \in \mathcal{J} \setminus \{i\}}$ l.u. sind. ⌘

Wie bereits zuvor angemerkt, induziert jeder Untervektorraum N über die Äquivalenzrelation $u \sim_N v \Leftrightarrow u - v \in N$ eine Klasseneinteilung von V in die mit N assoziierten affinen Unterräume. Diese bilden die Quotientenmenge $V/N = \{v + N : v \in V\}$. Überträgt man die Vektorraumaddition und skalare Multiplikation aus V auf V/N , erhält man wieder einen Vektorraum.

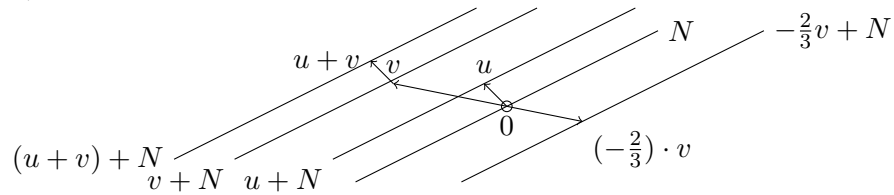
Beobachtung 3.35 Sei V ein K -Vektorraum und $N \subseteq V$ ein Untervektorraum. $V/N = \{v + N : v \in V\}$ bildet mit den Verknüpfungen

$$\begin{aligned} \hat{+}: V/N \times V/N &\rightarrow V/N & \hat{\cdot}: K \times V/N &\rightarrow V/N \\ (u + N, v + N) &\mapsto (u + v) + N & (\alpha, v + N) &\mapsto (\alpha \cdot v) + N \end{aligned}$$

einen K -Vektorraum, den Quotientenvektorraum von V nach N .

Beweis: s. Vorlesung □

In V/N rechnet man gewissermaßen mit affinen Unterräumen:



Im Quotientenvektorraum V/N werden alle Punkte eines affinen Unterraums als gleich betrachtet. Man ignoriert alle Anteile in Richtung des Untervektorraums N , und nur die nicht von N aufgespannten Anteile sind noch von Belang.

Kapitel 4

Lineare Abbildungen

4.1 Definition und erste Eigenschaften

Definition 4.1 Eine Abbildung $F: V \rightarrow W$ zwischen K -Vektorräumen V und W heißt (K) -linear oder $(K$ -Vektorraum)Homomorphismus oder linearer Operator, falls

$$(L) \quad F(\alpha v + \beta w) = \alpha F(v) + \beta F(w) \quad \text{für alle } \alpha, \beta \in K, v, w \in V \text{ gilt.}$$

(L) ist äquivalent dazu, folgende zwei Eigenschaften zu fordern:

$$(L1) \quad F(v+w) = F(v) + F(w) \quad \forall v, w \in V \quad [\text{Gruppenhom. für } (V, +), (W, +)],$$

$$(L2) \quad F(\alpha v) = \alpha F(v) \quad \forall \alpha \in K, v \in V \quad [\text{Veträglichkeit mit skalarer Mult.}].$$

Bemerkung: Auch hier kann in der Definition statt eines Körpers K nur ein Ring R verwendet werden, also R -Moduln V und W , aber wir werden nur über Körper sprechen. \diamond

Die Menge der K -linearen Abbildungen von V nach W bezeichnet man kurz mit

$$\text{Hom}_K(V, W).$$

Wie bei Gruppen heißt eine lineare Abb. $F: V \rightarrow W$

Monomorphismus, wenn F injektiv ist,

Epimorphismus, surjektiv,

Isomorphismus, bijektiv,

Endomorphismus, wenn $V = W$, $\text{End}_K(V) := \text{Hom}_K(V, V)$,

Automorphismus, wenn $V = W$ und F bijektiv, $\text{Aut}_K(V) \subseteq \text{End}_K(V)$.

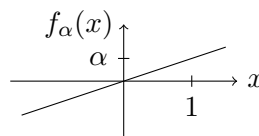
Beispiele:

- In \mathbb{R} ist für $\alpha \in \mathbb{R}$ die Abbildung

$$f_\alpha: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \alpha x$$

linear ($\in \text{Hom}_K(\mathbb{R}, \mathbb{R})$).

α gibt die konstante Steigung an.



- Die Streckung/Homothetie um einen Faktor $\alpha \in K$ ist linear:

$$F \in \text{End}_K(V), v \mapsto \alpha v \quad (\in \text{Hom}_K(V, V)).$$

- Jede Matrix $A \in M(m \times n; K)$ beschreibt eine lineare Abbildung

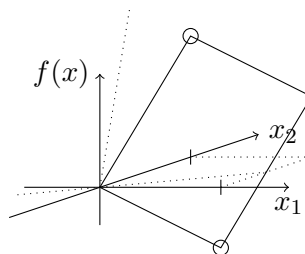
$$A: K^n \rightarrow K^m, x \mapsto A \cdot x \quad (\in \text{Hom}_K(K^n, K^m)).$$

Das erste Beispiel ist der Spezialfall $A \in \mathbb{R}^{1 \times 1}$.

Gut vorstellbar ist auch $A \in \mathbb{R}^{1 \times 2}$, also etwa konkret für $A = [-\frac{1}{2} \ 1]$ die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto f(x) = -\frac{1}{2}x_1 + x_2.$$

Wieder entspricht A der konstanten Steigung der Funktion Ax .



- Die Abbildung, die Polynomen den Vektor ihrer Koeffizienten zuordnet, ist linear

$$F: K[x] \rightarrow K^{\mathbb{N}}, f = a_0 + \dots + a_r x^r \mapsto \begin{pmatrix} a_0 \\ \vdots \\ a_r \\ 0 \\ \vdots \end{pmatrix} \quad (\in \text{Hom}_K(K[x], K^{\mathbb{N}})),$$

ebenso $F_d: K[x] \rightarrow K^d, f \mapsto \begin{pmatrix} a_0 \\ \vdots \\ a_d \end{pmatrix}$ ($\in \text{Hom}_K(K[x], K^d)$), die Projektion auf die ersten d Koeffizienten.

- Ist U ein Untervektorraum von V und $F: V \rightarrow W$ linear, dann ist die Einschränkung $F|_U: U \rightarrow W, u \mapsto F(u)$ linear.
- Differenzieren ist eine lineare Abbildung.

Zur Erinnerung: Für eine Menge M und einen K -Vektorraum V ist die Menge Abbildungen $V^M = \text{Abb}(M, V)$ mit punktweiser Addition $(f+g)(m) := f(m) + g(m)$ und Skalarmultiplikation $(\alpha f)(m) = \alpha \cdot f(m)$ ein K -Vektorraum.

Die Menge der stetig differenzierbaren Funktionen $C^1(\mathbb{R})$ ($f: \mathbb{R} \rightarrow \mathbb{R}$ stetig diffbar) und die Menge der stetigen Funktionen $C^0(\mathbb{R})$ ($f: \mathbb{R} \rightarrow \mathbb{R}$) sind Vektorräume.

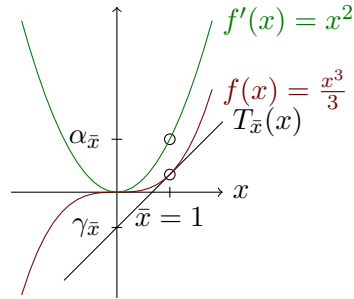
Das Ableiten von einer differenzierbaren Funktion ist eine Abbildung

$$\begin{aligned} \frac{d}{dx}: C^1(\mathbb{R}) &\rightarrow C^0(\mathbb{R}) \\ f &\mapsto f' = \frac{d}{dx} f \end{aligned}$$

Dabei gibt $f'(\bar{x})$ die Steigung in \bar{x} an, und die Tangente in \bar{x} ist der Graph von

$$T_{\bar{x}}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}).$$

$T_{\bar{x}}$ wird **Linearisierung** von f in \bar{x} genannt, denn $T_{\bar{x}}(x) = \alpha_{\bar{x}}x + \gamma_{\bar{x}}$ ist eine „um eine Konstante verschobene lineare“ (=affine) Funktion mit konstantem Anteil $\gamma_{\bar{x}} = f(\bar{x}) - f'(\bar{x})\bar{x}$ und linearem Anteil $\alpha_{\bar{x}}x = f'(\bar{x})x$.



Das Ableiten an sich ist selbst eine lineare Funktion vom \mathbb{R} -Vektorraum $C^1(\mathbb{R})$ in den \mathbb{R} -Vektorraum $C^0(\mathbb{R})$ (also $\in \text{Hom}_{\mathbb{R}}(C^1(\mathbb{R}), C^0(\mathbb{R}))$):

$$\text{für } f, g \in C^1(\mathbb{R}), \alpha, \beta \in \mathbb{R} \text{ gilt } \frac{d}{dx}(\alpha f + \beta g) = \alpha \left(\frac{d}{dx} f\right) + \beta \left(\frac{d}{dx} g\right).$$

Konkret ist etwa für $f(x) = x^3, g(x) = \sin(x), \alpha = 2, \beta = -1$

$$\frac{d}{dx}(2x^3 - \sin(x)) = 2 \frac{d}{dx}(x^3) - \frac{d}{dx} \sin(x) = 6x^2 - \cos(x).$$

Für Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto f(x_1, \dots, x_n)$ lernt man später in der Analysis „partielle Ableitungen“ kennen:

$$\frac{\partial}{\partial x_i} f := \frac{d}{d\xi} f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n).$$

Das Bilden des Gradienten

$$\begin{aligned} \nabla: C^1(\mathbb{R}^n) &\rightarrow C^0(\mathbb{R}^n, \mathbb{R}^n) \\ f &\mapsto \nabla f := \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) \end{aligned}$$

(hier als Zeile, sonst meist als Spalte) ist eine lineare Abbildung, also $\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g$. In einem Punkt $\bar{x} \in \mathbb{R}^n$ ausgewertet beschreibt der Gradient $\nabla f(\bar{x})$ die Steigung der Tangentialebene an den Graph von f in \bar{x} , die Tangentialebene ist also der Graph der affinen Funktion $T_{\bar{x}}(x) = f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) = Ax + b$. Dabei ist wieder $A = \nabla f(\bar{x}) \in M(1 \times n; \mathbb{R})$ die „Steigung“ des linearen Anteils und $b = f(\bar{x}) - A\bar{x}$ der konstante Anteil.

Für Funktionen

$$F: \mathbb{R}^n \rightarrow \mathbb{R}^m \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}, \quad f_i \in C^1(\mathbb{R}^n), i = 1 \dots m,$$

bildet man als Ableitung die „Jacobi-Matrix“,

$$J: C^1(\mathbb{R}^n, \mathbb{R}^m) \rightarrow C^0(\mathbb{R}^n, \mathbb{R}^{m \times n})$$

$$F \mapsto J_F := \begin{pmatrix} \nabla f_1 \\ \vdots \\ \nabla f_m \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Wieder ist $J_{\alpha F + \beta G} = \alpha J_F + \beta J_G$, also J eine lineare Abbildung und in einem Punkt $\bar{x} \in \mathbb{R}^n$ ausgewertet beschreibt die Matrix $A = J_F(\bar{x}) \in M(m \times n; \mathbb{R})$ die Steigung der Linearisierung von F in \bar{x} .

- Integrieren ist eine lineare Abbildung (die additive Konstante vergessen wir hier einmal)

$$\int: C^0(\mathbb{R}) \rightarrow C^1(\mathbb{R})$$

$$f \mapsto \int f dx,$$

denn für $f, g \in C^0(\mathbb{R}), \alpha, \beta \in \mathbb{R}$ ist $\int(\alpha f + \beta g)dx = \alpha \int f dx + \beta \int g dx$. Die konkrete Berechnung bestimmter Integrale ist linear,

$$\int_0^1: C^0(\mathbb{R}) \rightarrow \mathbb{R}$$

$$f \mapsto \int_0^1 f dx.$$

Auch das lässt sich wieder auf höhere Dimensionen ausdehnen.

♡

Ganz in Analogie zu den Gruppenhomomorphismen (B 2.10) gilt für die (Vektorraum-)Homomorphismen:

Beobachtung 4.2 Ist $F: V \rightarrow W$ linear ($F \in \text{Hom}_K(V, W)$), so gilt

- (a) $F(0) = 0$ und $F(v - w) = F(v) - F(w)$, insbes. $F(-v) = -F(v)$,
- (b) $F(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 F(v_1) + \cdots + \alpha_n F(v_n)$,
- (c) Ist F ein Isomorphismus, dann auch $F^{-1}: W \rightarrow V$.

Beweis: s. Vorlesung

□

Zwei K -Vektorräume V, W heißen isomorph, in Zeichen $V \simeq W$, wenn es einen Isomorphismus $F: V \rightarrow W$ gibt.

Beispiele:

- Ist V ein K -Vektorraum mit endlicher Basis $\mathcal{B} = (v_1, \dots, v_r)$, dann ist

$$\mathcal{K}_{\mathcal{B}}: V = \text{span } \mathcal{B} \rightarrow K^r$$

$$v = \sum_{i=1}^r \alpha_i v_i \mapsto \sum_{i=1}^r \alpha_i e_i \quad (\text{Koordinaten } \alpha_i \text{ von } v \text{ bzgl. } \mathcal{B})$$

ein Vektorraumisomorphismus, denn $\mathcal{K}_{\mathcal{B}}$ erfüllt (L) nach den Rechengesetzen, und nach S 3.13(c) ist jedes $v \in V, w \in K^r$ eindeutig als Linearkombination der v_i bzw. e_i darstellbar.

$$\Phi_{\mathcal{B}} := \mathcal{K}_{\mathcal{B}}^{-1}: K^r \rightarrow V, \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} \mapsto \sum_{i=1}^r a_i v_i \text{ heißt } \underline{\text{Koordinatensystem}}.$$

Merke: Jeder endlichdimensionale K -Vektorraum V ist isomorph zum K^r mit $r = \dim V$.

- Ist V ein K -Vektorraum mit *unendlicher* Basis $\mathcal{B} = (v_i)_{i \in \mathcal{J}}$, dann ist

$$F_{\mathcal{B}}: V = \text{span } \mathcal{B} \rightarrow K^{\mathcal{J}}$$

$$v = \sum_{\substack{i \in E \subseteq \mathcal{J} \\ |E| \text{ endlich}}} \alpha_i v_i \mapsto \sum_{i \in E} \alpha_i e_i$$

zwar linear ($\ddot{\text{U}}$), aber *kein* Vektorraumisomorphismus, denn

$$\text{span}(e_i)_{i \in \mathcal{J}} \neq K^{\mathcal{J}} \quad (\text{z. B. } \begin{pmatrix} 1 \\ \vdots \end{pmatrix} \notin \text{span}(e_i)_{i \in \mathcal{J}}).$$

ABER: $\text{span}(e_i)_{i \in \mathcal{J}}$ ist ein Untervektorraum von $K^{\mathcal{J}}$ und

$$F'_{\mathcal{B}}: V = \text{span } \mathcal{B} \rightarrow \text{span}(e_i)_{i \in \mathcal{J}}$$

$$v = \sum_{\substack{i \in E \subseteq \mathcal{J} \\ |E| \text{ endlich}}} \alpha_i v_i \mapsto \sum_{i \in E} \alpha_i e_i \quad \text{ist ein Isomorphismus.}$$

Umgekehrt hätte man beim Ansatz $f \in K^{\mathcal{J}} \rightarrow \sum_{i \in \mathcal{J}} f(i)v_i$ das Problem, dass die unendliche Summe i. A. kein Ergebnis liefert. In unendlichdimensionalen Vektorräumen sind Konvergenzeigenschaften meist wichtiger als Basisdarstellungen, s. Funktionalanalysis.

- Für beliebige Mengen M ist $(2^M, \Delta)$ über \mathbb{F}_2 isomorph zu \mathbb{F}_2^M ($\ddot{\text{U}}$), aber i. A. ist eine Basis schwer anzugeben.

♡

Wie alle Abbildungen kann man auch lineare Abbildungen selbst addieren und skalar multiplizieren, sie bilden damit selbst einen Vektorraum.

Beobachtung 4.3 Für K -Vektorräume V, W ist $\text{Hom}_K(V, W)$ ein Untervektorraum von $\text{Abb}(V, W)$.

Beweis: s. Vorlesung □

Lineare Abbildungen können problemlos hintereinandergeschaltet werden.

Beobachtung 4.4 Seien U, V, W K -Vektorräume, $G \in \text{Hom}_K(U, V)$ und $F \in \text{Hom}_K(V, W)$, dann ist $H = F \circ G \in \text{Hom}_K(U, W)$.

$$\begin{array}{ccc} U & \xrightarrow{F \circ G} & W \\ G \searrow & & \nearrow F \\ & V & \end{array}$$

Beweis: s. Vorlesung □

Ü: Zeige: Für $F, F' \in \text{Hom}_K(V, W)$, $G, G' \in \text{Hom}_K(U, V)$ gilt
 $(\alpha F + \beta F') \circ G = \alpha F \circ G + \beta F' \circ G$ und $F \circ (\alpha G + \beta G') = \alpha F \circ G + \beta F \circ G'$.
 Die Komposition linearer Abbildungen ist also linear im ersten und linear
 im zweiten Argument. Man sagt dazu kurz, sie ist bilinear. x

Mit der Komposition wird $\text{End}_K(V) = \text{Hom}_K(V, V)$ zu einem Ring:

Beobachtung 4.5 *Ist V ein K -Vektorraum, so ist $(\text{End}_K(V), +, \circ)$ ein
 (i. A. nicht kommutativer) Ring mit Eins.*

Beweis: s. Vorlesung □

Nullelement des Rings: Nullabbildung $\mathbb{0}: v \mapsto 0$

Einselement des Rings: Identität: $\text{id}_V: v \mapsto v$ (für $V = \{0\}$ ist $\text{id}_V = \mathbb{0}$)

Ein wichtiger Spezialfall des Vektorraums $\text{Hom}_K(V, W)$ ist der Vektorraum
 der linearen Abbildungen $\text{Hom}_K(V, K)$, die von V in den Körper von V
 abbilden.

Beispiel: Für $V = K^n$ und $A \in K^{m \times n}$ beschreibt jede einzelne Zeile
 $A_{i, \bullet}$ eine Abbildung $A_{i, \bullet} \cdot x$ aus $\text{Hom}_K(K^n, K)$ und wir haben beobachtet,
 dass der Lösungsraum $\text{Lös}(A, 0) \subseteq V$ durch den Untervektorraum
 $\text{span}(A_{i, \bullet})_{i \in \{1, \dots, m\}} \subseteq \text{Hom}_K(V, K)$ festgelegt wird. Die Vektorraumstrukturen
 von V und $\text{Hom}_K(V, K)$ sind eng miteinander verbunden. ♥

Man nennt $V^* := \text{Hom}_K(V, K)$ den Dualraum zu V .

$F \in V^*$ ($F \in \text{Hom}_K(V, K)$) heißt Linearform oder lineares Funktional.

Beispiele:

- Für $A \in K^{m \times n}$ bildet in Ax jede Zeile $A_{i, \bullet}: K^n \rightarrow K$, $x \mapsto A_{i, \bullet} \cdot x$ eine
 Linearform.
- $\int_0^1: C^0(\mathbb{R}) \rightarrow \mathbb{R}$, $f \mapsto \int_0^1 f(x)dx$ ist ein lineares Funktional, $\int_0^1 \in$
 $\text{Hom}_{\mathbb{R}}(C^0(\mathbb{R}), \mathbb{R})$ (denn \int ist linear, $\int(\alpha f + \beta g)dx = \alpha \int f dx + \beta \int g dx$).
- Für jedes $f \in C^1(\mathbb{R}^n)$ ($f: \mathbb{R}^n \rightarrow \mathbb{R}$) und jedes $\bar{x} \in \mathbb{R}^n$ ist $\nabla f(\bar{x})x$ die
 lineare Funktion mit der konstant gleichen Steigung wie f in \bar{x} , also
 kann auch $\nabla f(\bar{x})$ als Linearform in $\text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ interpretiert werden.

♥

4.2 Bild, Fasern und Kern

Wir verwenden die gleichen Bezeichnungen wie für Gruppenhomomorphismen
 für $F \in \text{Hom}_K(V, W)$:

$\text{Im}(F) = F(V) = \{F(v) : v \in V\} \subseteq W$ das Bild von F ,

$\text{Ker}(F) = F^{-1}(\{0\}) = \{v \in V : F(v) = 0\} \subseteq V$ der Kern von F .

Beobachtung 4.6 Sei $F \in \text{Hom}_K(V, W)$.

- (a) Sind $V' \subseteq V$ und $W' \subseteq W$ Untervektorräume, so auch $F(V') \subseteq W$ und $F^{-1}(W') = \{v \in V : F(v) \in W'\}$. Insbesondere sind $\text{Im } F$ und $\text{Ker } F$ Untervektorräume.
- (b) Ist $(v_i)_{i \in \mathcal{J}}$ eine l.a. Familie in V , so ist $(F(v_i))_{i \in \mathcal{J}}$ l.a. in W .
- (c) $\dim(\text{Im } F) \leq \dim(V)$. Mann nennt $\dim(\text{Im } F)$ den Rang von F .

Beweis: s. Vorlesung □

Beobachtung 4.7 Für $F \in \text{Hom}_K(V, W)$ gilt

- (a) F surjektiv $\Leftrightarrow \text{Im } F = W$,
- (b) F injektiv $\Leftrightarrow \text{Ker } F = \{0\}$,
- (c) Ist F injektiv und $(v_i)_{i \in \mathcal{J}}$ l.u., dann ist auch $(F(v_i))_{i \in \mathcal{J}}$ l.u.

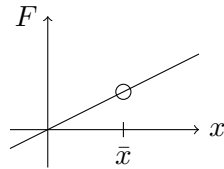
Beweis: s. Vorlesung □

Wenn F nicht injektiv ist, wird (c) i. A. nicht gelten, z. B. $F: K^n \rightarrow \{0\}$, $n > 0$.

Eine lineare Abbildung ist bereits durch wenige Vorgaben festgelegt.

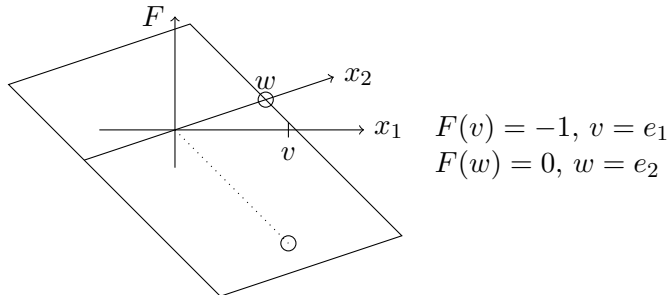
Beispiele:

- Für $F: \mathbb{R} \rightarrow \mathbb{R}$ linear kann man nur einen Wert vorgeben.



(Der Graph einer linearen Abb. muss immer durch 0 gehen und konstante Steigung haben!)

- Für $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ sind es maximal 2 Werte für linear unabh. Punkte.



♡

Eine lineare Abb. ist schon eindeutig durch ihre Werte auf einer Basis beschrieben.

Satz 4.8 (Existenz linearer Abb. bei vorgegebenen Werten) Für K -Vektorräume V, W seien Vektoren $(v_i)_{i \in \mathcal{J}}$ aus V und $(w_i)_{i \in \mathcal{J}}$ aus W gegeben.

(1) Ist $(v_i)_{i \in \mathcal{J}}$ eine Basis von V , dann gibt es genau ein $F \in \text{Hom}_K(V, W)$ mit $F(v_i) = w_i$ für $i \in \mathcal{J}$.

Dieses F hat folgende Eigenschaften:

- (a) $\text{Im } F = \text{span}(w_i)_{i \in \mathcal{J}}$,
- (b) F injektiv $\Leftrightarrow (w_i)_{i \in \mathcal{J}}$ l.u.

(2) Ist $(v_i)_{i \in \mathcal{J}}$ l.u. und zu einer Basis von V ergänzbar, dann gibt es mindestens ein $F \in \text{Hom}_K(V, W)$ mit $F(v_i) = w_i$ für $i \in \mathcal{J}$.

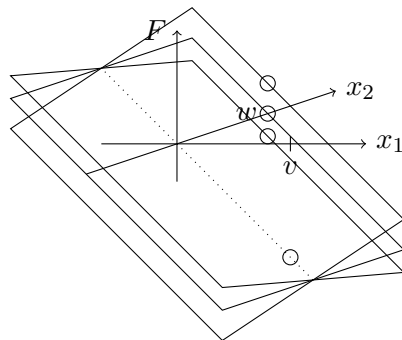
Merke: Stimmen zwei lineare Funktionen auf einer Basis überein, so sind sie identisch.

Beweis: s. Vorlesung □

Für endlichdimensionales V , $\dim V = n < \infty$, sieht man im Beweis von (2), dass für $|J| = r (\leq n)$ die Abbildung F genau $n - r$ Vorgaben von der Eindeutigkeit „entfernt“ ist.

Beispiele:

- Gibt man für $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ nur einen Wert vor, gibt es viele Möglichkeiten:



$$F(v) = -1, v = e_1,$$

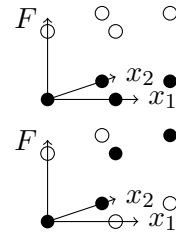
$$F(w) \in \{0, \frac{4}{10}, \frac{-3}{10}\}, w = e_2$$

(alle Ebenen, die die gepunktete Gerade enthalten, außer der vertikalen).

- Für $F: \mathbb{F}_2^2 \rightarrow \mathbb{F}_2$ gibt es nicht so viele Möglichkeiten:
 $\mathbb{F}_2^2 = \{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix})\}$. Wenn wir nur einen Wert vorgeben, z. B. $F(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}) = 0$, können wir uns für einen zweiten l.u. Vektor, z. B. $(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix})$ nur 2 Werte aussuchen, $F(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}) \in \{0, 1\}$, für jeden bekommen wir eine lineare Abbildung:

$F(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}) = 0$ die 0-Abb.,
 in Matrixform $Ax = [0 \ 0]x$,

$F(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}) = 1$,
 in Matrixform $Ax = [1 \ 0]x$ mit Werten
 $A(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}) = 1, A(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}) = 0, A(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}) = 1$.



♡

Schon für Gruppenhomomorphismen $\varphi: (V, +) \rightarrow (W, \tilde{+})$ ist für jedes $w \in \text{Im } \varphi$ und beliebiges $v \in \varphi^{-1}(\{w\})$ die Faser von φ über w von der Form

$\varphi^{-1}(\{w\}) = v + \text{Ker } \varphi$, wobei $\text{Ker } \varphi$ Normalteiler (also Untergruppe U mit $a + U = U + a$) und $v + \text{Ker } \varphi$ Rest-/Nebenklasse von v modulo $\text{Ker } \varphi$, vgl. B.2.13 und die Diskussion danach.

Ist nun $F \in \text{Hom}_K(V, W)$ sogar Vektorraumhomomorphismus, dann ist $\text{Ker } F$ zusätzlich Untervektorraum (B4.6(a)) und für $w \in \text{Im } F$ und beliebiges $v \in F^{-1}(\{w\})$ die Faser von $F^{-1}(\{w\}) = v + \text{Ker } F$ ein affiner Unterraum.

Diese Struktur zeigt sich besonders gut, wenn man die Basen von V und W passend zu $\text{Ker } F$ bzw. $\text{Im } F$ wählt.

Satz 4.9 (Dimensionsformel für lineare Abb.) *Ist $F: V \rightarrow W$ linear mit $\dim V < \infty$, so ist für beliebige Basen (v_1, \dots, v_k) von $\text{Ker } F$, (w_1, \dots, w_r) von $\text{Im } F$ und beliebige $u_1 \in F^{-1}(\{w_1\}), \dots, u_r \in F^{-1}(\{w_r\})$*

$$\mathcal{A} := \{u_1, \dots, u_r, v_1, \dots, v_k\} \text{ eine Basis von } V$$

und es gilt:

$$\dim V = \dim \text{Im } F + \dim \text{Ker } F.$$

Beweis: s. Vorlesung □

Korollar 4.10 *Ist V endlichdimensional und $F \in \text{Hom}_K(V, W)$, dann gilt für alle nichtleeren Fasern $\dim F^{-1}(\{w\}) = \dim V - \dim \text{Im } F$.*

Beispiel: Ist $a \in K^{1 \times n}$, also z. B. eine Zeile einer Matrix, dann ist ax eine Linearform, also gleichzeitig $a \in \text{Hom}_K(K^n, K)$. Der Kern von a ist der Untervektorraum $\text{Lös}(a, 0)$. Wegen $\dim \text{Ker } a = \dim \text{Lös}(a, 0) = \dim K^n - \dim \text{Im } a$ ist dieser für $a = 0$ der gesamte Raum V und für $a \neq 0$ eine Hyperebene in V durch 0.

Anschaulich hat a den Graph $\left\{ \begin{pmatrix} x \\ ax \end{pmatrix} : x \in K^n \right\}$, dieser ist eine nicht vertikale Hyperebene in K^{n+1} durch den Ursprung. Der Schnitt dieser Hyperebene in K^{n+1} mit der Hyperebene $K^n \times \{0\}$ beschreibt den Kern $ax = 0$. Nur für $a = 0$ fallen Graph und $K^n \times \{0\}$ zusammen. ♡

Korollar 4.11 *Zwischen zwei endlichdimensionalen K -Vektorräumen gibt es genau dann einen Isomorphismus, wenn $\dim V = \dim W$.*

Beweis: s. Vorlesung □

Korollar 4.12 *Ist $\dim V = \dim W < \infty$ und $F \in \text{Hom}_K(V, W)$, so sind äquivalent:*

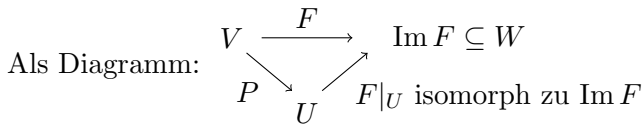
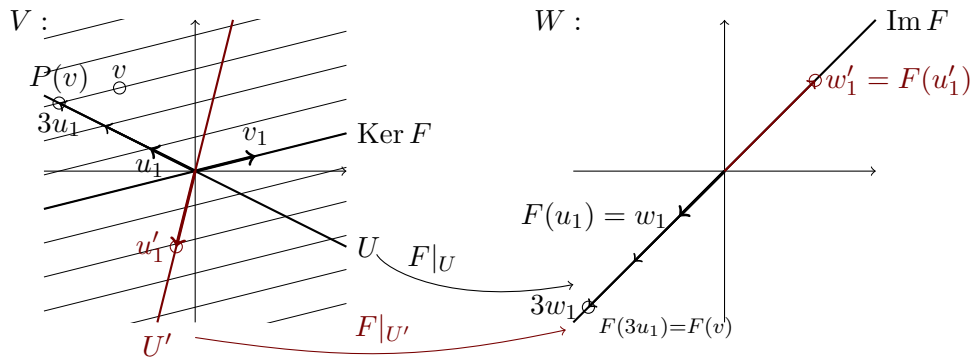
- (a) F ist injektiv,
- (b) F ist surjektiv,
- (c) F ist bijektiv.

Beweis: s. Vorlesung □

Wegen S4.9 haben lineare Abbildungen $F: V \rightarrow W$ mit endlichdimensionalem Vektorraum V stets folgende Struktur:

- Es gibt (mindestenes) einen Untervektorraum $U (= \text{span}(u_i)_{i=1}^r)$ von V der sich Eins-zu-Eins/isomorph auf den Untervektorraum $\text{Im } F$ in W abbildet.
- V ist die disjunkte Vereinigung der affinen Unterräume $u + \text{Ker } F$ für $u \in U$ (U bestimmt hier die Repräsentanten). Diese bilden die Faserung von F bilden.

Für jedes $v \in V$ gibt es genau ein $u \in U$ mit $v \in u + \text{Ker } F$, also mit $F(v) = F(u)$. Das definiert eine Projektion $P: V \rightarrow U$ von V auf den Untervektorraum U .



$(F|_U)^{-1}: \text{Im } F \rightarrow U$ nennt man Schnitt, denn das Bild U schneidet aus jeder Faser genau einen Punkt.

Satz 4.13 (Faktorisierungssatz)

Sei $F \in \text{Hom}_K(V, W)$ und $\mathcal{A} = (u_1, \dots, u_r, v_1, \dots, v_k)$ eine Basis von V mit $\text{Ker } F = \text{span}(v_1, \dots, v_k)$. Für $U = \text{span}(u_1, \dots, u_r)$ gilt:

- $V = U \oplus \text{Ker } F$.
- Die Einschränkung $F|_U: U \rightarrow \text{Im } F, u \mapsto F(u)$ ist ein Isomorphismus.
- Bezeichnet $P: V = U \oplus \text{Ker } F \rightarrow U, u + v' \mapsto u$ die Projektion auf den ersten Summanden, so ist $F = (F|_U) \circ P$. Insbesondere hat jede nichtleere Faser $F^{-1}(\{w\})$ genau einen Schnittpunkt mit U und $P(v) = F^{-1}(\{F(v)\}) \cap U$.

Beweis: s. Vorlesung □

Die Konstruktion funktioniert für jedes U mit $V = U \oplus \text{Ker } F$, aber obwohl die Wahl nicht wesentlich ist, hängt die Abbildung $F|_U$, die die wesentliche Bildinformation von F trägt, stark von U ab.

Ordnen wir den Wert direkt den Nebenklassen, hier also den affinen Unterräumen $u + \text{Ker } F$ bzw. den Elementen des Quotientenvektorraums $V / \text{Ker } F$,

zu, verschwindet die Abhängigkeit und $F' : V/\text{Ker } F \rightarrow \text{Im } F$, $v + \text{Ker } F \mapsto F(v)$ ist ein Isomorphismus unabhängig von der Wahl von U .

Ist nun N ein Untervektorraum von V , dann hat ganz allgemein der Quotientenvektorraum V/N folgende sogenannte „universelle Eigenschaft“: Zusammen mit der kanonischen Abbildung $\rho : V \rightarrow V/N$, $v \mapsto v + N$ gibt es für jede linear Abb. $F : V \rightarrow W$ mit $N \subseteq \text{Ker } F$ genau eine lineare Funktion $\bar{F} : V/N \rightarrow W$, sodass das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ & \searrow \rho & \nearrow \bar{F} \\ & V/N & \end{array} \quad \text{kommutiert.}$$

Falls $N = \text{Ker } F$, ist \bar{F} injektiv und weist jeder Faser von F ihren Wert zu.

Satz 4.14 (Homomorphiesatz für Vektorräume)

Sei V ein K -Vektorraum und $N \subseteq V$ ein Untervektorraum. Nur für den Quotientenvektorraum $(V/N, \hat{+}, \hat{\cdot})$ wird die kanonische Abbildung

$$\rho : V \rightarrow V/N, v \mapsto v + N$$

linear. Es gilt

- (a) ρ is surjektiv.
- (b) $\text{Ker } \rho = N$.
- (c) Ist $\dim V < \infty$, so gilt $\dim V/N = \dim V - \dim N$.
- (d) Der Quotientenvektorraum V/N hat folgende universelle Eigenschaft: Ist $F \in \text{Hom}_K(V, W)$ mit $N \subseteq \text{Ker } F$, so gibt es genau ein $\bar{F} \in \text{Hom}_K(V/N, W)$ mit $F = \bar{F} \circ \rho$. Weiter ist $\text{Ker } \bar{F} = (\text{Ker } F)/N$.

Beweis: s. Vorlesung □

Beachte, nur in (c) wird die Endlichkeit von $\dim V$ vorausgesetzt. Im Gegensatz zu Dimensionsformel und Faktorisierungssatz gilt der Homomorphiesatz sonst ganz allgemein. Kennt man eine Zerlegung von V in eine direkte Summe $V = U \oplus N$ wie im endlichdimensionalen Fall, dann kann man V/N durch U ersetzen wie schon im Faktorisierungssatz.

Satz 4.15 Ist $V = U \oplus N$ und $\rho : V \rightarrow V/N$ die kanonische Abbildung, dann ist $\rho' := \rho|_U : U \rightarrow V/N$ ein Isomorphismus.

Beweis: s. Vorlesung □

Im unendlichdimensionalen ist das nicht unbedingt hilfreich, weil man meist kein passendes U kennt.

Beispiele:

- Vektorraum $V = C(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} \text{ stetig}\}$. Ist f nur auf einem Teilbereich $X \subseteq \mathbb{R}$ wichtig, betrachten wir zwei Funktionen $f, g \in C(\mathbb{R})$ genau dann als „gleich“, wenn ihre Werte auf X übereinstimmen:
 $f \sim g :\Leftrightarrow f - g =: h \in \mathcal{I}(X) := \{h \in C(\mathbb{R}) : h(x) = 0 \text{ für } x \in X\}$

Ü: $\mathcal{I}(X)$ ist ein Untervektorraum von $C(\mathbb{R})$. x

$\bar{f} = f + \mathcal{I}(X) \in C(\mathbb{R})/\mathcal{I}(X)$ ist die Menge aller stetigen Funktionen, die auf X den gleichen Wert wie f haben. (So ein \bar{f} kann man z. B. als stetige Funktion auf einer abgeschlossenen Menge X definieren).

Wie verhält sich $C(\mathbb{R})/\mathcal{I}(X)$ zum Vektorraum $\mathbb{R}^X = \{\varphi: X \rightarrow \mathbb{R}\}$?

Sei dazu $\sigma: C(\mathbb{R}) \rightarrow \mathbb{R}^X, f \mapsto f|_X$ der Einschränkungshomomorphismus.

Setze $C(X) := \text{Im } \sigma = \{\varphi \in \mathbb{R}^X: \exists f \in C(\mathbb{R}) \text{ mit } \varphi = f|_X\}$,

„die auf \mathbb{R} stetig fortsetzbaren Funktionen“,

dann ist $\text{Ker } \sigma = \mathcal{I}(X)$ (diese werden auf die 0-Funktion abgebildet)

$$\text{und } \begin{array}{ccc} C(\mathbb{R}) & \xrightarrow{\sigma} & \mathbb{R}^X \\ \rho \searrow & & \nearrow \bar{\sigma} \\ & C(\mathbb{R})/\mathcal{I}(X) & \end{array} \quad \text{mit } \text{Ker } \bar{\sigma} = \text{Ker } \sigma/\mathcal{I}(X) = \{0\} \text{ (hier } 0 = \mathcal{I}(X)\text{!)},$$

also ist $C(\mathbb{R})/\mathcal{I}(X)$ isomorph zum Unterraum $C(X)$ von \mathbb{R}^X .

- Ähnliches geht für integrierbare Funktionen (Riemann, besser Lebesgue)

$$\mathcal{L}(\mathbb{R}) := \{f: \mathbb{R} \rightarrow \mathbb{R} \text{ integrierbar}\}$$

mit $\mathcal{N} := \{f \in \mathcal{L}(\mathbb{R}): \int_{\mathbb{R}} |f(x)| dx = 0\} \subset \mathcal{L}(\mathbb{R})$.

Für $f, g \in \mathcal{N}$ und $\alpha, \beta \in \mathbb{R}$ ist auch $(\alpha f + \beta g) \in \mathcal{N}$ [Eigenschaft des Integrals!], damit ist \mathcal{N} ein Untervektorraum. Bei Lebesgue-Integration enthält \mathcal{N} alle Funktionen, die „fast überall“ gleich 0 sind, und \mathcal{N} ist unendlichdimensional (z. B. $\uparrow \overset{\cdot}{\times} \rightarrow$ für jedes $x \in \mathbb{R}$).

$$f \sim_{\mathcal{N}} g \Leftrightarrow \int_{\mathbb{R}} |f(x) - g(x)| dx = 0 \quad f \text{ ist „fast überall“ gleich } g.$$

Ein $\bar{f} \in \mathcal{L}(\mathbb{R})/\mathcal{N}$ ist die Äquivalenzklasse der fast überall zu f gleichen Funktionen und $L(\mathbb{R}) := \mathcal{L}(\mathbb{R})/\mathcal{N}$ ist als Quotientenvektorraum selbst ein Vektorraum.

Es gibt dafür aber keine ähnlich konstruktive Beschreibung wie für $C(X)$.

♡

4.3 Der Dualraum $V^* = \text{Hom}_K(V, K)$

Die linearen Funktionale/Linearformen $\varphi: V \rightarrow K$ bilden mit ihrem eindimensionalen Bildraum so etwas wie die Grundbausteine der linearen Funktionen. Vorrangig ist $V^* = \text{Hom}_K(V, K)$ auch einfach nur ein Vektorraum, in dem die Vektoren nun sehr spezielle Funktionen sind, die V linear in K abbilden. Wie drückt sich diese Kopplung an V aus, wenn man sich Gedanken über Basen, Dimension oder Untervektorräume von V^* macht?

Ist es möglich, gut aufeinander abgestimmte Basen für V und seine Linearformen V^* zu finden, gewinnt man besonders vertändliche und effiziente

Darstellungen aller linearen Funktionale auf V , die auch das strukturelle Wechselspiel z. B. zwischen linearen Abbildungen und Kern besonders deutlich machen.

Beobachtung 4.16 Sei V ein endlichdimensionaler K -Vektorraum.

- (a) Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , so ist $\mathcal{B}^* = (\varphi_1, \dots, \varphi_n)$ mit $\varphi_i(v_j) = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$ eine Basis von V^* , die zu \mathcal{B} duale Basis \mathcal{B}^* .
- (b) Ist $\mathcal{A} = (\psi_1, \dots, \psi_n)$ eine Basis von V^* , so gibt es eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V mit dualer Basis $\mathcal{B}^* = \mathcal{A}$.

Beweis: s. Vorlesung □

Für n -dimensionales V ist V^* daher isomorph zu V ! Das bedeutet auch, dass lineare Funktionale eine sehr kleine Teilmenge aller mögliche Abbildungen K^V bilden, denn $|V| < |2^V| \leq |K^V|$. Im endlichdimensionalen haben sie nicht nur genau die gleiche Dimension n , es gibt auch zu jeder Basis des einen eine besonders gut passende Basis des anderen. Seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}^* = (\varphi_1, \dots, \varphi_n)$ so ein Paar, dann ist jede lineare Abbildung $\varphi: V \rightarrow K$ einfach durch ihre Koordinatendarstellung $\varphi = \sum_{i=1}^n \gamma_i \varphi_i$ beschreib- und auswertbar. Denn ist $c = [\gamma_1, \dots, \gamma_n] \in K^{1 \times n}$ die Zeile der Koordinaten von φ bezüglich \mathcal{B}^* und enthält für $v = \sum_{i=1}^n \xi_i v_i \in V$ der Spaltenvektor $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ die Koordinaten von v bezüglich \mathcal{B} , dann ist $\varphi(v) = cx$, weil

$$\varphi(v) = \sum_{i=1}^n \gamma_i \varphi_i(v) = \sum_{i=1}^n \gamma_i \varphi_i\left(\sum_{j=1}^n \xi_j v_j\right) = \sum_{i=1}^n \sum_{j=1}^n \gamma_i \xi_j \underbrace{\varphi_i(v_j)}_{=\delta_{ij}} = \sum_{i=1}^n \gamma_i \xi_i = cx.$$

Für n -dimensionales V beschreiben also Matrixzeilen die gesamte Welt der linearen Funktionale. Zwischen den Vektorräumen V und V^* besteht anscheinend im endlichdimensionalen kein wesentlicher Unterschied, wie zwei Seiten der gleichen Medaille.

Beispiel: Für $V = \mathbb{R}^n$ mit kanonischer Basis $\mathcal{B} = (e_i)_{i=1}^n$ ist die duale Basis $\mathcal{B}^* = (\varphi_i)_{i=1}^n$ formal durch $\varphi_i(e_j) = \delta_{ij}$ beschrieben, aber wegen $\varphi_i(e_j) = [0, \dots, 0, \underset{i}{1}, 0, \dots, 0]e_j = \delta_{ij} \forall ij$ dürfen wir φ_i einfach mit dem Zeilenvektor zu e_i identifizieren.

Wenn wir für \mathbb{R}^2 die Basis $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ wählen, welche Zeilenvektoren bilden dann \mathcal{B}^* ? ♡

Vorsicht! Für $K^{\mathbb{N}}$ funktionieren diese Argumente schon nicht mehr, denn $\text{span}(e_i)_{i \in \mathbb{N}} \neq K^{\mathbb{N}}$. Ebenso ist auf $V = \text{span}(e_i)_{i \in \mathbb{N}}$ die Funktion

$$\varphi: \text{span}(e_i)_{i \in \mathbb{N}} \rightarrow K, \quad x \mapsto \sum_{i \in \mathbb{N}} x_i$$

ein gut erklärtes lineares Funktional $\in V^*$ (warum?), das der unendlich langen 1er Zeile entspräche, aber $\varphi \notin \text{span}(\varphi_i: V \rightarrow K, \sum_{j \in \mathbb{N}} \xi_j e_j \mapsto \xi_i)_{i \in \mathbb{N}}$.

φ ist also nicht im dualen Untervektorraum, der von den „Zeilen zu e_i “ aufgespannt wird. ∇

Wie im Beweis zu (b) und im homogenen Gleichungssystem $Ax = 0$ ist der gemeinsame Kern mehrerer Linearformen von besonderer Bedeutung. Auch das kann man von beiden Seiten aus gleich betrachten. Wenn man eine Teilmenge von V vorgibt, welche Eigenschaften hat der Untervektorraum der Linearformen in V^* , die diese Teilmenge in ihrem Kern enthalten?

Definition 4.17 Sei V ein K -Vektorraum.

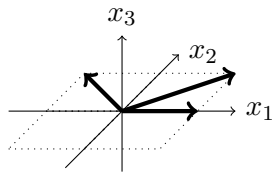
- Für eine Teilmenge $\Phi \subseteq V^*$ heißt $\Phi^\circ := \{v \in V : \varphi(v) = 0 \text{ für alle } \varphi \in \Phi\} = \bigcap_{\varphi \in \Phi} \text{Ker } \varphi$ der Annulator von Φ .
- Für eine Teilmenge $Y \subseteq V$ heißt $Y^\circ := \{\varphi \in V^* : \varphi(v) = 0 \text{ für alle } v \in Y\}$ der Annulator von Y .

[Vgl. die Struktur mit dem Dualitätslemma für Relationen L 1.6!]

Y° sammelt alle Linearformen φ , die $Y \subseteq \text{Ker } \varphi$ und (weil $\text{Ker } \varphi$ immer Untervektorraum ist) damit auch $\text{span } Y \subseteq \text{Ker } \varphi$ erfüllen. Man darf erwarten, dass Annulatoren „zueinander duale“ Untervektorräume $U \subseteq V$ und $U^* \subseteq V^*$ beschreiben, die jeweils all die Vektoren aus V bzw V^* enthalten, für die $\varphi(u) = 0$ für jede beliebige Wahl von $u \in U$ bzw $\varphi \in U^*$ gilt.

Beispiele:

- $\mathbb{R}^3, (\mathbb{R}^3)^* \simeq \mathbb{R}^{1 \times 3}$, sei $\Phi = \{[1 \ 0 \ 0], [1 \ 1 \ 0], [-1 \ 1 \ 0]\}$, dann ist $\Phi^\circ = \{x \in \mathbb{R}^3 : \begin{aligned} [1 \ 0 \ 0]x &= x_1 = 0, \\ [1 \ 1 \ 0]x &= x_1 + x_2 = 0, \\ [-1 \ 1 \ 0]x &= -x_1 + x_2 = 0 \end{aligned}\}$
 $= \left\{ \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \right\}$ „die z -Achse“.



Geometrisch entspricht der Schnitt der Kerne der φ_i hier dem Schnitt der Ebenen von Vektoren, die „orthogonal“ zum jeweiligen Zeilenvektor von φ_i stehen.

- $\mathbb{R}^n, (\mathbb{R}^n)^* \simeq \mathbb{R}^{1 \times n}$, $\Phi = \{a_1, \dots, a_m\} \subseteq \mathbb{R}^{1 \times n}$, $A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \in \mathbb{R}^{m \times n}$
 $\Phi^\circ = \{x \in \mathbb{R}^n : Ax = 0\}$ „Lösungsmenge des homogenen Systems“
 $=$ Schnitt aller Hyperebenen durch 0 mit „Normalvektor“ aus Φ
 $=$ alle Vektoren, die „orthogonal“ auf alle Vektoren aus Φ stehen.
Für $Y \subseteq \mathbb{R}^n$ ist entsprechend $Y^\circ = \{a \in \mathbb{R}^{1 \times n} : ax = 0 \ \forall x \in Y\}$
(in \mathbb{R}^3 : $Y = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$, $Y^\circ = \{[\alpha_1 \ \alpha_2 \ 0] \in \mathbb{R}^{1 \times 3}\} = \text{span}\{[1 \ 0 \ 0], [1 \ 1 \ 0]\}$).

♡

Ganz ohne Forderung an die Dimension gelten folgende Eigenschaften.

Beobachtung 4.18 Für $Y, Y_1, Y_2 \subseteq V$, $\Phi, \Phi_1, \Phi_2 \subseteq V^*$ gilt:

- (a) Φ° ist Untervektorraum von V , Y° ist Untervektorraum von V^* ,
- (b) $\Phi \subseteq (\Phi^\circ)^\circ$, $Y \subseteq (Y^\circ)^\circ$,
- (c) aus $\Phi_1 \subseteq \Phi_2$ folgt $\Phi_1^\circ \supseteq \Phi_2^\circ$, aus $Y_1 \subseteq Y_2$ folgt $Y_1^\circ \supseteq Y_2^\circ$,
- (d) $\Phi^\circ = (\text{span } \Phi)^\circ$, $Y^\circ = (\text{span } Y)^\circ$,
- (e) $(\Phi_1 \cup \Phi_2)^\circ = \Phi_1^\circ \cap \Phi_2^\circ$, $(Y_1 \cup Y_2)^\circ = Y_1^\circ \cap Y_2^\circ$,
- (f) $\text{span } \Phi \subseteq \Phi^{\circ\circ}$, $\text{span } Y \subseteq Y^{\circ\circ}$.

Beweis: s. Vorlesung □

Beispiele:

- Für homogene Gleichungssysteme besagt (d) $\Phi^\circ = (\text{span } \Phi)^\circ$ folgendes:
In $Ax = 0$ mit $A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \in K^{m \times n}$ entspricht jede Zeile einem Funktional $\psi: K^n \rightarrow K$ und für $\Phi = \{\psi_i: i = 1, \dots, m\}$ ist $\Phi^\circ = \text{Lös}(A, 0)$ die Lösungsmenge.

Ist nun $\dim(\text{span } \Phi) = r < m$, braucht man gar nicht alle ψ_i , es reicht eine Basis $(\varphi_1, \dots, \varphi_r)$ aus Φ auszuwählen (= r l.u. Elemente). Schon diese r Gleichungen beschreiben die gleiche Lösungsmenge

$$\Phi^\circ = (\text{span } \Phi)^\circ = (\text{span}(\varphi_1, \dots, \varphi_r))^\circ$$

Wir werden sehen: r ist gerade der Zeilenrang von A .

- Zum \mathbb{R}^3 -Beispiel vorher: $Y = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$,
 $Y^\circ = \{[\alpha_1 \ \alpha_2 \ 0] \in \mathbb{R}^{1 \times 3}\} = \text{span}\{[1 \ 0 \ 0], [0 \ 1 \ 0]\} = \text{span}\{[1 \ 0 \ 0], [1 \ 1 \ 0], [-1 \ 1 \ 0]\}$,
 $Y^{\circ\circ} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 = 0, x_2 = 0 \right\} = \text{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

♡

Bei Existenz einer Basis (immer mit Auswahlaxiom) gilt für Untervektorräume Y immer $Y^{\circ\circ} = Y$. Wir zeigen das für den endlichdimensionalen Fall und benötigen dazu zuerst eine Dimensionsformel.

Satz 4.19 Sei V ein endlichdimensionaler Vektorraum.

Ist U ein Untervektorraum von V , gilt $\dim U + \dim U^\circ = \dim V$.

Ist Γ ein Untervektorraum von V^* , gilt $\dim \Gamma + \dim \Gamma^\circ = \dim V^* (= \dim V)$.

Beweis: s. Vorlesung □

Beispiel: Zum homogenen Gleichungssystem $Ax = 0$ mit $A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in K^{m \times n}$:

hier ist $V = K^n$, $\Gamma = \text{span}(a_i)_{i=1}^m$ und $F: K^n \rightarrow K^m$, $x \mapsto Ax$,

dann ist $\Gamma^\circ = \text{Ker } F$ und

$$\dim V \stackrel{S 4.9}{=} \dim \text{Im } F + \dim \text{Ker } F \stackrel{S 4.19}{=} \dim \Gamma + \dim \Gamma^\circ$$

also ist $\dim \text{Im } F = \dim \Gamma$. Das bedeutet

- (a) Der Rang der durch A definierten Abbildung ist gleich der Dimension des von den Zeilen von A aufgespannten Unterraums.

(b) Sind die Zeilen der Matrix A linear unabhängig, reduziert jede Zeile die Dimension des Lösungsraumes um 1:

Für eine einzelne Gleichung $ax = 0$ mit $a \neq 0$ gilt

$$\underbrace{\dim \operatorname{span} a}_{=1} + \underbrace{\dim(\operatorname{span} a)^\circ}_{=n-1} = n.$$

Daher ist $(\operatorname{span} a)^\circ$ eine Hyperebene durch den Ursprung.

Im \mathbb{R}^3 ist das wieder gut über Normalvektoren vorstellbar. ♡

Korollar 4.20 Sei V ein endlichdimensionaler Vektorraum.

Ist U Untervektorraum von V , gilt $U = U^{\circ\circ}$.

Ist Γ Untervektorraum von V^* , gilt $\Gamma = \Gamma^{\circ\circ}$.

Beweis: s. Vorlesung □

Damit lassen sich Rechengesetze zur Bestimmung des Annulators vom Schnitt von Untervektorräumen herleiten.

Korollar 4.21 Sei V ein endlichdimensionaler Vektorraum.

Sind U_1, U_2 UVRe von V , gilt $(U_1 \cap U_2)^\circ = U_1^\circ + U_2^\circ$, also $U_1 \cap U_2 = (U_1^\circ + U_2^\circ)^\circ$.

Sind Γ_1, Γ_2 UVRe von V^* , gilt $(\Gamma_1 \cap \Gamma_2)^\circ = \Gamma_1^\circ + \Gamma_2^\circ$, also $\Gamma_1 \cap \Gamma_2 = (\Gamma_1^\circ + \Gamma_2^\circ)^\circ$.

Beweis: s. Vorlesung □

Beispiel: Will man den Schnitt zweier Unterräume U_1 und U_2 durch ein möglichst kleines Gleichungssystem beschreiben und kennt man schon Gleichungssysteme für U_1 und U_2 , etwa $U_1 = \{x \in K^n : A^{(1)}x = 0\}$, $U_2 = \{x \in K^n : A^{(2)}x = 0\}$ mit $A^{(1)} \in K^{m_1 \times n}$, $A^{(2)} \in K^{m_2 \times n}$, empfiehlt das Korollar, eine Basis $(a_1^{(3)}, \dots, a_r^{(3)})$ von $U_1^\circ + U_2^\circ = \operatorname{span}(\{A_{i,\bullet}^{(1)} : i = 1, \dots, m_1\} \cup$

$\{A_{i,\bullet}^{(2)} : i = 1, \dots, m_2\})$ auszuwählen. Dann ist für $A^{(3)} = \begin{bmatrix} a_1^{(3)} \\ \vdots \\ a_r^{(3)} \end{bmatrix} \in K^{r \times n}$

$U_1 \cap U_2 = \{x \in K^n : A^{(3)}x = 0\}$. ♡

Mit diesen Resultaten wird klar: Im endlichdimensionalen kann man jeden r -dimensionalen Untervektorraum nicht nur „von innen“ durch die r aufspannenden Vektoren, sondern genauso „von außen“ (dual) durch $n - r$ einzelne lineare Gleichungen (die Kerne der linearen Funktionale) beschreiben.

Korollar 4.22 Sei $U \subseteq V$ ein r -dimensionaler Untervektorraum eines n -dimensionalen Vektorraums. Dann gibt es Hyperebenen H_1, \dots, H_{n-r} durch Null mit $U = \bigcap_{j=1}^{n-r} H_j$.

Beweis: s. Vorlesung □

Da V^* ein Vektorraum ist, können wir auch V^{**} bilden. Die nächste Beobachtung zeigt, dass wir für endlichdimensionales V den Vektorraum V^{**} mit V identifizieren können, ganz ohne eine Basis kennen zu müssen.

Beobachtung 4.23 Sei V ein endlichdimensionaler K -Vektorraum, dann ist

$$\mu: V \rightarrow V^{**}$$

$$v \mapsto \begin{cases} \mu(v): V^* \rightarrow K \\ \varphi \mapsto (\mu(v))(\varphi) := \varphi(v) \end{cases}$$

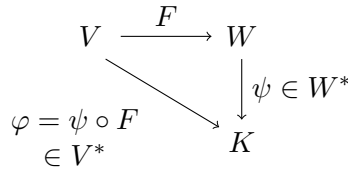
ein Isomorphismus, also $V \simeq V^{**}$.

μ in Worten: μ bildet jedes $v \in V$ auf die Linearform $\mu(v) \in V^{**}$ ab, die jedem $\varphi \in V^*$ ihren Wert in v zuweist (also einfach v in φ verwendet, $v \equiv \mu(v)$).

Beweis: s. Vorlesung □

Die duale Abbildung F^* zu $F: V \rightarrow W$

Wenn eine lineare Abbildung $F: V \rightarrow W$ Elementen aus V Elemente aus W zuordnet um diesen „dort“ über ein lineares Funktional $\psi: W \rightarrow K$ Werte aus K zuzuweisen, ist es günstiger, gleich ein entsprechendes Funktional auf V zu bestimmen und anzuwenden.



Definition 4.24 Sind V und W K -Vektorräume und $F \in \text{Hom}_K(V, W)$, dann heißt $F^*: W^* \rightarrow V^*$ die duale Abbildung (zu F).

$$\psi \mapsto \psi \circ F$$

Beispiel: Die 2×3 Matrix $F = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 3 & 2 \end{bmatrix}$ diene als Abbildung $F: V = \mathbb{R}^3 \rightarrow W = \mathbb{R}^2$. Jedes lineare Funktional $\psi: W = \mathbb{R}^2 \rightarrow K = \mathbb{R}$ kann als 1×2 Zeilenvektor $a = [a_1 \ a_2]$ aufgefasst werden. Für jeden Vektor $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ muss nun in $\psi(F(x))$ zuerst $y = Fx = \begin{bmatrix} 2x_1 - x_3 \\ 3x_2 + 2x_3 \end{bmatrix} \in \mathbb{R}^2$ und dann $\psi(y) = a_1 y_1 + a_2 y_2$ berechnet werden, aber das ist einfach

$$\varphi(x) = a_1(2x_1 - x_3) + a_2(3x_2 + 2x_3) = [2a_1 \ 3a_2 \ a_2 - a_1] x,$$

d.h. $F^*: W^* \rightarrow V^*$, $a \mapsto [2a_1 \ 3a_2 \ a_2 - a_1]$.

Viel einfacher sieht man das direkt mit der Matrixmultiplikation:

$$\psi(F(x)) = a \cdot (F \cdot x) = (a \cdot F) \cdot x = (F^*(a)) \cdot x = \varphi(x).$$

Tatsächlich ist $F^*(a) = [2a_1 \ 3a_2 \ a_2 - a_1] = a \cdot F$. ♡

Beobachtung 4.25 $F^* \in \text{Hom}_K(W^*, V^*)$ (d.h., F^* ist linear).

Beweis: s. Vorlesung □

Da unter F alle $v \in \text{Ker } F$ den Wert $F(v) = 0$ haben und somit $\psi(F(v)) = 0$ für jedes $\psi \in W^*$ gilt, müssen die Funktionale im Bild von F^* in ihrem Kern mindestens den Kern von F enthalten, d.h., diese Funktionale müssen im Annulator von $\text{Ker } F$ liegen.

Beobachtung 4.26 *Sind V und W endlichdimensionale K -Vektorräume und $F \in \text{Hom}_K(V, W)$, dann gilt: $\text{Im } F^* = (\text{Ker } F)^\circ$.*

Beweis: s. Vorlesung □

In Worten: Das Bild der dualen Abbildung ist der Annulator (im \mathbb{R}^n als das „orthogonale“ Komplement vorstellbar) des Kerns der Abbildung.

Beispiel: Wie man im vorigen Beispiel an $a \cdot F$ für $a \in \mathbb{R}^{1 \times 2}$ sieht, ist für eine Matrix F mit $\text{Ker } F = \text{Lös}(F, 0)$ der Annulator $(\text{Ker } F)^\circ$ einfach der von den Zeilen von F aufgespannte Untervektorraum in V^* .

Das sollte auch auf Grund der Betrachtungen zu den homogenen Gleichungssysteme nicht mehr überraschen. ♡

Die Komposition dualer Abbildungen ist ganz einfach:

Beobachtung 4.27 *Für $G \in \text{Hom}_K(U, V)$, $F \in \text{Hom}_K(V, W)$ gilt $(F \circ G)^* = G^* \circ F^*$.*

Beweis: s. Vorlesung □

Kapitel 5

Lineare Abbildungen und Matrizen

5.1 Koordinatensysteme und Matrizenrechnung

Nach B 4.3 ist $\text{Hom}_K(V, W)$ ein Untervektorraum von $\text{Abb}(V, W) = W^V$.

Für endlichdimensionale Vektorräume wird $\text{Hom}_K(V, W)$ über Matrizen sehr konkret vorstellbar: Sei dazu

$\mathcal{A} = (v_1, \dots, v_n)$ eine Basis von V ,
 $\mathcal{B} = (w_1, \dots, w_m)$ eine Basis von W ,

$$V \simeq K^n \text{ über Koordinatenabb. } \quad \mathcal{K}_{\mathcal{A}}: \quad V \quad \rightarrow K^n$$

$$v = \sum_{j=1}^n \alpha_j v_j \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} (= \sum \alpha_j e_j),$$

$$W \simeq K^m \text{ über Koordinatenabb. } \quad \mathcal{K}_{\mathcal{B}}: \quad W \quad \rightarrow K^m$$

$$w = \sum_{i=1}^m \beta_i w_i \mapsto \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} (= \sum \beta_i e_i).$$

Für gegebenes $F \in \text{Hom}_K(V, W)$ soll $A = (a_{ij}) \in K^{m \times n}$ so bestimmt werden, dass folgendes Diagramm kommutiert:

$$\begin{array}{ccc}
 v_j \mapsto & \xrightarrow{\quad} & \sum_{i=1}^m a_{ij} w_i \\
 \downarrow \mathcal{K}_{\mathcal{A}} \simeq & \begin{array}{ccc} V & \xrightarrow{F} & W \\ & \simeq \downarrow \mathcal{K}_{\mathcal{B}} & \\ K^n & \xrightarrow{A} & K^m \end{array} & \downarrow \\
 e_j \mapsto & \xrightarrow{\quad} & \sum_{i=1}^m a_{ij} e_i = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}
 \end{array}$$

Nach S 4.8 reicht es aus, den Weg der Basisvektoren $v_j, j = 1, \dots, n$, zu verfolgen:

Hat v_j das Bild $F(v_j) = \hat{w}_j = \sum_{i=1}^m a_{ij}w_i$ (Koordinaten von \hat{w}_j bzgl. \mathcal{B}), muss $\mathcal{K}_{\mathcal{A}}(v_j) = e_j$ unter A das Bild $A(e_j) = \mathcal{K}_{\mathcal{B}}(\hat{w}_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$ haben.

Ist nun allgemein $v = \sum_{j=1}^n \alpha_j v_j$, dann ist $\mathcal{K}_{\mathcal{A}}(v) = \sum_{j=1}^n \alpha_j e_j$ und

$$\mathcal{K}_{\mathcal{B}}(F(v)) = A(\sum_{j=1}^n \alpha_j e_j) = \sum_{h=1}^m \alpha_j \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in K^m.$$

In Matrixschreibweise,

$$A \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Merke: Spalte $A_{\bullet, j}$ enthält die Koordinaten des Bildes des j -ten Basisvektors bezüglich der Basis des Bildraumes.

Tatsächlich sind für K -Vektorräume V und W der Dimension n und m die linearen Abb. $\text{Hom}_K(V, W)$ isomorph zum Vektorraum der linearen Abb., die durch $K^{m \times n}$ -Matrizen beschrieben werden. Bevor wir das zeigen, sollen kurz die Unterschiede im unendlichdimensionalen Fall beleuchtet werden.

Sind V, W unendlichdimensional mit unendlichen Basen $\mathcal{A} = (v_j)_{j \in \mathcal{J}}$ und $\mathcal{B} = (w_i)_{i \in \mathcal{I}}$, ändert sich Folgendes:

Die Koordinatenabb. lauten $\mathcal{K}_{\mathcal{A}}: V \rightarrow \text{span}_K(e_j)_{j \in \mathcal{J}} \subsetneq K^{\mathcal{J}}$
 bzw. $\mathcal{K}_{\mathcal{B}}: W \rightarrow \text{span}_K(e_i)_{i \in \mathcal{I}} \subsetneq K^{\mathcal{I}}$, die Matrizen bilden sich sonst wie zuvor,

- lassen sich aber nicht mehr so schön anordnen,
- umfassen nicht mehr alle möglichen Matrizen aus $K^{\mathcal{I} \times \mathcal{J}}$, sondern nur die, die in jeder „Spalte“ $j \in \mathcal{J}$ höchstens endlich viele Koeffizienten $\neq 0$ haben.

Die Nutzung von Matrixdarstellungen linearer Abbildung zwischen unendlichdimensionalen Räumen ist daher etwas komplizierter, wir konzentrieren uns vorwiegend auf den endlichdimensionalen Fall.

Als direkte Folgerung von S 4.8 lassen sich Koordinatensysteme so einführen:

Korollar 5.1 *Ist V ein K -Vektorraum mit einer Basis $\mathcal{B} = (v_1, \dots, v_n)$, gibt es genau einen Isomorphismus $\Phi_{\mathcal{B}}: K^n \rightarrow V$ mit $\Phi_{\mathcal{B}}(e_j) = v_j$ für $j = 1, \dots, n$, wobei (e_1, \dots, e_n) die kanonische Basis von K^n bezeichnet.*

Beweis: s. Vorlesung □

$\mathcal{K}_{\mathcal{B}}$ von zuvor ist $\mathcal{K}_{\mathcal{B}} = \Phi_{\mathcal{B}}^{-1}$. Zur Erinnerung: $\Phi_{\mathcal{B}}$ heißt Koordinatensystem.

Korollar 5.2 Zu jeder linearen Abb. $F: K^n \rightarrow K^m$ gibt es genau eine Matrix $A \in K^{m \times n} = M(m \times n; K)$, sodass $F(x) = Ax$ für alle Spaltenvektoren $x \in K^n$.

Beweis: s. Vorlesung □

Korollar 5.3 Seien K -Vektorräume V mit Basis $\mathcal{A} = (v_1, \dots, v_n)$ und W mit Basis $\mathcal{B} = (w_1, \dots, w_m)$ gegeben. Zu jedem $F \in \text{Hom}_K(V, W)$ gibt es genau eine Matrix $A = (a_{ij}) \in K^{m \times n}$ mit

$$(Mat) \quad F(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{für } j = 1, \dots, n,$$

und die Abbildung $M_{\mathcal{B}}^{\mathcal{A}}: \text{Hom}_K(V, W) \rightarrow K^{m \times n}$, $F \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(F)$ ist ein Isomorphismus von K -Vektorräumen. Insbesondere gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(\alpha F + \beta G) = \alpha M_{\mathcal{B}}^{\mathcal{A}}(F) + \beta M_{\mathcal{B}}^{\mathcal{A}}(G).$$

Man bezeichnet $M_{\mathcal{B}}^{\mathcal{A}}(F)$ als die Matrix, die F bezüglich der Basen \mathcal{A} und \mathcal{B} darstellt.

Beweis: s. Vorlesung □

Man verliert nichts, wenn man sich lineare Abbildungen zwischen endlichdimensionalen Vektorräumen als Matrix-Vektor Multiplikation vorstellt!

Für ein konkretes F hängen die Werte der Matrix $M_{\mathcal{B}}^{\mathcal{A}}(F)$ von der Wahl der Basen ab. S 4.9 besagt, dass es immer Basen gibt, für die die Matrix besonders einfache Gestalt hat.

Korollar 5.4 Sei $F \in \text{Hom}_K(V, W)$ mit endlichen $n = \dim V$, $m = \dim W$, $r = \dim \text{Im } F$, dann gibt es Basen \mathcal{A} von V und \mathcal{B} von W , sodass

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \quad \text{mit} \quad I_r = (\delta_{ij}) = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in K^{r \times r}.$$

Beweis: s. Vorlesung □

Wie sich die Matrixdarstellung ändert, wenn man von einer Basis zu einer anderen übergeht, und wie man Basen für die Darstellung aus K 5.4 algorithmisch findet, wird später ein wichtiges Thema sein.

Komposition linearer Abb. und Matrixmultiplikation

Für lineare Abb. $G \in \text{Hom}_K(U, V)$, $F \in \text{Hom}_K(V, W)$ ist nach B 4.4 auch $F \circ G \in \text{Hom}_K(U, W)$ linear. Wie bekommt man die Matrix zu $F \circ G$ aus den Matrizen zu F und G ?

O. E. sei dazu $U = K^r$, $V = K^n$, $W = K^m$ mit den kanonischen Basen (e_j^U) , (e_h^V) , (e_i^W) , $A \in K^{m \times n}$ die Matrix zu F , $B \in K^{n \times r}$ die Matrix zu G , und $C \in K^{m \times r}$ beschreibe $F \circ G: K^r \rightarrow K^m$.

Um die Darstellung von C über (Mat) zu erhalten, verfolgen wir das Bild von $e_j^U \in K^r$ durch B und dann durch A :

$$G(e_j^U) = B \cdot e_j^U = \sum_{h=1}^n b_{hj} e_h^V \text{ und } F(e_h^V) = A \cdot e_h^V = \sum_{i=1}^m a_{ih} e_i^W, \text{ also}$$

$$(F \circ G)(e_j^U) = A \cdot (B \cdot e_j^U) = \sum_{h=1}^n b_{hj} (A e_h^V) = \sum_{h=1}^n \sum_{i=1}^m b_{hj} a_{ih} e_i^W$$

Daher ist

$$(F \circ G)(e_j^U) = \sum_{i=1}^m \underbrace{\sum_{h=1}^n a_{ih} b_{hj}}_{=c_{ij} \quad (\text{Mat})} e_i^W = C \cdot e_j^U$$

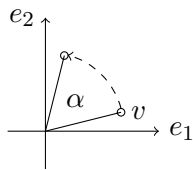
und nach K 5.3 ist C damit eindeutig festgelegt durch

$$C = (c_{ij}) = \left(\sum_{h=1}^n a_{ih} b_{hj} \right)_{ij} = A \cdot B, \quad \text{die bekannte Matrixmultiplikation.}$$

Merke: In endlichdimensionalen Vektorräumen beschreibt die Matrixmultiplikation genau die Hintereinanderausführung der entsprechenden linearen Abbildungen (bei kompatibler Wahl der Basis „in der Mitte“).

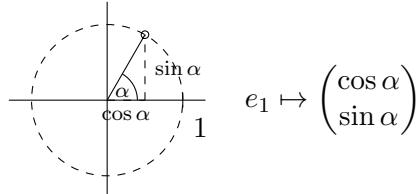
Da für \circ das Assoziativgesetz gilt, $F \circ (G \circ H) = (F \circ G) \circ H$, gilt es auch für die Matrixmultiplikation: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

Beispiel: Drehung um Winkel α um den Ursprung in der Ebene \mathbb{R}^2 :



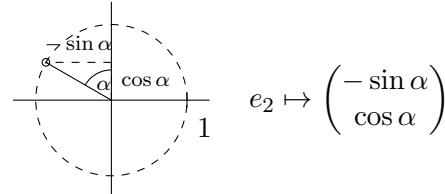
Welche lineare Abb. dreht beliebige $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ um α ?
 Bilde e_1 und e_2 mit einem $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ auf um α gedrehte Versionen $e_1^\alpha = F(e_1)$ und $e_2^\alpha = F(e_2)$ ab, dann ist auch $F(v) = x_1 \cdot F(e_1) + x_2 \cdot F(e_2)$ um α gedreht.

Drehung von e_1 :



$$e_1 \mapsto \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$$

Drehung von e_2 :



$$e_2 \mapsto \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$$

$A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$ beschreibt die lineare Abbildung, die alle Vektoren um α gegen den Uhrzeigersinn „in mathematisch positiver Richtung“ dreht.
 ($= [F(e_1) \quad F(e_2)]$)

Sind A, B, C Drehungen um α, β, γ , besagt das Assoziativgesetz, dass es egal ist, ob man mit $A(BC)$ zuerst um $\beta + \gamma$ und dann um α , oder mit $(AB)C$ zuerst um γ und dann um $\alpha + \beta$ dreht.

Hier gilt aber (geometrisch einsichtig) sogar *ausnahmsweise* Kommutativität:

Sei $B = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}$, dann ist

$$A \cdot B = \begin{bmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -\cos \alpha \sin \beta - \sin \alpha \cos \beta \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta & -\sin \alpha \sin \beta + \cos \alpha \cos \beta \end{bmatrix} = \begin{bmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{bmatrix} = B \cdot A.$$

Für die Drehung zurück wählt man $\beta = -\alpha$, dann ist

$$A \cdot B = \begin{bmatrix} \cos 0 & -\sin 0 \\ \sin 0 & \cos 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{die Identität.}$$

Hier sind A und B inverse Elemente bzgl. der Matrixmultiplikation, $B = A^{-1}$. Die Matrizen $\left\{ \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$ bilden mit der Matrixmultiplikation daher sogar eine kommutative Gruppe. ♡

Für Endomorphismen, also lineare Abb. $F, G: V \rightarrow V$ wie im Beispiel, ist auch $F \circ G: V \rightarrow V$ wieder ein Endomorphismus, und $\text{End}_K(V)$ ist ein Ring mit Eins (B 4.5). Für Matrizen:

Beobachtung 5.5 *Ist V ein K -Vektorraum mit $\dim V = n < \infty$, so ist der Ring $(\text{End}_K(V), +, \circ)$ isomorph zu $(K^{n \times n}, +, \cdot)$.*

Beweis: s. Vorlesung □

Das Nullelement	in $\text{End}_K(V)$	ist die Nullabbildung	$\mathbb{0}: v \mapsto 0,$
	in $K^{n \times n}$	die Nullmatrix	0 oder $0^{n \times n}$.
Das Einselement	in $\text{End}_K(V)$	ist die Identität	$\text{id}_V: v \mapsto v,$
	in $K^{n \times n}$	die Identität	$I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$
		oder Einheitsmatrix	(auch E_n oder I).

Bereits bei Ringen wurde gezeigt, das $(K^{n \times n}, +, \cdot)$ *nicht* nullteilerfrei ist. Wie das geometrisch zu interpretieren ist, wird später klar werden.

Invertierbare/Reguläre Matrizen

Eine Matrix/lineare Abbildung muss ein Isomorphismus sein (linear und bijektiv), damit die Umkehrabbildung existiert, dies ist dann auch ein Isomorphismus, s. B 4.2(c).

Dafür kommen somit nur Matrizen $A \in K^{n \times n}$ in Frage. Für quadratische Matrizen gilt $K^{n \times n} \simeq \text{End}_K(V)$ (für $\dim V = \dim W$ wäre natürlich auch $K^{n \times n} \simeq \text{Hom}_K(V, W)$ korrekt, die Vorstellung der Abbildung in denselben Vektorraum ist hier aber wegen der Ringstruktur günstiger).

Ist $F \in \text{End}(K^n)$ Automorphismus, dann ist $F^{-1} \in \text{End}(K^n)$ und $F \circ F^{-1} = F^{-1} \circ F = \text{id}_{K^n}$.

Definition 5.6 $A \in K^{n \times n}$ heißt invertierbar/regular, wenn es $A^{-1} \in K^{n \times n}$ gibt mit $A^{-1} \cdot A = I_n (= A \cdot A^{-1})$ und Eindeutigkeit aus Isomorphie zu den Abbildungen $\text{End}(K^n)$.

Ist A nicht invertierbar, nennt man A auch singulär.

Beobachtung 5.7 Die Menge $\text{GL}(n; K) = \{A \in K^{n \times n} : A \text{ invertierbar}\}$ mit der Multiplikation als Verknüpfung ist eine (i.A. nicht kommutative) Gruppe mit neutralem Element I_n . Sie wird allgemeine lineare Gruppe (generalised linear) genannt und ist isomorph zur Gruppe der Vektorraumautomorphismen $(\text{Aut}(K^n), \circ)$.

Beweis: s. Vorlesung □

Aus dem Beweis erkennt man $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$,
und allgemein gilt für Gruppen $(A^{-1})^{-1} = A$.

Die Transposition

Die Transposition bildet aus einer Matrix $A \in K^{m \times n}$ (bzw. einem Spaltenvektor $v \in K^n$) eine neue Matrix $A^\top \in K^{n \times m}$ (bzw. einen Zeilenvektor $v^\top \in K^{1 \times n}$), indem sie die Spalten von A zu Zeilen von A^\top macht,

$$\begin{aligned} \cdot^\top : K^{m \times n} &\rightarrow K^{n \times m} \\ A = (a_{ij}) &\mapsto A^\top := (a'_{ij}) \text{ mit } a'_{ij} = a_{ji}, i = 1, \dots, n, j = 1, \dots, m. \end{aligned}$$

Konkret ist etwa $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^\top = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Die Transposition ist eine lineare Abbildung, $\cdot^\top \in \text{Hom}(K^{m \times n}, K^{n \times m})$, A^\top heißt die Transponierte von A .

Es gilt $(A \cdot B)^\top = B^\top \cdot A^\top$,

denn: für $C = AB$ ist $(C)_{ij} = A_{i,\bullet} \cdot B_{\bullet,j} \overset{\top}{\rightsquigarrow} (C^\top)_{ji} = (B^\top)_{j,\bullet} \cdot (A^\top)_{\bullet,i}$.

Ist $A \in K^{n \times n}$ invertierbar ($A \in \text{GL}(n; K)$), dann ist auch $A^\top \in \text{GL}(n; K)$,

denn: $I_n = I_n^\top = (A \cdot A^{-1})^\top = (A^{-1})^\top \cdot A^\top$ und wegen der Eindeutigkeit folgt $(A^\top)^{-1} = (A^{-1})^\top =: A^{-\top}$.

Die Diskussion der Dualräume zeigte $(K^n)^* \simeq K^{1 \times n} \simeq K^n$. Jedes lineare Funktional $\varphi \in \text{Hom}_K(K^n, K) = (K^n)^*$ lässt sich bezüglich der dualen Basis durch einen Koordinatenvektor $c \in K^n$ über $\varphi(x) = c^\top x$ darstellen. Ebenso können für Matrizen A die Zeilen $A_{i,\bullet}$ als Elemente von $(K^n)^*$ aufgefasst werden, für die wieder etwa der span betrachtet werden kann. Transponieren „wechselt“ gewissermaßen zwischen Raum und Dualraum.

Für Abbildungen beschreibt das Transponieren gerade den Übergang zur dualen Abbildung. Zur Verdeutlichung sei $F \in \text{Hom}_K(V, W)$ mit Matrixdarstellung $A \in K^{m \times n}$ (für gewählte Basen) und $\psi \in \text{Hom}_K(W, K) = W^*$ ein

lineares Funktional mit Koordinatendarstellung $d \in K^m$ zur dualen Basis von W . Die duale Abbildung F^* bildet das lineare Funktional ψ auf das lineare Funktional $\varphi = F^*(\psi) \in \text{Hom}_K(V, K) = V^*$ ab, sodass $\varphi(v) = \psi(F(v))$ für $v \in V$ gilt. Ist x der Koordinatenvektor zu v und $y = Ax$ der Koordinatenvektor zu $w = F(v)$, dann ist $c = A^\top d$ der Koordinatenvektor von φ zur dualen Basis von V , denn

$$\psi(F(v)) = d^\top(Ax) = (A^\top d)^\top x = c^\top x = \varphi(v).$$

A^\top ist damit gerade Matrixdarstellung von F^* bezüglich der Dualbasen.

Beobachtung 5.8 Seien K -Vektorräume V mit Basis $\mathcal{A} = (v_1, \dots, v_n)$ und W mit Basis $\mathcal{B} = (w_1, \dots, w_m)$ sowie $F \in \text{Hom}_K(V, W)$ gegeben. Dann gilt mit den dualen Basen \mathcal{A}^* und \mathcal{B}^* aus B 4.16

$$M_{\mathcal{B}}^{\mathcal{A}}(F)^\top = M_{\mathcal{A}^*}^{\mathcal{B}^*}(F^*).$$

Inbesondere ist $\cdot^* : \text{Hom}_K(V, W) \rightarrow \text{Hom}_K(W^*, V^*)$, $F \mapsto F^*$ ein Vektorraumisomorphismus.

Beweis: s. Vorlesung □

Die Rechenregel $(A \cdot B)^\top = B^\top \cdot A^\top$ entspricht $(F \circ G)^* = G^* \circ F^*$ von B 4.27.

Spaltenrang, Zeilenrang und Rang einer Matrix

Jede Matrix $A \in K^{m \times n}$ definiert eine lineare Abb. $A: K^n \rightarrow K^m$, $x \mapsto Ax$; man nennt

$$\left. \begin{array}{l} \dim \text{Im } A = \dim \text{span}(A_{\bullet, j})_{j=1, \dots, n} \text{ den } \underline{\text{Spaltenrang}} \\ \dim \text{Im } A^\top = \dim \text{span}(A_{i, \bullet})_{i=1, \dots, m} \text{ den } \underline{\text{Zeilenrang}} \end{array} \right\} \text{ der Matrix } A.$$

Satz 5.9 Für $A \in K^{m \times n}$ ist $\text{Spaltenrang} = \text{Zeilenrang} = \text{der } \underline{\text{Rang}} \text{ der Matrix } A$.

Beweis: s. Vorlesung □

Man erkennt das auch sofort aus K 5.4, aber das wird erst mit Beherrschung der Basistransformationen klar.

Rechenregeln 5.10 (für Matrizen) Sind $A, A' \in K^{m \times n}$, $B, B' \in K^{n \times r}$, $C \in K^{r \times s}$ und $\alpha \in K$, so gilt:

- (1) *Distributivgesetz:* $A(B + B') = AB + AB'$ und $(A + A')B = AB + A'B$,
- (2) $A(\alpha B) = (\alpha A)B = \alpha(AB)$,

- (3) Assoziativgesetz: $(AB)C = A(BC)$,
- (4) $(AB)^\top = B^\top A^\top$,
- (5) $I_m \cdot A = A \cdot I_n = A$.

Sind $A, B \in GL(n; K)$, gilt zusätzlich

- (6) $(A^{-1})^{-1} = A$,
- (7) $(AB)^{-1} = B^{-1}A^{-1}$,
- (8) $(A^\top)^{-1} = (A^{-1})^\top (= A^{-\top})$.

Beweis: s. Vorlesung □

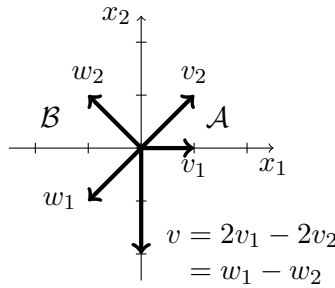
5.2 Koordinatentransformationen

Ist V ein endlichdimensionaler K -Vektorraum mit $n = \dim V$ und ist $\mathcal{A} = (v_1, \dots, v_n)$ eine Basis von V , dann beschreibt nach K 5.1 das Koordinatensystem $\Phi_{\mathcal{A}}: K^n \rightarrow V$ mit $\Phi_{\mathcal{A}}(e_j) = v_j, j = 1, \dots, n$ genau einen Isomorphismus. Dessen Umkehrabbildung

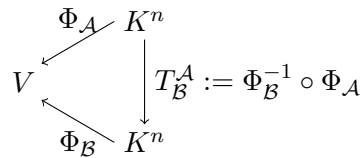
$$\Phi_{\mathcal{A}}^{-1}: V \rightarrow K^n, v = x_1v_1 + \dots + x_nv_n \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

weist jedem $v \in V$ seine Koordinaten bezüglich \mathcal{A} zu.

Eine andere Basis $\mathcal{B} = (w_1, \dots, w_n)$ von V ergibt ein anderes Koordinatensystem $\Phi_{\mathcal{B}}$ und andere Koordinaten für das gleiche v .



Die Übergänge beschreibt das Diagramm



(alle Abbildungen sind Isomorphismen).

$T_{\mathcal{B}}^{\mathcal{A}}$ heißt Transformationsmatrix des Basiswechsels (von Basis \mathcal{A} zu Basis \mathcal{B}).

Der Name Matrix ist berechtigt, denn $T_{\mathcal{B}}^{\mathcal{A}} \in GL(n; K)$, weil $T_{\mathcal{B}}^{\mathcal{A}}$ als Automorphismus auf K^n für die kanonische Basis e_1, \dots, e_n direkt als Matrix aufgefasst werden kann.

Für $v = x_1v_1 + \dots + x_nv_n = y_1w_1 + \dots + y_nw_n$ erfüllt sie $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_{\mathcal{B}}^{\mathcal{A}} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Beispiel: Für $V = \mathbb{R}^2$, $\mathcal{A} = (v_1, v_2)$, $\mathcal{B} = (w_1, w_2)$ (wie im Bild oben) ist mit $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, w_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ und Matrixdarstellung zu

$\Phi_{\mathcal{A}}: A = [v_1 \ v_2] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad v = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} \in \mathbb{R}^2,$
 $\Phi_{\mathcal{B}}: B = [w_1 \ w_2] = \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}, \quad v = \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} \in \mathbb{R}^2.$
 $T_{\mathcal{B}}^{\mathcal{A}} = B^{-1} \cdot A$, wobei B^{-1} nun e_1, e_2 bzgl. der Basis w_1, w_2 beschreibt:

$$\rightsquigarrow \quad B^{-1} = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \quad (\text{stimmt, denn } B \cdot B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}),$$

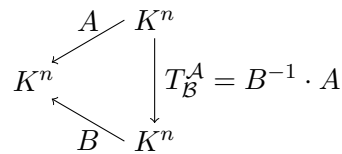
$$T_{\mathcal{B}}^{\mathcal{A}} = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & -2 \\ -1 & 0 \end{bmatrix}.$$

$T_{\mathcal{B}}^{\mathcal{A}}$ bildet die Koordinaten von v bzgl. \mathcal{A} auf die von v bzgl. \mathcal{B} ab:

$$T_{\mathcal{B}}^{\mathcal{A}} \cdot \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \frac{1}{2} \begin{bmatrix} -1 & -2 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{bmatrix} -1 & -2 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

♡

Ist speziell $V = K$ und beschreiben A und B die Koordinatensysteme $\Phi_{\mathcal{A}}$ und $\Phi_{\mathcal{B}}$ (die Vektoren sind dann jeweils die Spalten von A und B) erhält man (wie im Beispiel)



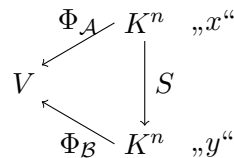
Sind nun für allgemeines V die Vektoren von \mathcal{A} bereits als Linearkombinationen von \mathcal{B} gegeben, also

$$v_j = s_{1j}w_1 + \dots + s_{nj}w_n, \quad j = 1, \dots, n,$$

dann ist mit $S = \begin{bmatrix} s_{11} & \dots & s_{1n} \\ \vdots & & \vdots \\ s_{n1} & \dots & s_{nn} \end{bmatrix}$ ($S \cdot x$ bildet Koordinaten für \mathcal{A} auf Koordinaten für \mathcal{B} ab)

$$\Phi_{\mathcal{A}} = \Phi_{\mathcal{B}} \circ S,$$

$$\text{damit } T_{\mathcal{B}}^{\mathcal{A}} = S.$$



Beispiel: In Fortsetzung des vorigen Beispiels ist

$$\left. \begin{array}{l} v_1 = -\frac{1}{2} \cdot w_1 - \frac{1}{2} \cdot w_2 \rightarrow S_{\bullet,1} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \\ v_2 = -1 \cdot w_1 + 0 \cdot w_2 \rightarrow S_{\bullet,2} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{array} \right\} S = \begin{bmatrix} -\frac{1}{2} & -1 \\ -\frac{1}{2} & 0 \end{bmatrix}$$

$$\Rightarrow T_{\mathcal{B}}^{\mathcal{A}} = S = \begin{bmatrix} -\frac{1}{2} & -1 \\ -\frac{1}{2} & 0 \end{bmatrix}.$$

Sind umgekehrt die Koordinaten von w_1 bzw. w_2 bezüglich v_1 und v_2 bekannt, geben sie die Spalten von S^{-1} an,

$$\left. \begin{array}{l} w_1 = 0 \cdot v_1 - 1 \cdot v_2 \rightarrow S_{\bullet,1}^{-1} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ w_2 = -2 \cdot v_1 + 1 \cdot v_2 \rightarrow S_{\bullet,2}^{-1} = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \end{array} \right\} S^{-1} = \begin{bmatrix} 0 & -2 \\ -1 & 1 \end{bmatrix}$$

$$\Rightarrow T_{\mathcal{A}}^{\mathcal{B}} = (T_{\mathcal{B}}^{\mathcal{A}})^{-1} = S^{-1} = \begin{bmatrix} 0 & -2 \\ -1 & 1 \end{bmatrix} \text{ (denn } S^{-1} \cdot S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{)}.$$

♡

Für Matrizen $A \in \text{GL}(n; K)$ möchte man oft die Inverse A^{-1} explizit bestimmen oder die inverse Abbildung numerisch ausführen. Wie das geht, wird später ein Thema sein.

Wir wollen nun die Matrixdarstellung $M_{\mathcal{B}}^{\mathcal{A}}(F)$ einer linearen Abbildung $F: V \rightarrow W$ sowie die Matrixmultiplikation als Komposition linearer Abbildungen noch einmal in Bezug auf die Abhängigkeit von der Wahl der Basis beleuchten.

Beobachtung 5.11 Für $F \in \text{Hom}_K(V, W)$ mit endlichen Basen \mathcal{A} und \mathcal{B} von V und W kommutiert das folgende Diagramm linearer Abbildungen,

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi_{\mathcal{A}}} & V \\ M_{\mathcal{B}}^{\mathcal{A}}(F) \downarrow & & \downarrow F \\ K^m & \xrightarrow{\Phi_{\mathcal{B}}} & W \end{array} \quad \text{bzw.} \quad \begin{array}{l} \text{und es gilt} \\ \Phi_{\mathcal{B}} \circ M_{\mathcal{B}}^{\mathcal{A}}(F) = F \circ \Phi_{\mathcal{A}} \\ M_{\mathcal{B}}^{\mathcal{A}}(F) = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}. \end{array}$$

Beachte: Für $T_{\mathcal{B}}^{\mathcal{A}}$ ist $V = W$ und $F = \text{id}_V$, also $M_{\mathcal{B}}^{\mathcal{A}}(\text{id}_V) = T_{\mathcal{B}}^{\mathcal{A}}$.

Beweis: s. Vorlesung

□

Zur Matrixmultiplikation:

Satz 5.12 Für $G \in \text{Hom}_K(U, V)$, $F \in \text{Hom}_K(V, W)$ mit endlichen Basen $\mathcal{A}, \mathcal{B}, \mathcal{C}$ für U, V, W gilt: $M_{\mathcal{C}}^{\mathcal{A}}(F \circ G) = M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}(G)$.

Beachte: Die Matrizen müssen die Abbildungen für die gleiche Basis \mathcal{B} von V darstellen!

Beweis: s. Vorlesung

□

Korollar 5.13 Für $F, G \in \text{End}_K(V)$ mit $\dim V = n \in \mathbb{N}$ und Basis \mathcal{B} ist $M_{\mathcal{B}}(F \circ G) = M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}(G)$. Daher ist $M_{\mathcal{B}}: \text{End}_K(V) \rightarrow M(n \times n; K)$ ein Ring-mit-Eins-Isomorphismus.

Nun lässt sich gut beschreiben, wie sich die Matrixdarstellung je nach Wahl der Basis ändert.

Satz 5.14 (Transformationsformel) Für $F \in \text{Hom}_K(V, W)$ mit endlichen Basen $\mathcal{A}, \mathcal{A}'$ und $\mathcal{B}, \mathcal{B}'$ für V, W kommutiert das folgende Diagramm:

$$\begin{array}{ccc}
 K^n & \xrightarrow{A := M_{\mathcal{B}}^{\mathcal{A}}(F)} & K^m \\
 \downarrow T := T_{\mathcal{A}'}^{\mathcal{A}} & \begin{array}{c} \searrow \Phi_{\mathcal{A}} \\ V \xrightarrow{F} W \\ \nearrow \Phi_{\mathcal{A}'} \end{array} & \begin{array}{c} \Phi_{\mathcal{B}} \\ \searrow \\ T_{\mathcal{B}'}^{\mathcal{B}} =: S \\ \nearrow \Phi_{\mathcal{B}'} \end{array} & \downarrow \\
 K^n & \xrightarrow{B := M_{\mathcal{B}'}^{\mathcal{A}'}(F)} & K^m
 \end{array}$$

Insbesondere gilt

$$M_{\mathcal{B}'}^{\mathcal{A}'}(F) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}$$

bzw.

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

Beweis: s. Vorlesung □

Korollar 5.15 Für $F \in \text{End}_K(V)$ mit endlichen Basen $\mathcal{A} = (v_j)_{j=1}^n, \mathcal{B} = (w_i)_{i=1}^n$ von V gilt

$$\begin{array}{cccc}
 M_{\mathcal{B}}(F) & = & T_{\mathcal{B}}^{\mathcal{A}} \cdot M_{\mathcal{A}}(F) \cdot T_{\mathcal{A}}^{\mathcal{B}} \\
 \parallel & & \parallel & \parallel \\
 \text{bzw. } B & = & S \cdot A \cdot S^{-1} & \text{ wobei } v_j = \sum_{i=1}^n s_{ij} w_i, j = 1, \dots, n.
 \end{array}$$

Eine Matrix S ist genau dann eine Transformationsmatrix auf V , wenn sie einen Automorphismus beschreibt $\Leftrightarrow S \in \text{GL}(\dim V; K)$.

Interpretiert man $S = T_{\mathcal{B}}^{\mathcal{A}}$ als Transformation von Koordinaten bzgl. \mathcal{A} auf Koordinaten bzgl. \mathcal{B} , geben die Spalten von S die Koordinaten der alten Basisvektoren von \mathcal{A} bzgl. der neuen Basisvektoren von \mathcal{B} an. Die Spalten von S^{-1} enthalten hingegen die Koordinaten der neuen Basisvektoren von \mathcal{B} bzgl. der alten Basisvektoren von \mathcal{A} .

Definition 5.16 Zwei Matrizen $A, B \in K^{m \times n}$ heißen äquivalent, wenn es ein $S \in \text{GL}(m; K)$ und ein $T \in \text{GL}(n; K)$ gibt mit $B = SAT^{-1}$.

Zwei Matrizen $A, B \in K^{n \times n}$ heißen ähnlich, wenn es ein $S \in \text{GL}(n; K)$ gibt mit $B = SAS^{-1}$.

Beachte: Zwei verschiedene Matrizen sind genau dann äquivalent, wenn sie bezüglich *verschiedener Paare* von Basen die gleiche lineare Abbildung beschreiben. Bei der Matrixdarstellung von Endomorphismen darf man jeweils nur *eine* Basis wählen, und zwei verschiedene *quadratische* Matrizen sind genau dann ähnlich, wenn sie bezüglich *verschiedener einzelner* Basen den gleichen Endomorphismus beschreiben.

Beide Begriffe beschreiben Äquivalenzreaktionen:

	äquivalent	ähnlich
Reflexivität	$A = I_m \cdot A \cdot I_n$	$A = I_n \cdot A \cdot I_n$
Symmetrie	$A = S^{-1} \cdot B \cdot T$	$A = S^{-1} \cdot B \cdot S$
Transitivität	$C = S_C \underbrace{(S_B A T_B^{-1})}_{=B} T_C^{-1}$ $= S A T^{-1}$	$C = S_C \underbrace{(S_B A S_B^{-1})}_{=B} S_C^{-1}$ $= S A S^{-1}$

Es stellt sich die Frage, ob man zu jeder Äquivalenzklasse von Matrizen besonders einfache Matrizen, sogenannte **Normalformen**, als Repräsentanten finden kann.

Für äquivalente Matrizen ist das über K 5.4 sehr leicht möglich.

Lemma 5.17 *Zwei Matrizen aus $K^{m \times n}$ sind genau dann äquivalent, wenn sie den gleichen Rang haben. Insbesondere ist jede Matrix vom Rang r äquivalent zu*

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Beweis: s. Vorlesung □

Wie man die Basen findet, die zu dieser Normalform führen, wird im nächsten Abschnitt erklärt.

Für *Ähnlichkeit* ist die Sachlage deutlich komplizierter, weil die Abbildungen jeweils bezüglich *einer* Basis in möglichst einfacher Form beschrieben werden sollen. Diese Frage wird daher erst in einem späteren Kapitel behandelt.

Zum Ende dieses Abschnitts kommen wir nocheinmal auf Zeilenrang und Spaltenrang zurück.

Lemma 5.18 *Sei $A \in K^{m \times n}$, $S \in GL(m; K)$, $T \in GL(n; K)$. Es gilt*

- (a) *Spaltenrang(SAT) = Spaltenrang(A),*
- (b) *Zeilenrang(SAT) = Zeilenrang(A).*

Beweis: s. Vorlesung □

Dies liefert einen weiteren Beweis für $\text{Zeilenrang}(A) = \text{Spaltenrang}(A)$ für $A \in K^{m \times n}$, **denn:** nach L 5.17 ist A äquivalent zu $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ für ein $r \in \mathbb{N}_0$, diese Matrix hat $\text{Zeilenrang} = \text{Spaltenrang}$, also folgt der Satz aus L 5.18. □

Merke: Multiplikationen von rechts oder links mit regulären Matrizen verändern niemals den Rang, **denn:** in L 5.18 entspricht dies jeweils entweder dem Fall $S = I_m$ oder $T = I_n$. □

$$P_{ij}^{-1} = P_{ij}.$$

Dies entspricht dem Wechsel der Plätze von w_i und w_j . Im Gaußalgorithmus war dies Operation (O2).

Bemerkung: Eine quadratische Matrix P , die in jeder Zeile und Spalte genau eine 1, sonst aber nur den Wert 0 hat, nennt man Permutationsmatrix. \diamond

Ü: Zeige, dass die Menge \mathcal{P}_n der $n \times n$ Permutationsmatrizen mit der Matrixmultiplikation eine Gruppe bildet und dass diese isomorph zur Gruppe der Permutationen ist. \times

Die Matrizen

$$S \in \{S_i(\alpha), Q_{ij}(\alpha), P_{ij} : 0 \neq \alpha \in K, i, j = 1, \dots, m, i \neq j\}$$

werden Elementarmatrizen genannt.

Im Eliminationsverfahren Alg 2.24 können die Operationen auf A jeweils durch Matrixmultiplikationen *von links* mit Elementarmatrizen S_h ersetzt werden. Nach L 5.18 ändern diese nichts am Rang der Matrix.

Satz 5.19 *Jede Matrix $A \in K^{m \times n}$ lässt sich durch Multiplikation mit einer endlichen Folge von Elementarmatrizen $S_i \in \text{GL}(m; K)$, $i = 1, \dots, h$ in Zeilenstufenform bringen, sodass*

$$S_h \cdot \dots \cdot S_1 \cdot A = \bar{A} = \left[\begin{array}{cccc} 0 & \dots & 0 & \underline{*} \dots \\ 0 & & \dots & 0 \underline{*} \dots \\ & & & \ddots \\ & & & 0 \underline{*} \dots \\ 0 & & & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \end{array} \right] \left. \begin{array}{l} \left. \right\} r \\ \left. \right\} m - r \end{array} \right.$$

Es gilt $r = \text{Rang } A$.

Beweis: s. Vorlesung \square

Ü: Inwiefern unterscheidet sich dieser Beweis vom Vorgehen in Alg 2.24? \times

2. Operationen auf den Spalten von $A = M_{\mathcal{B}}^{\mathcal{A}}(F) \in K^{m \times n}$ mit $F: V \rightarrow W$

Ist $\mathcal{A} = (v_1, \dots, v_n)$ die Basis von V , gibt Spalte j die Koordinaten des Bildes von v_j bzgl. \mathcal{B} an. Die Spaltenoperationen sind Multiplikationen von rechts mit $n \times n$ Elementarmatrizen und lassen sich jeweils als Basiswechsel von \mathcal{A} zu einer leicht geänderten Basis \mathcal{A}' interpretieren (**Ü**).

(\mathcal{S}_s) Multiplikation/Skalierung der Spalte i mit $\alpha \in K \setminus \{0\}$: $A \cdot S_i(\alpha)$

(\mathcal{S}_a) Addition des $\alpha \in K$ -fachen der Spalte i zur Spalte j : $A \cdot Q_{ij}(\alpha)$ ($i \neq j$)
(Beachte: Gegenüber (\mathcal{Z}_a) ist die Reihenfolge von i und j vertauscht.)

(S_p) Vertauschen von Spalte i und Spalte j : $A \cdot P_{ij}$ ($i \neq j$)

Multipliziert man die Zeilenstufenmatrix \bar{A} aus S 5.19 von rechts geeignet mit $T_k^{-1} \in \{S_i(\alpha), Q_{ij}(\alpha), P_{ij} \in K^{n \times n} : 0 \neq \alpha \in K, i, j = 1, \dots, n, i \neq j\}$, erhält man (\ddot{U})

$$\underbrace{S_h \cdot \dots \cdot S_1}_{S \in \text{GL}(m; K)} \cdot A \cdot \underbrace{T_1^{-1} \cdot \dots \cdot T_k^{-1}}_{T^{-1} \in \text{GL}(n; K)} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = SAT^{-1}.$$

Gleichzeitig bekommt man eine faktorierte (zerlegte) Darstellung von A

$$A = S^{-1} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} T$$

mit $S^{-1} = S_1^{-1} \cdot \dots \cdot S_h^{-1} \in \text{GL}(m; K)$, $T = T_k \cdot \dots \cdot T_1 \in \text{GL}(n; K)$.

Jede invertierbare Matrix ist ein Automorphismus $A \in \text{GL}(n; K)$ und hat vollen Rang $\text{Rang } A = n$. Sie ist daher äquivalent zu I_n und kann wie oben als Produkt von Elementarmatrizen dargestellt werden. Dies kann sogar nur durch Multiplikation von links (also nur durch Zeilenumformungen) geschehen. Dies wird später von besonderer Bedeutung sein und daher im Detail gezeigt.

Tatsächlich werden nicht alle Elementarmatrizen benötigt, bereits $S_i(\alpha)$, $\alpha \in K \setminus \{0\}$, und $Q_{ij} := Q_{ij}(1)$ genügen vollständig!

Beobachtung 5.20 Die Elementarmatrizen $Q_{ij}(\alpha)$ und P_{ij} sind Produkte von Elementarmatrizen des Typs $S_i(\alpha)$ und Q_{ij} : Für $\alpha \in K \setminus \{0\}$ ist

$$Q_{ij}(\alpha) = S_j\left(\frac{1}{\alpha}\right) \cdot Q_{ij} \cdot S_j(\alpha), \quad P_{ij} = Q_{ji} \cdot Q_{ij}(-1) \cdot Q_{ji} \cdot S_j(-1).$$

Beweis: s. Vorlesung □

Lemma 5.21 Die Elementarmatrizen sind invertierbar und ihre Inversen sind wieder Elementarmatrizen.

Beweis: s. Vorlesung □

Satz 5.22 Jede invertierbare Matrix $A \in K^{n \times n}$ ist (endliches) Produkt von Elementarmatrizen.

Man sagt: Die Gruppe $\text{GL}(n; K)$ wird von den Elementarmatrizen erzeugt.

Beweis: s. Vorlesung □

Algorithmus 5.23 (Invertierbarkeitstest und Berechnung der Inversen)

Input: $A \in K^{n \times n}$

Output: A^{-1} oder Nachweis der Singularität von A

1. Erzeuge Zeilenstufenform $S_h \cdot \dots \cdot S_1 \cdot A$ nach S 5.19 und bestimme $r = \text{Rang } A$. Ist $r < n$ hat A keine Inverse, STOP.

2. Lösche das Dreieck über der Diagonale und normiere diese wie im Beweis zu S 5.22 durch Multiplikation von links mit Elementarmatrizen:

$$S_k \cdot \dots \cdot S_1 \cdot A = I_n$$

3. $A^{-1} = S_k \cdot \dots \cdot S_1$.

Bemerkungen:

- Für Computerberechnungen ist das Vefahren in dieser Form fast immer schlecht! Es ist ein zentrales Thema der *Numerik*, wie man mit den endlichen Fließkommazahlen die Inverse möglichst *stabil* bestimmt.
- Wenn man wirklich einmal von Hand die Inverse berechnen muss, ist es günstig die S_i direkt auf die Matrix $[A \ I_n]$ anzuwenden, dann berechnet man

$$[S_k \cdot \dots \cdot S_1 \cdot A \ S_k \cdot \dots \cdot S_1 \cdot I_n] = [I_n \ A^{-1}].$$

- Man kann statt „nur Zeilenoperationen“ auch „nur Spaltenoperationen“ verwenden. Mischen geht i.A. nicht.

◇

Matrizen in Dreiecksgestalt sind allgemein ein nützliches Hilfsmittel.

$R \in K^{n \times n}$ heißt (rechte) obere Dreiecksmatrix falls $r_{ij} = 0$ für $i > j$,

$L \in K^{n \times n}$ heißt (linke) untere Dreiecksmatrix falls $l_{ij} = 0$ für $i < j$.

Die folgenden Überlegungen sind nur für obere Dreiecksmatrizen formuliert, über Transposition kann man jedoch alle analogen Ergebnisse für untere Dreiecksmatrizen sofort herleiten.

Beobachtung 5.24 Sind $R, R' \in K^{n \times n}$ obere Dreiecksmatrizen, gilt:

- $R \cdot R'$ ist obere Dreiecksmatrix.
- R ist genau dann invertierbar ($\in \text{GL}$), wenn $r_{ii} \neq 0$ für alle $i = 1, \dots, n$.
- Ist $R \in \text{GL}$, dann ist auch R^{-1} obere Dreiecksmatrix.

Beweis: s. Vorlesung

□

Die oberen Dreiecksmatrizen bilden also einen Unterring von $(K^{n \times n}, +, \cdot)$, die invertierbaren oberen Dreiecksmatrizen eine Untergruppe von $\text{GL}(n; K)$.

5.4 Lineare Gleichungssysteme

Mit Hilfe der entwickelten Theorie lassen sich nun lineare Gleichungssysteme sehr elegant in voller Allgemeinheit diskutieren.

Die Koeffizientenmatrix $A \in K^{m \times n}$ und die rechte Seite $b \in K^m$ beschreiben das

$$\underline{\text{lineare Gleichungssystem}} \quad A \cdot x = b \quad (\text{LGS})$$

Das zu Gleichung (LGS) gehörende

$$\text{homogene (Gleichungs-)System} \quad A \cdot x = 0 \quad (\text{HGS})$$

hat mindestens die Lösung 0. Ist $b \neq 0$, nennt man Gleichung (LGS) inhomogen.

$\text{Lös}(A, b) := \{x \in K^n : Ax = b\}$ ist der Lösungsraum von Gleichung (LGS).

Sei $F: K^n \rightarrow K^m$, $x \mapsto A \cdot x$ die durch A beschriebene lineare Abbildung.

Gleichung (LGS) hat genau dann (mindestens) eine Lösung, wenn

$$b \in \text{Im } F = \text{span}(A_{\bullet,j})_{j=1}^n.$$

Die Faser von F über b ist der Lösungsraum,

$$F^{-1}(b) = \text{Lös}(A, b) = \begin{cases} \emptyset & \text{falls } b \notin \text{span}(A_{\bullet,j})_{j=1}^n, \\ v + \text{Ker } F & \text{falls } v \in F^{-1}(b) \neq \emptyset. \end{cases}$$

Insbesondere ist $F^{-1}(0) = \text{Lös}(A, 0) = \text{Ker } F$ ein Untervektorraum und die Lösungsmenge ein affiner Unterraum.

Korollar 5.25 Sei $Ax = b$ ein Gleichungssystem mit m Gleichungen, n Unbekannten und $r = \text{Rang } A$. Dann gilt für die Lösungsräume

- (a) $\text{Lös}(A, 0) \subseteq K^n$ ist Untervektorraum der Dimension $n - r$.
- (b) $\text{Lös}(A, b)$ ist entweder leer oder ein affiner Unterraum der Dimension $n - r$. Ist $v \in \text{Lös}(A, b)$ beliebig, so ist $\text{Lös}(A, b) = v + \text{Lös}(A, 0)$.

Beweis: s. Vorlesung □

Merke: Die allgemeine Lösung eines inhomogenen Gleichungssystems erhält man durch Addition einer speziellen Lösung des inhomogenen zur allgemeinen Lösung des homogenen Systems.

Wie erkennt man algorithmisch, ob es eine Lösung gibt?

Man vergleicht den Rang von A mit dem Rang der erweiterten Matrix $[A \ b]$.

Für $b \begin{cases} \in \\ \notin \end{cases} \text{span}(A_{\bullet,j})_{j=1}^n$ ist $\text{Rang } A \begin{cases} = \\ < \end{cases} \dim \text{span}(b, A_{\bullet,j})_{j=1}^n = \text{Rang } [A \ b]$.

Das zeigt

Satz 5.26 Der Lösungsraum von $Ax = b$ ist genau dann nicht leer, wenn

$$\text{Rang } A = \text{Rang } [A \ b].$$

Eine Basis (w_1, \dots, w_{n-r}) des homogenen Lösungsraums nennt man auch Fundamentalsystem von Lösungen des homogenen Systems.

Spezialfälle

- Enthält $\text{Lös}(A, b)$ nur genau ein Element, sagt man, das System ist eindeutig lösbar.

Beobachtung 5.27 Für $A \in K^{m \times n}$, $b \in K^m$ sind äquivalent

- $Ax = b$ ist eindeutig lösbar,
- $\text{Rang } A = \text{Rang } [A \ b] = n$.

Die Zeilenstufenform hat dann das Bild $\begin{bmatrix} * & \dots & * \\ 0 & \dots & * \\ 0 & & \end{bmatrix}, \begin{bmatrix} b_1 \\ \vdots \\ b_n \\ 0 \end{bmatrix}$.

- Ist $\text{Rang } A = n$, hat das homogene System nur die triviale Lösung 0.
- Für $A \in K^{n \times n}$ ist $\text{Rang } A = n \Leftrightarrow A \in \text{GL}(n; K)$, also A bijektiv, und dann ist

$$x = A^{-1} \cdot b \quad \text{die eindeutige Lösung.}$$

[Numerisch ist es viel besser, die Lösung über den Gaußalgorithmus zu bestimmen!]

- Hat $A \in K^{m \times n}$ vollen Zeilenrang, also $\text{Rang } A = m$, dann ist $A: K^n \rightarrow K^m$ surjektiv, also für jedes $b \in K^m$ lösbar.
Man sagt dann: $Ax = b$ ist universell lösbar.

Kapitel 6

Die Determinante

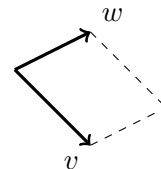
Die Determinante ist eine Funktion, die jeder *quadratischen* Matrix $A \in K^{n \times n}$ (bzw. jedem Endomorphismus) eine Zahl $\delta \in K$ zuordnet und dabei besonders schöne Eigenschaften besitzt:

- $\delta \neq 0 \Leftrightarrow A$ invertierbar/regular,
- $\det(A \cdot B) = \det A \cdot \det B$, insbesondere $1 =: \det I_n = \det A \cdot \det A^{-1} = \frac{\det A}{\det A}$,
- sie liefert ein für die Theorie wichtiges Verfahren zur Lösung von $Ax = b$ für A regulär,
- im \mathbb{R}^n lässt sie sich als Volumen mit Vorzeichen interpretieren.

6.1 Definition und erste Eigenschaften

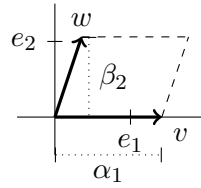
Die Determinante wird axiomatisch eingeführt, indem man schöne Eigenschaften für die Funktion fordert und dann zeigt, dass es so eine Funktion gibt. Dieser Ansatz geht auf Weierstrass zurück und lässt sich besser motivieren als die ursprüngliche Formel von Leibniz.

Die Ausgangsaxiome passen gut zum Volumen eines Parallelotops, das von gegebenen Vektoren aufgespannt wird, etwa wie im nebenstehenden Beispiel mit zwei Vektoren im \mathbb{R}^2 .



Für die Anschauung ist es günstig, das Koordinatensystem so zu legen, dass ein Vektor die Richtung von e_1 wiedergibt, etwa $v = \alpha_1 e_1$,

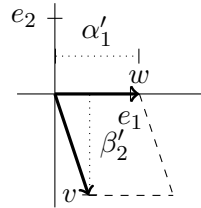
$$[v \ w] = \begin{bmatrix} \alpha_1 & \beta_1 \\ 0 & \beta_2 \end{bmatrix}$$



Das Volumen des Parallelogramms $(0, v, w, v + w)$ ist $\alpha_1 \cdot \beta_2$ (mit Vorzeichen).

Mit der gleichen Berechtigung könnte man $w = \alpha'_1 e_1$ wählen,

$$[w \ v] = \begin{bmatrix} \alpha'_1 & \beta'_1 \\ 0 & \beta'_2 \end{bmatrix},$$

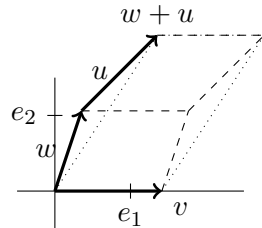


dann ist $\alpha'_1 \cdot \beta'_2 = -\alpha_1 \cdot \beta_2$ das negative Volumen!

Sind beide Vektoren linear abhängig, $\begin{matrix} \longrightarrow & \longrightarrow \\ v & w \end{matrix}$, ist das Volumen = 0.

Ist $v = \begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix}$ und addiert man zu $w = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ einen Vektor $u = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$,

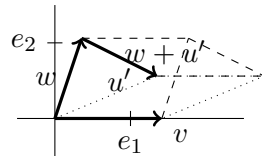
$$[v \ w + u] = \begin{bmatrix} \alpha_1 & \beta_1 + \gamma_1 \\ 0 & \beta_2 + \gamma_2 \end{bmatrix},$$



berechnet sich für das Parallelogramm $(0, v, w + u, v + w + u)$ das „Vorzeichenvolumen“

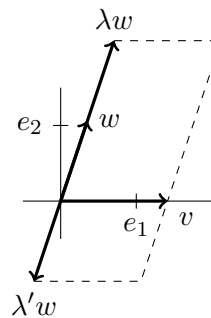
$$\alpha_1 \cdot (\beta_2 + \gamma_2)$$

als die Summe der beiden Vorzeichenvolumen zu v, w und v, u .



Streckung von w um einen Faktor $\lambda \in \mathbb{R}$ ergibt

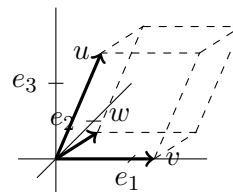
$$[v \ \lambda w] = \begin{bmatrix} \alpha_1 & \lambda\beta_1 \\ 0 & \lambda\beta_2 \end{bmatrix},$$



Vorzeichenvolumen $(v, \lambda w) = \lambda \cdot$ Vorzeichenvolumen (v, w) .
Das Vorzeichenvolumen ist also linear in der 2. Komponente, aber das gleiche Argument funktioniert auch in der 1. Komponente.

Die gleiche Vorstellung funktioniert im \mathbb{R}^3 mit drei Vektoren v, w, u , wobei für die Anschauung die Basis wieder günstig gelegt wird,

$$[v \ w \ u] = \begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ 0 & \beta_2 & \gamma_2 \\ 0 & 0 & \gamma_3 \end{bmatrix}.$$



Das Vorzeichenvolumen $(v, w, u) = \alpha_1\beta_2\gamma_3$ ist wieder linear in jeder Komponente.

Abbildungen, die in jeder Komponente linear sind, nennt man multilinear, es lohnt sich diesen Begriff gleich näher zu betrachten.

Definition 6.1 Seien K -Vektorräume V_1, \dots, V_h und W gegeben. Eine Abbildung $\varphi: V_1 \times \dots \times V_h \rightarrow W$ heißt **multilinear** oder **h -fach linear**, wenn für jedes $i \in \{1, \dots, h\}$ und jede feste Wahl von $v_j \in V_j, j \in \{1, \dots, h\} \setminus \{i\}$ die Abbildung $\varphi_i: V_i \rightarrow W, v \mapsto \varphi(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_h)$ K -linear ist.

Multilineare Abbildungen bilden wieder Untervektorräume von $W^{\times_{i=1}^h V_i}$. Hier sind vorerst nur Abbildungen mit $V_1 = \dots = V_h = V$, also $V^h \rightarrow W$ interessant.

Definition 6.2 Für gegebene K -Vektorräume V und W heißt eine h -fach lineare Abbildung $\varphi: V^h \rightarrow W$

- **symmetrisch**, falls für jede Permutation $\sigma \in S_h$

$$\varphi(v_1, \dots, v_h) = \varphi(v_{\sigma(1)}, \dots, v_{\sigma(h)}) \quad \text{gilt,}$$

- **alternierend**, falls aus $v_i = v_j$ für ein Paar (i, j) mit $i \neq j$ folgt, dass $\varphi(v_1, \dots, v_h) = 0$.

Das „Vorzeichenvolumen“ soll *nicht* symmetrisch (vertauschen wechselt eventuell das Vorzeichen) wohl aber alternierend sein, denn bei zwei gleichen Vektoren soll das Volumen 0 sein. Für alternierende multilineare Abbildungen gilt generell, dass linear abhängige Argumente stets das Bild 0 ergeben.

Beobachtung 6.3 Ist $\varphi: V^h \rightarrow W$ eine alternierende h -fach lineare Abbildung und sind $u_1, \dots, u_h \in V$ linear abhängig, gilt $\varphi(u_1, \dots, u_h) = 0$.

Beweis: s. Vorlesung □

Wie sich Vertauschungen in den Argumenten auf das Vorzeichen auswirken, lässt sich für alternierende multilineare Abbildungen gut über Permutationen beschreiben.

Beobachtung 6.4 Ist $\varphi: V^h \rightarrow W$ eine alternierende h -fach lineare Abbildung auf K -Vektorräumen V und W , dann wechselt jeweils das Vorzeichen, wenn zwei Argumente vertauscht werden,

$$\varphi(\dots, v_i, \dots, v_j, \dots) = -\varphi(\dots, v_j, \dots, v_i, \dots).$$

Allgemein gilt für jede Permutation $\sigma \in S_h$

$$\varphi(v_{\sigma(1)}, \dots, v_{\sigma(h)}) = \operatorname{sgn}(\sigma) \cdot \varphi(v_1, \dots, v_h).$$

Beweis: s. Vorlesung □

Für alternierende multilineare Abbildungen gilt sogar, dass sie auf einem durch (v_1, \dots, v_h) aufgespannten Unterraum schon vollständig durch den Wert $\varphi(v_1, \dots, v_h)$ festgelegt sind.

Lemma 6.5 Sei $\varphi: V^h \rightarrow W$ eine alternierende h -fach lineare Abbildung auf K -Vektorräumen V und W und $v_1, \dots, v_h \in V$. Für beliebige $u_j = \sum_{i=1}^h a_{ji}v_i$, $j = 1, \dots, h$ gilt

$$\varphi(u_1, \dots, u_h) = \sum_{\sigma \in S_h} \operatorname{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot \dots \cdot a_{h,\sigma(h)} \cdot \varphi(v_1, \dots, v_h).$$

Dies entspricht der geometrischen Vorstellung, dass das „Vorzeichenvolumen des Basisparallelotops“ bereits den Wert aller anderer Volumina festlegt.

Beweis: s. Vorlesung □

Ist $W = K^1 = K$, nennt man eine multilineare Abbildung $\varphi: V^h \rightarrow K$ auch **Multilinearform** oder **h -fache Linearform**. Man beachte, dass in den bisherigen Begriffen und Überlegungen nur im Beweis von B 6.3 das multiplikative Inverse gebraucht wurde, alles andere lässt sich ebensogut über kommutativen Ringen (also über Moduln statt Vektorräumen) formulieren.

Die Determinante soll nun genau so eine alternierende Multilinearform auf den Zeilen quadratischer Matrizen sein.

Im Folgenden ist $A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in K^{n \times n}$ mit Zeilen $a_i \in K^{1 \times n}$ (der einzige Grund, Zeilen statt Spalten zu wählen, ist, dass die Zeilenoperationen bereits vertrauter sein dürften).

Definition 6.6 Sei K ein Körper (oder kommutativer Ring mit Eins) und $n \in \mathbb{N}$. Eine Abbildung

$$\det: K^{n \times n} \rightarrow K, \quad A \mapsto \det A$$

heißt **Determinante**, falls folgendes gilt:

(D1) \det ist linear in jeder Zeile, d. h., für $i \in \{1, \dots, n\}$ gilt

$$\det \begin{bmatrix} \vdots \\ \alpha a_i + \beta a'_i \\ \vdots \end{bmatrix} = \alpha \det \begin{bmatrix} \vdots \\ a_i \\ \vdots \end{bmatrix} + \beta \det \begin{bmatrix} \vdots \\ a'_i \\ \vdots \end{bmatrix},$$

(D2) \det ist alternierend, d. h., gibt es in A zwei gleiche Zeilen, ist $\det A = 0$,

(D3) \det ist normiert mit $\det I_n = 1$ [für $K = \mathbb{R}$ entspricht das dem Volumen des Einheitswürfels].

Man schreibt die Determinante auch gern so:

$$|A| := \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} := \det A.$$

Vorsicht! Das sind *keine* Betragsstriche! Das Ergebnis ist *oft* negativ! ∇

Um den Namen „die Determinante“ zu rechtfertigen, fehlt noch der Nachweis, dass es eine *eindeutige* Funktion gibt, die (D1)–(D3) erfüllt.

Für $\det: K^{n \times n} \rightarrow K$ bedeutet L 6.5, dass der Wert dieser alternierenden n -fachen Linearform auf den Zeilen durch (D3)

$$\det(e_1^\top, \dots, e_n^\top) = \det I_n = 1$$

eindeutig festgelegt ist, denn für jede Zeile ist $a_i \in \text{span}(e_i^\top)_{i=1}^n = K^{1 \times n}$. Die Formel aus L 6.5 ergibt damit genau die folgende Formel von Leibniz, die zwar für die Praxis ungeeignet, für die Theorie aber oft hilfreich ist. Es bleibt nur zu zeigen, dass mit der Formel auch wirklich (D1)–(D3) erfüllt sind.

Satz 6.7 Ist K ein Körper (oder kommutativer Ring mit Eins) und $n \in \mathbb{N}$, so gibt es genau eine Determinante $\det: K^{n \times n} \rightarrow K$ und für $A = (a_{ij}) \in K^{n \times n}$ ist ihr Wert

$$\det A = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}. \quad (\text{LEIBNIZ-Formel})$$

Beweis: s. Vorlesung □

Für $n = 2$ berechnet sich die Determinante durch

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \underbrace{a_{11} \cdot a_{22}}_{\sigma=\text{id}} - \underbrace{a_{12} \cdot a_{21}}_{\sigma=(1\ 2)}.$$

Beweis: s. Vorlesung □

Bemerkung 6.10 Die Beweise von (D9) und (D10) zeigen eine gegenüber (LEIBNIZ-Formel) deutlich bessere Möglichkeit auf, die Determinante $\det A$ für Körper über (D8) zu berechnen:

Bringe A mit (D7)/(\mathcal{Z}_a) und h -facher Anwendung von (D6)/(\mathcal{Z}_p) in obere Dreiecksform $\bar{A} = \begin{bmatrix} \bar{a}_{11} & \cdots & * \\ 0 & \ddots & \\ & & \bar{a}_{nn} \end{bmatrix}$, dann ist $\det A = (-1)^h \cdot \prod_{i=1}^n \bar{a}_{ii}$. ◇

Jede Matrix $A \in K^{n \times n}$ stellt für eine bestimmte Basis \mathcal{B} einen Endomorphismus $F: V \rightarrow V$ mit $M_{\mathcal{B}}(F) = A$ dar. Damit lässt sich der Begriff der Determinante direkt auf Endomorphismen übertragen.

Definition 6.11 Sei $F \in \text{End}_K(V)$ und \mathcal{B} eine endliche Basis von V , dann heißt $\det F := \det M_{\mathcal{B}}(F)$ Determinante (des Endomorphismus).

Ist die Determinante in dieser Form wohldefiniert, also unabhängig von der Wahl der Basis? Sei \mathcal{A} eine andere Basis, dann gibt es nach K 5.15 ein $S \in \text{GL}(n; K)$ mit $M_{\mathcal{B}}(F) = SM_{\mathcal{A}}(F)S^{-1}$ und wegen (D11) folgt nun $\det(S) \det(S^{-1}) = \det(SS^{-1}) = \det I_n = 1$ und

$$\det M_{\mathcal{B}}(F) = \det(SM_{\mathcal{A}}(F)S^{-1}) \stackrel{(D11)}{=} \det(S) \det M_{\mathcal{A}}(F) \det(S^{-1}) = \det M_{\mathcal{A}}(F).$$

Also ist $\det: \text{End}_K(V) \rightarrow K$, $F \mapsto \det F$ für jeden endlichdimensionalen K -Vektorraum wohldefiniert.

6.2 Minoren und der Entwicklungssatz von Laplace

Minoren sind Determinanten von quadratischen Untermatrizen, eine genaue Definition erfolgt später. Die Beziehungen zwischen Determinante und den Determinanten der Untermatrizen sind vielfältig und nützlich und lassen sich gut über schrittweises Streichen von Zeilen-Spalten-Paaren (i, j) aus der Matrix untersuchen.

Sei R ein kommutativer Ring mit Eins. Für $A \in R^{n \times n}$ und $(i, j) \in n \times n$ bezeichne

$$A^{(i,j)} := \begin{bmatrix} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{bmatrix}.$$

Die zu A komplementäre/adjunkte Matrix $A^{\#} \in R^{n \times n}$ ist definiert durch

$$(A^{\#})_{ij} := \det A^{(j,i)}. \quad [\text{Vorsicht: } (j, i) \text{ statt } (i, j)!]$$

Wird Zeile i und Spalte j vollständig gestrichen, bezeichnet

$$A^{\overline{(i,j)}} := \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} \in R^{(n-1) \times (n-1)} \quad \text{die Streichungsmatrix .}$$

Beobachtung 6.12 $\det A^{(i,j)} = (-1)^{i+j} \det A^{\overline{(i,j)}}$.

Beweis: s. Vorlesung □

Beobachtung 6.13 Sei $A = [a^1, \dots, a^n] \in R^{n \times n}$ mit Spalten a^k , $k = 1, \dots, n$.

$$\det A^{(i,j)} = \det[a^1, \dots, a^{j-1}, e_i, a^{j+1}, \dots, a^n].$$

Beweis: s. Vorlesung □

Satz 6.14 Für $A \in R^{n \times n}$ und ihre Komplementärmatrix $A^\#$ gilt

$$A^\# \cdot A = A \cdot A^\# = (\det A) \cdot I_n.$$

Beweis: s. Vorlesung □

Satz 6.15 (Entwicklungssatz von Laplace) Sei $A \in R^{n \times n}$ mit $n \geq 2$. Für jedes $i \in \{1, \dots, n\}$ gilt (Entwicklung nach der i -ten Zeile)

$$\det A = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det A^{\overline{(i,j)}}.$$

Für jedes $j \in \{1, \dots, n\}$ gilt (Entwicklung nach der j -ten Spalte)

$$\det A = \sum_{i=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det A^{\overline{(i,j)}}.$$

Beweis: s. Vorlesung □

Der Entwicklungssatz hilft, die Leibnizformel in günstiger Reihenfolge abzu-
arbeiten, wenn es Zeilen oder Spalten mit vielen 0-Einträgen gibt.

Beispiel: $\begin{vmatrix} 3 & 2 & -4 \\ 0 & 1 & 0 \\ 1 & 7 & 0 \end{vmatrix} \stackrel{\text{Zeile 2}}{=} -0 \cdot \begin{vmatrix} 2 & -4 \\ 7 & 0 \end{vmatrix} + 1 \cdot \underbrace{\begin{vmatrix} 3 & -4 \\ 1 & 0 \end{vmatrix}}_{=4} - 0 \cdot \begin{vmatrix} 3 & 2 \\ 1 & 7 \end{vmatrix} = 4. \quad \heartsuit$

Das zu verwendende Vorzeichen merkt man sich gut über das

„Schachbrettmuster“ $\begin{vmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{vmatrix} .$

Statt des Rings R wird im folgenden wieder ein Körper K vorausgesetzt. Es steht damit wieder ein multiplikatives Inverses zur Verfügung.

Für reguläres $A \in K^{n \times n}$ liefert die Definition der adjunkten $A^\#$ mit B 6.12 über S 6.14 direkt eine Beschreibung der Inversen.

Satz 6.16 Für $A \in GL(n; K)$ ist

$$A^{-1} = \frac{1}{\det A} C^\top \quad \text{mit } c_{ij} = \det A^{(i,j)} = (-1)^{i+j} \det A^{\overline{(i,j)}} \text{ für } i, j = 1, \dots, n.$$

Beispiel: $n = 2$: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \cdot \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^\top = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. \heartsuit

Für größere n ist das weder numerisch noch praktisch sinnvoll, aber der Satz zeigt für $K = \mathbb{R}$ wegen der stetigen Abhängigkeit der Determinanten von den Matrixeinträgen, dass auch die Inverse stetig von den Matrixeinträgen abhängt (in einer Umgebung von $A \in GL(n, \mathbb{R})$) und sogar differenzierbar ist.

Die explizite Formel für A^{-1} erlaubt die Lösung von Gleichungssystemen $Ax = b$ mit $A \in GL(n; K)$ über $x = A^{-1} \cdot b$ explizit anzugeben:

$$\begin{aligned} x_i &= \sum_{j=1}^n (A^{-1})_{ij} b_j \stackrel{\text{S 6.16}}{=} \frac{1}{\det A} \sum_{j=1}^n \det A^{(j,i)} \cdot b_j \\ &\stackrel{\text{B 6.13}}{=} \frac{1}{\det A} \sum_{j=1}^n \det [a^1, \dots, a^{i-1}, e_j, a^{i+1}, \dots, a^n] \cdot b_j \\ &\stackrel{(D1)}{=} \frac{1}{\det A} \det [a^1, \dots, a^{i-1}, \underbrace{\sum_{j=1}^n b_j e_j}_{=b}, a^{i+1}, \dots, a^n]. \end{aligned}$$

Damit ist folgende für die Theorie wichtige analytische Beschreibung der Lösung von Gleichungssystemen bewiesen.

Satz 6.17 (CRAMERsche Regel) Sei $A = [a^1, \dots, a^n] \in GL(n; K)$, $b \in K^n$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ die eindeutige Lösung von $Ax = b$. Für jedes $i \in \{1, \dots, n\}$ ist

$$x_i = \frac{\det [a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n]}{\det A}.$$

Dies zeigt für $K = \mathbb{R}$, dass die Lösung stetig von den Koeffizienten in A und der rechten Seite b abhängt.

Beispiel:

$$\begin{array}{rcl} 2x_1 & +3x_2 & = 9 \\ & \frac{1}{2}x_2 & +x_3 = \frac{5}{2} \\ \frac{1}{3}x_1 & +\frac{1}{2}x_2 & +x_3 = \frac{7}{2} \end{array} \quad \det A = 1 + 1 + 0 - 1 - 0 - 0 = 1$$

$$\begin{aligned}
 x_1 &= \det \begin{bmatrix} 9 & 3 & 0 \\ \frac{5}{2} & \frac{1}{2} & 1 \\ \frac{7}{2} & \frac{1}{2} & 1 \end{bmatrix} / 1 = -1 \cdot \underbrace{\begin{vmatrix} 9 & 3 \\ 7 & \frac{1}{2} \end{vmatrix}}_{\frac{9-21}{2}=-6} + 1 \cdot \underbrace{\begin{vmatrix} 9 & 3 \\ \frac{5}{2} & \frac{1}{2} \end{vmatrix}}_{\frac{9-15}{2}=-3} = 3 \\
 x_2 &= \det \begin{bmatrix} 2 & 9 & 0 \\ 0 & \frac{5}{2} & 1 \\ \frac{1}{3} & \frac{7}{2} & 1 \end{bmatrix} / 1 = -1 \cdot \underbrace{\begin{vmatrix} 2 & 9 \\ \frac{1}{3} & \frac{7}{2} \end{vmatrix}}_{7-3=4} + 1 \cdot \underbrace{\begin{vmatrix} 2 & 9 \\ 0 & \frac{5}{2} \end{vmatrix}}_{=5} = 1 \\
 x_3 &= \det \begin{bmatrix} 2 & 3 & 9 \\ 0 & \frac{1}{2} & \frac{5}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{7}{2} \end{bmatrix} / 1 = 2 \cdot \underbrace{\begin{vmatrix} \frac{1}{2} & \frac{5}{2} \\ \frac{1}{2} & \frac{7}{2} \end{vmatrix}}_{\frac{7-5}{4}=\frac{1}{2}} + \frac{1}{3} \cdot \underbrace{\begin{vmatrix} 3 & 9 \\ \frac{1}{2} & \frac{5}{2} \end{vmatrix}}_{\frac{15-9}{2}=3} = 2
 \end{aligned}$$

♡

Kann man den Rang von $A \in K^{m \times n}$ aus den Determinanten von Untermatrizen ablesen?

Eine Matrix $A' \in K^{h \times h}$ mit $h < \min\{m, n\}$ heißt *h-reihige Teil-/Untermatrix* von A , wenn man A durch Zeilen und Spaltenvertauschungen in die Form $\begin{bmatrix} A' & * \\ * & * \end{bmatrix}$ bringen kann.

Z. B. erhält man durch Streichung von $m - h$ Zeilen und $n - h$ Spalten eine h -reihige Teilmatrix.

det A' heißt dann (*h-reihiger*) Minor von A .

Satz 6.18 Für $A \in K^{m \times n}$, $r \in \mathbb{N}$ mit $0 < r \leq \min\{m, n\}$ sind äquivalent:

- (a) $r = \text{Rang } A$.
- (b) Es gibt einen r -reihigen Minor $\neq 0$ und für $h > r$ ist jeder h -reihige Minor = 0.

Beweis: s. Vorlesung □

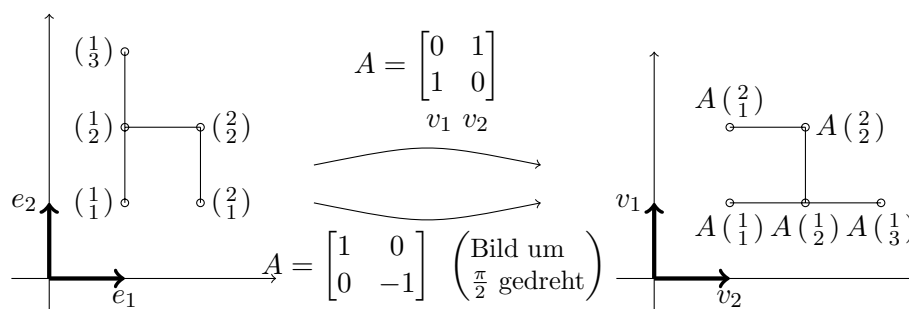
6.3 Geometrische Interpretation im \mathbb{R}^n

Für $K = \mathbb{R}$ hat der Absolutbetrag der Determinante die Interpretation als Volumen des von den Zeilen- oder Spaltenvektoren aufgespannten Paralleleptops.

Wegen $\det I_n = 1$ gibt $|\det A|$ an, um wieviel durch die Abbildung $A \cdot x$ (die Spalten von A sind die Bilder der kanonischen Basisvektoren) das Volumen des Bildes des Einheitswürfels vergrößert bzw. verkleinert wird.

Welche Bedeutung hat das Vorzeichen von $\det A$?

Das Vorzeichen wechselt z. B., wenn man zwei Spaltenvektoren vertauscht, also in der Basis des Bildes zwei Basisvektoren ihren Platz wechseln, oder wenn man das Vorzeichen eines der Basisvektoren ändert.



Das Bild kann man nicht mehr aus dem Original durch „Drehen oder Strecken“ erhalten, man muss auch einmal „Spiegeln“.

Im \mathbb{R}^3 hilft die Vorstellung, mit Daumen, Zeigefinger und Mittelfinger der rechten Hand e_1, e_2 und e_3 in dieser Reihenfolge nachzuformen. Bildet man nun e_1 auf $v_1 = e_3, e_3$ auf $v_3 = e_1$ ab und lässt e_2 mit $v_2 = e_2$ unverändert, kann man die rechte Hand nicht so drehen, dass Daumen, Zeige- und Mittelfinger mit v_1, v_2, v_3 zur Deckung kommen, man muss dazu die linke Hand verwenden. Tauscht man nun nochmals zwei Vektoren aus, geht es wieder nur mit der rechten Hand.

Definition 6.19 Ein Automorphismus F eines endlichen \mathbb{R} -Vektorraumes $V (\simeq \mathbb{R}^n)$ heißt

- orientierungstreu, falls $\det F > 0$, und
- orientierungsuntreu, falls $\det F < 0$.

Zwei Basen $\mathcal{A} = (v_1, \dots, v_n), \mathcal{B} = (w_1, \dots, w_n)$ von V heißen gleichorientiert, in Zeichen $\mathcal{A} \sim \mathcal{B}$, falls $\det F > 0$ für den nach S 4.8 eindeutig bestimmten Automorphismus F mit $F(v_i) = w_i, i = 1, \dots, n$, gilt.

Die Relation \sim erklärt eine Äquivalenzrelation auf der Menge X aller Basen von V [warum?] und eine Partition in zwei Äquivalenzklassen

$$X = X^\bullet \cup X_\bullet.$$

Für $V = \mathbb{R}^n$ kann man diese Zerlegung über die bijektive Abbildung

$$M: X \rightarrow \text{GL}(n; \mathbb{R}), \mathcal{A} \mapsto A = M(\mathcal{A}) \text{ (mit Basisvektoren als Spalten)}$$

beschreiben: $A \sim B \Leftrightarrow \det M(\mathcal{A}) \cdot \det M(\mathcal{B}) > 0$.

$\text{GL}(n; \mathbb{R})$ wird dadurch wie folgt partitioniert:

$$\begin{aligned} G_+ &:= \{A \in \text{GL}(n; \mathbb{R}) : \det A > 0\}, \\ G_- &:= \{A \in \text{GL}(n; \mathbb{R}) : \det A < 0\}. \end{aligned}$$

Man kann zeigen, dass zwei Basen des \mathbb{R}^n genau dann in der gleichen Klasse G_+ bzw. G_- liegen, wenn sie sich „stetig ineinander verformen“ lassen.

Definition 6.20 Sind $A, B \in \text{GL}(n; \mathbb{R})$, so versteht man unter einem Weg von A nach B eine stetige Abbildung von einem Intervall $[a, b] \subset \mathbb{R}$ nach GL ,

$$\psi: [a, b] \rightarrow \text{GL}(n; \mathbb{R}), \quad t \mapsto \psi(t) = (\psi_{ij}(t))$$

mit $\psi(a) = A$ und $\psi(b) = B$.

Die Matrix $[\psi_{ij}(t)]$ muss für jedes $t \in [a, b]$ invertierbar ($\in \text{GL}$) sein. Zwei Matrizen heißen verbindbar/zusammenhängend, $A \sim B$, wenn es einen Weg von A nach B gibt.

Mit topologischen Hilfsmitteln erhält man (siehe die Beweisskizze in Fischer 3.4.4)

Satz 6.21 Für $A, B \in \text{GL}(n; \mathbb{R})$ sind äquivalent

- (a) $\det A \cdot \det B > 0$,
- (b) A und B sind verbindbar.

Also sind G_+ und G_- genau die zwei Zusammenhangskomponenten in Bezug auf stetige Transformierbarkeit.

Kapitel 7

Polynomdivision und Primfaktorzerlegung

Für die ganzen Zahlen ist das Konzept der Primfaktorzerlegung sehr vertraut. Genau das gleiche Konzept lässt sich auch für Polynome aus dem Polynomring $K[x]$ über einem beliebigen Körper K entwickeln.

Zur Erinnerung: $K[x]$ ist ein nullteilerfreier kommutativer Ring mit Eins (B 2.25). Für ein Polynom $f = \sum_{i=0}^n a_i x^i \in K[x]$ mit $a_n \neq 0$ ist $\deg f = n$ der Grad von f und $a_{\deg f}(f) = a_n$ der führende Koeffizient von f . Das Nullpolynom $f = 0$ hat $\deg f = -\infty$, ein f mit $\deg f \leq 0$ heißt konstant, ein f mit $\deg f = 0$ heißt invertierbar (denn $f = a_0 \neq 0$ und $f^{-1} = a_0^{-1}$). Ein Polynom $f \neq 0$ heißt normiert, falls $a_{\deg f} = 1$. Wegen der Nullteilerfreiheit gilt für $f, g \in K[x]$ stets $\deg(f \cdot g) = \deg f + \deg g$ (wobei $\deg 0 = -\infty = -\infty + n$ für $n \in \mathbb{N}_0$) sowie die Kürzungsregel, d. h., für $g \neq 0$ ist $fg = hg \Leftrightarrow f = h$.

Gilt $h \cdot g = f$, nennt man g und h Teiler von f und schreibt $g|f$ bzw. $h|f$; Teilbarkeit ist transitiv und wegen der 1 reflexiv, aber i. A. nicht symmetrisch. Für $g \neq 0$ testet man $g|f$, indem man denselben Divisionsalgorithmus wie in der Grundschule nun auf Polynomen durchführt und feststellt, ob dabei ein Rest bleibt.

Beispiel: $f = 4x^3 - 2x^2 + 5$, $g = 2x^2 - 3x$,

$$\begin{array}{r} (4x^3 - 2x^2 + 5) \div (2x^2 - 3x) = 2x + 2 \quad \text{mit Rest } 6x + 5, \\ \begin{array}{r} 4x^3 - 2x^2 + 5 \\ -4x^3 + 6x^2 \\ \hline 4x^2 + 0x \\ -4x^2 + 6x \\ \hline 6x + 5 \end{array} \end{array}$$

also ist $f = (2x + 2)g + 6x + 5$. ♡

Satz 7.1 (Polynomdivision) Für $f, g \in K[x]$ mit $g \neq 0$ gibt es eindeutig bestimmte Polynome $q, r \in K[x]$ mit $f = g \cdot q + r$ und $\deg r < \deg g$.

Beweis: s. Vorlesung □

Wie für ganze Zahlen lassen sich nun ganz analoge Teilbarkeitsresultate für Polynome entwickeln. Dabei übernimmt $\deg f$ für $f \in K[x]$ die Rolle von $|z|$ für $z \in \mathbb{Z}$.

Beobachtung 7.2 Seien $f, g, f_1, f_2 \in K[x]$.

- (a) $g|f \wedge f|h \Rightarrow g|h$,
- (b) $g|f \wedge f|g \Leftrightarrow f = ag$ für ein invertierbares $a \in K[x]$,
- (c) $g|f_1 \wedge g|f_2 \Rightarrow g|(p_1f_1 + p_2f_2)$ für alle $p_1, p_2 \in K[x]$.

Beweis: s. Vorlesung □

Polynome mit $f = a \cdot g$ für ein $a \in K \setminus \{0\}$ haben die gleichen Teilbarkeits-eigenschaften, sie heißen assoziiert, und das erklärt eine Äquivalenzrelation auf $K[x]$.

Definition 7.3 Seien $d, f_1, \dots, f_n \in K[x], n \in \mathbb{N}$.

- d heißt gemeinsamer Teiler von f_1, \dots, f_n , falls $d|f_i$ für $1 \leq i \leq n$.
- d heißt größter gemeinsamer Teiler (ggT) von f_1, \dots, f_n , falls d gemeinsamer Teiler von f_1, \dots, f_n ist und $d'|d$ für alle $d' \in K[x]$ mit $d'|f_i, i = 1, \dots, n$, gilt.
- $\text{ggT}(f_1, \dots, f_n)$ bezeichne den normierten ggT von f_1, \dots, f_n .
- f_1, \dots, f_n heißen teilerfremd, falls $\text{ggT}(f_1, \dots, f_n) = 1$.

Für $n \geq 2$ ist nicht klar, dass es einen ggT gibt. Wir zeigen das zuerst für $n = 2$.

Satz 7.4 Zu $f, g \in K[x] \setminus \{0\}$ gibt es einen ggT $d \in K[x]$, und für d gibt es $p, q \in K[x]$ mit $d = pf + qg$.

Beweis: s. Vorlesung □

Um den normierten $\text{ggT}(f, g)$ zu berechnen, müssen nachträglich d, p, q mit $(a_{\deg_d}(d))^{-1}$ multipliziert werden.

Korollar 7.5 Für $f_1, \dots, f_n \in K[x] \setminus \{0\}, n \in \mathbb{N}$, gibt es einen normierten größten gemeinsamen Teiler $d = \text{ggT}(f_1, \dots, f_n) \in K[x]$ sowie $h_1, \dots, h_n \in K[x]$ mit $d = h_1f_1 + \dots + h_nf_n$.

Beweis: s. Vorlesung □

Definition 7.6 Ein nicht konstantes Polynome $f \in K[x]$ heißt irreduzibel, falls aus $f = gh$ mit $g, h \in K[x]$ folgt, dass g oder h invertierbar ist, sonst heißt f reduzibel.

Wie Primzahlen in \mathbb{Z} hat ein irreduzibles f nur „triviale“ Teiler d , die selbst invertierbar sind oder die Gestalt $d = a \cdot f$ für ein $a \in K$ haben. Insbesondere sind irreduzible Polynome $f, g \in K[x] \setminus \{0\}$ mit $f \neq g$ stets teilerfremd, $\text{ggT}(f, g) = 1$.

Lemma 7.7 Für $f, g, h \in K[x] \setminus \{0\}$ mit $\text{ggT}(f, g) = 1$ gilt: $f|gh \Rightarrow f|h$.

Beweis: s. Vorlesung □

Satz 7.8 Jedes nicht konstante Polynome $f \in K[x]$ lässt sich als Produkt $f = a \prod_{i=1}^n p_i$ von normierten irreduziblen Polynomen p_i und einem $a \in K$ schreiben. Die Polynome sind bis auf die Reihenfolge eindeutig bestimmt.

Beweis: s. Vorlesung □

Definition 7.9 Für $f \in K[x] \setminus \{0\}$ heißt die Darstellung $f = a \prod_{i=1}^k p_i^{r_i}$ mit $a \in K$, paarweise verschiedenen normierten irreduziblen Polynomen p_i , $r_i \in \mathbb{N}$, $i = 1, \dots, k$, $k \in \mathbb{N}_0$ Primfaktorzerlegung oder Primfaktordarstellung von f .

Ein $g \in K[x] \setminus \{0\}$ heißt ein kleinstes gemeinsames Vielfaches (kgV) von $f_1, \dots, f_n \in K[x] \setminus \{0\}$, $n \in \mathbb{N}$, falls $f_i|g$, $i = 1, \dots, n$ und für $h \in K[x]$ mit $f_i|h$, $i = 1, \dots, n$, stets $g|h$ gilt.

$\text{kgV}(f_1, \dots, f_n)$ bezeichne das normierte kgV von f_1, \dots, f_n .

Ü: Um irreduzible Polynome besser zusammenfassen zu können, seien hier Potenzen mit dem Wert 0 in der Primfaktorzerlegung erlaubt. Zeige: Ist $f_i = a_i \prod_{j=1}^m p_j^{r_j^{(i)}}$ die Primfaktorzerlegung von f_i , $i = 1, \dots, n$, mit $r_j^{(i)} \in \mathbb{N}_0$, dann ist

$$\begin{aligned} \text{ggT}(f_1, \dots, f_n) &= \prod_{j=1}^m p_j^{\min\{r_j^{(i)} : i=1, \dots, n\}}, \\ \text{kgV}(f_1, \dots, f_n) &= \prod_{j=1}^m p_j^{\max\{r_j^{(i)} : i=1, \dots, n\}}. \end{aligned}$$

⊠

Ein normiertes irreduzibles Polynom vom Grad 1 heißt Linearfaktor und kann immer in der Form $(x - \lambda)$ mit $\lambda \in K$ geschrieben werden. Enthält die Primfaktorzerlegung von f den Linearfaktor $(x - \lambda)$, dann ergibt Einsetzen von λ in f stets $f(\lambda) = 0$.

Ein $\lambda \in K$ heißt Nullstelle oder Wurzel von $f \in K[x]$, falls $f(\lambda) = 0$.

Lemma 7.10 $\lambda \in K$ ist genau dann Nullstelle von $f \in K[x]$, wenn es ein (dann eindeutiges) $q \in K[x]$ gibt mit $f = (x - \lambda) \cdot q$. Für dieses q gilt $\deg q = \deg f - 1$.

Beweis: s. Vorlesung □

Das Beispiel $K = \mathbb{R}$ und $f = x^2 + 1$ zeigt, dass es i. A. keine Nullstellen geben muss.

Ist $f = (x - \lambda)q$ und λ auch Nullstelle von q , heißt λ mehrfache Nullstelle von f und für $f \neq 0$ bezeichnet

$$\mu(f; \lambda) := \max\{r \in \mathbb{N}_0 : f = (x - \lambda)^r \cdot g, g \in K[x]\}$$

die Vielfachheit der Nullstelle. Sie ist der Exponent von $(x - \lambda)$ in der Primfaktorzerlegung von f .

Korollar 7.11 *Ist $0 \neq f \in K[x]$ und sind $\lambda_1, \dots, \lambda_k \in K$ sämtliche paarweise verschiedenen Nullstellen von f mit Vielfachheiten $r_i = \mu(f; \lambda_i)$, so gibt es ein $q \in K[x] \setminus \{0\}$ mit*

$$f = q \cdot \prod_{i=1}^k (x - \lambda_i)^{r_i}, \deg q = \deg f - \sum_{i=1}^k r_i (\geq 0) \text{ und } q \text{ hat keine Nullstellen.}$$

Beweis: s. Vorlesung □

Korollar 7.12 *Stimmen zwei Polynome $f, g \in K[x]$ mit $\deg f < n$, $\deg g < n$ für n unterschiedliche Werte $\lambda_1, \dots, \lambda_n \in K$ überein, so ist $f = g$.*

Beweis: s. Vorlesung □

Korollar 7.13 *Für unendliches K ist $\hat{\cdot} : K[x] \rightarrow \text{Abb}(K, K)$, $f \mapsto \hat{f}$ injektiv.*

Beweis: s. Vorlesung □

Für das q aus K 7.11 gilt im „besten“ Fall $q = a_0 \in K$, im „schlechtesten“ $q = f$. Für $K = \mathbb{C}$ bewies Gauß 1799, dass immer der beste eintritt.

Satz 7.14 (Fundamentalsatz der Algebra) *Jedes Polynom $f \in \mathbb{C}[x]$ mit $\deg f > 0$ hat mindestens eine Nullstelle.*

Der Beweis erfordert Hilfsmittel der Analysis und erfolgt dort oder in der Vorlesung Funktionentheorie.

Korollar 7.15 *Jedes $0 \neq f \in \mathbb{C}[x]$ zerfällt in Linearfaktoren: Für $n = \deg f$ gibt es (nicht notwendig verschiedene) $\lambda_1, \dots, \lambda_n, a \in \mathbb{C}$ mit*

$$f = a(x - \lambda_1) \cdots (x - \lambda_n) = a \prod_{i=1}^n (x - \lambda_i).$$

Ein Körper K mit der Eigenschaft, dass jedes $f \in K[x]$ wenigstens eine Nullstelle hat, heißt algebraisch abgeschlossen. Die Menge der reellen Zahlen \mathbb{R} ist leider nicht algebraisch abgeschlossen (z. B. $x^2 + 1$).

Faktorisierung von Polynomen über $K = \mathbb{R}$

Da \mathbb{R} als Unterkörper von \mathbb{C} aufgefasst werden kann, erlaubt $\mathbb{R}[x] \subset \mathbb{C}[x]$ zu zeigen, dass in $\mathbb{R}[x]$ Polynome noch recht gut zerlegt werden können.

Lemma 7.16 *Sei $f \in \mathbb{R}[x]$ und $\lambda \in \mathbb{C}$ eine Nullstelle von f über \mathbb{C} , dann ist auch deren konjugiert komplexe $\bar{\lambda}$ eine Nullstelle von f mit gleicher Vielfachheit $\mu(f; \bar{\lambda}) = \mu(f; \lambda)$.*

Beweis: s. Vorlesung □

Zusammen mit Fundamentalsatz der Algebra zeigt das Lemma

Satz 7.17 *Jedes reelle Polynom $f \in \mathbb{R}[x]$ ist darstellbar als*

$$f = a(x - \lambda_1) \cdot \dots \cdot (x - \lambda_r) \cdot g_1 \cdot \dots \cdot g_m = a \prod_{i=1}^r (x - \lambda_i) \cdot \prod_{i=1}^m g_i$$

mit $a, \lambda_1, \dots, \lambda_r \in \mathbb{R}$ und $g_i = (x - z_i)(x - \bar{z}_i)$, $i = 1, \dots, m$.

Insbesondere ist $\deg f = r + 2m$, und $g_i \in \mathbb{R}[x]$ sind normierte Polynome vom Grad 2 ohne reelle Nullstellen.

In der Praxis ist es schwer, die Nullstellen zu bestimmen.

Für Polynome vom Grad 2 ist aus der Schule quadratisches Ergänzen bekannt, das hier o. E. für den normierten Fall betrachtet wird,

$$\begin{aligned} x^2 + px + q &= 0 && | + \left(\frac{p^2}{4} - q\right), \\ x^2 + 2\frac{p}{2}x + \frac{p^2}{4} &= \frac{p^2}{4} - q, \\ \left(x + \frac{p}{2}\right)^2 &= \frac{p^2}{4} - q && \stackrel{p^2/4 - q \geq 0}{\Rightarrow} x_{\pm} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \end{aligned}$$

Es gibt keine reelle Nullstellen, falls die Diskriminante $\frac{p^2}{4} - q$ echt kleiner Null ist, dann sind $z = -\frac{p}{2} + i\sqrt{q - \frac{p^2}{4}} \in \mathbb{C}$ und \bar{z} die komplexen Nullstellen.

Ähnliches geht noch für Grad 3 und Grad 4 (komplizierter), aber ab Grad 5 kann es aus algebraischen Gründen keine allgemeine Formel geben. Man benötigt dann numerische Verfahren (z. B. das Newton-Verfahren) zur näherungsweise Berechnung.

Ist $f = \prod_{i=1}^n (x - \lambda_i) \in K[x]$ ein normiertes Polynom mit bekannter Faktorisierung, lassen sich die Koeffizienten sehr leicht berechnen (*Wurzelsatz von*

Vieta):

$$\begin{aligned}
 a_0 &= (-1)^n \prod_{i=1}^n \lambda_i \\
 a_1 &= (-1)^{n-1} (\lambda_2 \lambda_3 \cdots \lambda_n + \lambda_1 \lambda_3 \cdots \lambda_n + \cdots + \lambda_1 \cdots \lambda_{n-1}) \\
 &= (-1)^{n-1} \sum_{1 \leq i_1 < \cdots < i_{n-1} \leq n} \prod_{j=1}^{n-1} \lambda_{i_j} \\
 a_2 &= (-1)^{n-2} \sum_{1 \leq i_1 < \cdots < i_{n-2} \leq n} \prod_{j=1}^{n-2} \lambda_{i_j} \\
 &\vdots \\
 a_{n-2} &= \sum_{1 \leq i < j \leq n} \lambda_i \lambda_j, \quad a_{n-1} = - \sum_{i=1}^n \lambda_i, \quad a_n = 1.
 \end{aligned}$$

Mit den elementarsymmetrischen Funktionen

$$s_k(\lambda_1, \dots, \lambda_n) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \prod_{j=1}^k \lambda_{i_j} \quad (\text{mit } \sum_{\emptyset} = 0, \prod_{\emptyset} = 1)$$

ergibt das die kurze Schreibweise

$$a_k = (-1)^{n-k} s_{n-k}(\lambda_1, \dots, \lambda_n), \quad k = 0, \dots, n.$$

Für die Matrizenrechnung ist es wichtig, für ein Polynom $f \in \mathbb{C}[x]$ mit nur reellen Nullstellen einfach erkennen zu können, ob alle Nullstellen das gleiche Vorzeichen haben.

Beobachtung 7.18 (Vorzeichenregel)

Hat $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n \in \mathbb{C}[x]$ [normiert!] nur reelle Nullstellen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, dann gilt:

- (a) $\lambda_i < 0$ (≤ 0) für alle $1 \leq i \leq n \Leftrightarrow a_j > 0$ (≥ 0) für alle $0 \leq j \leq n-1$.
- (b) $\lambda_i > 0$ (≥ 0) für alle $1 \leq i \leq n \Leftrightarrow (-1)^{n-j} a_j > 0$ (≥ 0) für alle $0 \leq j \leq n-1$ (alternierendes Vorzeichen!).

Beweis: s. Vorlesung

□

Kapitel 8

Endomorphismen von Vektorräumen

Sind V und W K -Vektorräume mit $\dim V = n$, $\dim W = m$, und ist $F: V \rightarrow W$ linear mit $\text{Rang } F = r$, dann gibt es nach K 5.4 immer Basen \mathcal{A} von V und \mathcal{B} von W , sodass $M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in K^{m \times n}$. Jede Matrix $A \in K^{m \times n}$ mit $\text{Rang } A = r$ ist äquivalent zu dieser Normalform unter geeigneter Koordinatentransformation SAT^{-1} (s. D 5.16, L 5.17).

Für $F: V \rightarrow V$ linear, also $F \in \text{End}_K(V)$, sucht man nicht unterschiedliche Basen \mathcal{A} und \mathcal{B} , sondern *eine* Basis \mathcal{A} mit $M_{\mathcal{A}}(F) \in K^{n \times n}$ in einer geeigneten Normalform, dass man den Charakter von F möglichst unmittelbar ablesen kann, d. h., man möchte schon aus der Koordinatendarstellung in dieser Basis erkennen, wie sich ein Vektor „geometrisch“ verändert, wenn F einmal oder mehrmals hintereinander auf den Vektor angewandt wird.

Für eine Matrix $A \in K^{n \times n}$ beschreibt die Ähnlichkeitstransformation SAS^{-1} den Übergang zu einer neuen Basis (K 5.15). Man sucht also innerhalb jeder Äquivalenzklasse von ähnlichen Matrizen nach besonders schönen Repräsentanten. In voller Allgemeinheit wird dies die sogenannte JORDANsche Normalform sein, die in wichtigen Spezialfällen zu einer Diagonalmatrix von sogenannten Eigenwerten für eine Basis dazupassender Eigenvektoren wird.

8.1 Invariante Untervektorräume, Eigenwerte und Eigenvektoren

Definition 8.1 Sei V ein K -Vektorraum [auch ∞ -dim.] und $F \in \text{End}_K(V)$.

Ein Untervektorraum U von V heißt *invarianter Untervektorraum* von F , kurz *F-invariant*, wenn $F(U) \subseteq U$ gilt, also $u \in U \Rightarrow F(u) \in U$.

Sind $U_i, i = 1, \dots, m$ invariante Untervektorräume von F mit $V = \bigoplus_{i=1}^m U_i$, nennt man sie reduzierende (invariante) Untervektorräume von F .

Beispiel: Für $F \in \text{End}_K(V)$ sind $V, \text{Ker } F, \text{Im } F$ stets F -invariant. \heartsuit

Jeder Anteil eines Vektors, der einem invarianten Untervektorraum angehört, wird durch F auch wieder in diesen Unterraum abgebildet. Aus diesem Unterraum kommt auch bei mehrmaligem Anwenden von F nichts mehr heraus, wohl aber kann bei jeder Anwendung ein neuer Anteil von „außerhalb“ in diesen Unterraum hineingelangen. Invariante Untervektorräume sind so etwas wie schwarze Löcher der Abbildung. Bei reduzierenden invarianten Untervektorräumen ist das Weltall vollständig in Einflussgebiete der Schwarzen Löcher von F aufgeteilt: F lässt nicht zu, dass ein Anteil eines Einflussgebiets je in ein anderes Einflussgebiet wechselt und jedes Einflussgebiet ist ein kleines Weltall für sich ohne Interaktion mit dem Rest.

Ist U ein invarianter Untervektorraum von F mit Basis $\mathcal{U} = (u_1, \dots, u_k)$ und $\dim V = n$, ergänze \mathcal{U} mit $\mathcal{W} = (w_1, \dots, w_{n-k})$ zu einer Basis $\mathcal{A} = (u_1, \dots, u_k, w_1, \dots, w_{n-k})$ von V , dann ist (wegen (Mat) in K 5.3)

$$M_{\mathcal{A}}(F) = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \in K^{n \times n} \quad \text{mit} \quad \begin{array}{l} A_{11} \in K^{k \times k}, \quad A_{12} \in K^{k \times (n-k)}, \\ A_{22} \in K^{(n-k) \times (n-k)}. \end{array}$$

und $M_{\mathcal{U}}(F|_U) = A_{11} \in K^{k \times k}$. Der A_{11} -Block gibt also F auf dem Unterraum U vollständig wieder.

Normalerweise ist $A_{12} \neq 0$, denn die Bilder der w_j werden meist auch einige Koordinaten $\neq 0$ für Basisvektoren u_j in U haben. Damit $A_{12} = 0$ sein kann, muss man \mathcal{W} so geschickt wählen können (das geht sehr oft nicht), dass $W = \text{span } \mathcal{W}$ selbst invarianter Untervektorraum von F ist, also mit U ein Paar reduzierender Untervektorräume bildet. In diesem Fall gilt dann

$$M_{\mathcal{A}}(F) = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \in K^{n \times n}, \quad A_{11} = M_{\mathcal{U}}(F|_U), A_{22} = M_{\mathcal{W}}(F|_W).$$

Beobachtung 8.2 Ist $F \in \text{End}_K(V)$ und sind $U_i, i = 1, \dots, m$ reduzierende Untervektorräume von F , so gilt $F|_{U_i} \in \text{End}_K(U_i), i = 1, \dots, m$, und für $v \in V$ mit $v = w_1 + \dots + w_m, w_i \in U_i, i = 1, \dots, m$, ist $F(v) = \sum_{i=1}^m F|_{U_i}(w_i)$.

Falls $\dim V = n$ endlich ist und $\mathcal{U}_i = (u_1^{(i)}, \dots, u_{n_i}^{(i)})$ Basen der U_i bilden, ist für $\mathcal{A} = (u_1^{(1)}, \dots, u_{n_m}^{(m)})$ die Matrix $M_{\mathcal{A}}(F) = \begin{bmatrix} A^{(1)} & & 0 \\ & \ddots & \\ 0 & & A^{(m)} \end{bmatrix}$ in Blockdiagonalform mit $A^{(i)} = M_{\mathcal{U}_i}(F|_{U_i}) \in K^{n_i \times n_i}$.

Beweis: s. Vorlesung □

Um die Struktur von F zu verstehen reicht es daher, die Struktur von $F|_U$ auf nicht weiter reduzierbaren Untervektorräumen U zu untersuchen.

Beispiel: Die Matrix $A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$ (Drehung um Winkel α) beschreibt für $\alpha \neq k\pi, k \in \mathbb{Z}$ eine Abbildung $A \in \text{End}(\mathbb{R}^2)$, die keinen invarianten eindimensionalen Untervektorraum hat.

Für $\alpha \in \{k\pi : k \in \mathbb{Z}\}$ bilden $\{\mu e_1 : \mu \in \mathbb{R}\}$ und $\{\mu e_2 : \mu \in \mathbb{R}\}$ ein Paar reduzierender Untervektorräume von \mathbb{R}^2 . ♡

Eindimensionale invariante Untervektorräume entsprechen geometrisch Geraden, die durch F auf sich selbst abgebildet werden. Weil F linear ist, können die Vektoren dieser Geraden durch Anwendung von F nur um konstante Faktoren gestreckt/skaliert werden.

Ist U ein eindimensionaler invarianter Untervektorraum (Basis u_1), dann gilt

$$F(u) = \lambda \cdot u \quad \text{für alle } u \in U = \{\alpha u_1 : \alpha \in K\} \text{ und ein festes } \lambda \in K,$$

und jedes $u \neq 0$ mit $F(u) = \lambda u$ führt zu einem eindimensionalen invarianten Untervektorraum.

Definition 8.3 Sei $F \in \text{End}_K(V)$ für einen K -Vektorraum V .

Ein $\lambda \in K$ heißt Eigenwert (EW), wenn es ein $v \in V \setminus \{0\}$ gibt mit $F(v) = \lambda v$.

Jedes $v \in V \setminus \{0\}$ mit $F(v) = \lambda v$ heißt Eigenvektor (EV) von F (zum Eigenwert λ).

F heißt diagonalisierbar, wenn V eine Basis aus Eigenvektoren besitzt.

Beobachtung 8.4 Sei V ein K -Vektorraum mit $\dim V = n \in \mathbb{N}$. Ein $F \in \text{End}_K(V)$ ist genau dann diagonalisierbar, wenn es eine Basis $\mathcal{A} = (v_1, \dots, v_n)$ gibt, sodass

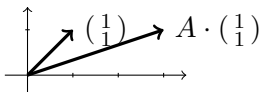
$$M_{\mathcal{A}}(F) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} =: \text{Diag}(\lambda_1, \dots, \lambda_n) \text{ eine } \underline{\text{Diagonalmatrix}} \text{ ist.}$$

v_i ist dann Eigenvektor von F zu λ_i für $i = 1, \dots, n$.

Insbesondere ist eine Matrix $A \in K^{n \times n}$ genau dann diagonalisierbar, wenn es ein $S \in \text{GL}(n; K)$ mit $S \cdot A \cdot S^{-1}$ Diagonalmatrix gibt. Die Spalten von S^{-1} enthalten dabei die Koordinaten zu den Vektoren einer Basis aus Eigenvektoren von A .

Beweis: s. Vorlesung □

Vorsicht! Selbst für diagonalisierbare Endomorphismen sind i. A. die meisten Vektoren *keine* Eigenvektoren! So ist etwa für $A = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

$$A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \neq \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$


Es ist gut zu erkennen, dass die jeweiligen Anteile des Vektors an den Untervektorräumen jeweils für sich unterschiedlich gestreckt und dann wieder zusammengesetzt werden. ∇

Eigenvektoren zu unterschiedlichen Eigenwerten sind stets linear unabhängig.

Lemma 8.5 Sind v_1, \dots, v_m Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m$ von $F \in \text{End}_K(V)$, dann sind sie linear unabhängig und $m \leq \dim V$.

Beweis: s. Vorlesung \square

Das führt zu der folgenden hinreichenden Bedingung für Diagonalisierbarkeit.

Satz 8.6 Ist V ein n -dimensionaler K -Vektorraum und hat $F \in \text{End}_K(V)$ n unterschiedliche Eigenwerte $\lambda_1, \dots, \lambda_n$, dann ist F diagonalisierbar.

Beweis: s. Vorlesung \square

Für $\dim V = n$ kann es höchstens n unterschiedliche Eigenwerte geben, aber oft gibt es deutlich weniger, denn wie im Drehungsbeispiel kann es gar keine Eigenwerte oder, wie bei λI_n , einen großen Untervektorraum zum selben Eigenwert geben.

Definition 8.7 Für $F \in \text{End}_K(V)$ und $\lambda \in K$ heißt

$$\text{Eig}(F; \lambda) := \{v \in V : F(v) = \lambda v\} \quad \text{der \underline{Eigenraum} von } F \text{ bzgl. } \lambda.$$

Beobachtung 8.8 Sei V ein K -Vektorraum [auch ∞ -dim.], $F \in \text{End}_K(V)$.

- (a) $\text{Eig}(F; \lambda)$ ist ein Untervektorraum.
- (b) λ ist Eigenwert von $F \Leftrightarrow \text{Eig}(F; \lambda) \neq \{0\}$.
- (c) $\text{Eig}(F; \lambda) \setminus \{0\}$ ist die Menge der Eigenvektoren von F zu λ .
- (d) $\text{Eig}(F; \lambda) = \text{Ker}(\lambda \text{id}_V - F)$.
- (e) Für $\lambda_1, \lambda_2 \in K$ mit $\lambda_1 \neq \lambda_2$ ist $\text{Eig}(F; \lambda_1) \cap \text{Eig}(F; \lambda_2) = \{0\}$.

Beweis: s. Vorlesung \square

Definition 8.9 Sei $F \in \text{End}_K(V)$, dann heißt $\dim \text{Eig}(F; \lambda)$ die geometrische Vielfachheit von $\lambda \in K$ und $\text{sp } F := \{\lambda \in K : \text{Ker}(\lambda \text{id}_V - F) \neq \{0\}\}$ das Spektrum von } F.

8.2 Das charakteristische Polynom

In diesem Abschnitt wird nun immer $\dim(V) < \infty$ vorausgesetzt. Damit lassen sich Eigenräume mit $\text{Ker}(\lambda \text{id}_V - F) \neq \{0\}$ gerade durch $\det(\lambda \text{id}_V - F) = 0$ charakterisieren.

Beobachtung 8.10 Für $F \in \text{End}_K(V)$ ist $\lambda \in K$ genau dann ein Eigenwert von F , wenn $\det(\lambda \text{id}_V - F) = 0$.

Beweis: s. Vorlesung □

Wie lässt sich $\det(\lambda \text{id}_V - F)$ in Abhängigkeit von λ beschreiben?

Wie bei Einführung der Determinante empfiehlt es sich, das zuerst in der Matrixwelt zu betrachten, also $A \in K^{n \times n}$ für F und I_n für id_V zu schreiben. Ersetzen von $\lambda \in K$ durch eine Unbestimmte/Variable ξ führt zu

$$P_A(\xi) := \det(\xi I_n - A) = \begin{vmatrix} \xi - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \xi - a_{22} & & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & \xi - a_{nn} \end{vmatrix},$$

also einer Determinante einer Matrix $B = \xi I_n - A \in K[\xi]^{n \times n}$ mit Einträgen aus dem Polynomring $K[\xi]$. Die Determinante ist nun selbst ein Element aus $K[\xi]$, also ein Polynom in ξ , und hat hier gemäß (LEIBNIZ-Formel) $\det B = \sum_{\sigma \in S_n} \text{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)}$ die spezielle Form

$$\begin{aligned} P_A(\xi) &= \underbrace{\prod_{i=1}^n (\xi - a_{ii})}_{\sigma=\text{id}} + \underbrace{\sum_{\sigma \in S_n \setminus \{\text{id}\}} \text{sgn}(\sigma) (\xi I - A)_{1,\sigma(1)} \cdots (\xi I - A)_{n,\sigma(n)}}_{=: Q(\xi)} \\ &= \xi^n + \alpha_{n-1} \xi^{n-1} + \cdots + \alpha_1 \xi + \alpha_0. \end{aligned}$$

Weil jede Permutation $\sigma \in S_n \setminus \{\text{id}\}$ höchstens $n-2$ Fixpunkte (die Fixpunkte $i = \sigma(i)$ führen zu Diagonalelementen) enthalten kann, ist $\deg Q(\xi) \leq n-2$. Daher sind die Werte folgender Koeffizienten sehr gut beschreibbar:

$$\begin{aligned} \alpha_n &= 1, \\ \alpha_{n-1} &= (-1) \cdot (a_{11} + \cdots + a_{nn}), \\ \alpha_0 &= (-1)^n \det A. \end{aligned}$$

Man nennt $P_A(\xi) = \det(\xi I - A) \in K[\xi]$ das charakteristische Polynom von $A \in K^{n \times n}$ und $(a_{11} + \cdots + a_{nn})$ die Spur von A (= die Summe der Diagonalelemente).

Wie der Wert der Determinante einer Abbildung F , so ist auch das charakteristische Polynom unabhängig von der Wahl der Basis.

Lemma 8.11 *Ähnliche Matrizen haben das gleiche charakteristische Polynom (also auch die gleiche Spur).*

Beweis: s. Vorlesung □

$\det(\lambda \text{id}_V - F)$ weist jedem Wert $\lambda \in K$ für eine beliebige Basis \mathcal{A} den Wert $\det M_{\mathcal{A}}(\lambda \text{id}_V - F)$ zu, siehe D 6.11 und die Überlegung zur Unabhängigkeit von der Wahl von \mathcal{A} .

Ist $A = M_{\mathcal{A}}(F)$, dann ist $M_{\mathcal{A}}(\lambda \text{id}_V - F) = \lambda M_{\mathcal{A}}(\text{id}_V) - M_{\mathcal{A}}(F) = \lambda I - A$, also ist

$$\det(\cdot \text{id}_V - F): K \rightarrow K$$

$$\lambda \mapsto \det(\lambda I - A) = \widehat{P}_A(\lambda)$$

die gleiche Funktion wie die Abbildungsvariante $\widehat{P}_A(\lambda)$ des charakteristischen Polynoms. Wegen L 8.11 heißt daher

$$P_F(\xi) := P_A(\xi) \quad \underline{\text{das charakteristische Polynom von } F \text{ (} \in K[\xi]\text{)}}.$$

Damit ist folgendes gezeigt:

Satz 8.12 *Sei $F \in \text{End}_K(V)$ mit K -Vektorraum V und $n = \dim V < \infty$.*

Das charakteristische Polynom $P_F(\xi) \in K[\xi]$ hat folgende Eigenschaften:

- (a) $\deg P_F = n$.
- (b) P_F beschreibt die Abbildung $\widehat{P}_F: K \rightarrow K, \lambda \mapsto \det(\lambda \text{id}_V - F)$.
- (c) Die Nullstellen von P_F sind die Eigenwerte von F .
- (d) Ist $A = M_{\mathcal{A}}(F)$ für eine Basis \mathcal{A} , so ist $P_F(\xi) = P_A(\xi) = \det(\xi I_n - A)$.

Eine Möglichkeit Eigenwerte aufzuspüren ist daher, nach Nullstellen von $P_F(\xi)$ zu suchen [das ist i. A. nicht einfach].

Ist ein Eigenwert gefunden, lässt sich der Eigenraum leicht bestimmen.

Beobachtung 8.13 *Ist $A \in K^{n \times n} [\simeq \text{End}_K(K^n)]$, so ist*

$$\text{Eig}(A; \lambda) = \text{Lös}(\lambda I - A, 0) = \{x \in K^n : (\lambda I - A)x = 0\}$$

der Lösungsraum des homogenen linearen Gleichungssystems zu $\lambda I - A$.

Beispiele:

$$1. \ A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \ P_A(\xi) = (\xi - \cos \alpha)^2 + (\sin \alpha)^2 = \xi^2 - \xi \underbrace{2 \cos \alpha}_{=\text{Spur}} + 1,$$

$${}_1\xi_2 = \cos \alpha \pm \sqrt{(\cos \alpha)^2 - 1}.$$

Für $K = \mathbb{R}$ gibt es nur für $\alpha \in \{0, \pi\}$ Nullstellen, für diese α hat $A \in \mathbb{R}^{2 \times 2}$ Eigenwerte.

Für $K = \mathbb{C}$ hat A für jedes α zwei (nur für $\alpha \in \{0, \pi\}$ identische) Eigenwerte und ist diagonalisierbar.

2. $A = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix},$

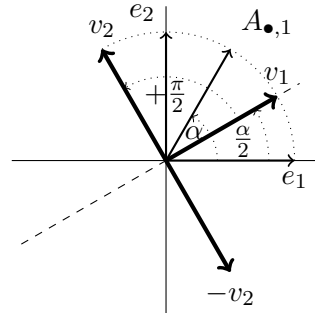
$P_A(\xi) = (\xi - \cos \alpha)(\xi + \cos \alpha) - (\sin \alpha)^2 = \xi^2 - 1 = (\xi + 1)(\xi - 1),$

A hat die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$.

Eigenvektoren sind

für $\lambda_1 = 1$: $v_1 = \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix},$

für $\lambda_2 = -1$: $v_2 = \begin{pmatrix} \cos \frac{\alpha+\pi}{2} \\ \sin \frac{\alpha+\pi}{2} \end{pmatrix}.$



Im \mathbb{R}^2 entspricht das der Spiegelung entlang der Achse $\{t \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix} : t \in \mathbb{R}\}.$

Wie in B 8.4 führt Basis $\mathcal{A} = (v_1, v_2)$ zu $M_{\mathcal{A}}(A) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$ mit Transformationsmatrix $S^{-1} := T_{(e_i)}^{\mathcal{A}} = [v_1, v_2]$ gilt $SAS^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$

3. $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, P_A(\xi) = (\xi - 2)^2 - 1 = \xi^2 - 4\xi + 3 = (\xi - 1)(\xi - 3),$

$\text{Eig}(A; 1) = \text{Ker} \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 + x_2 = 0 \} = \{ \alpha \underbrace{\begin{pmatrix} 1 \\ -1 \end{pmatrix}}_{=:v_1} : \alpha \in \mathbb{R} \}$

$\text{Eig}(A; 3) = \text{Ker} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 - x_2 = 0 \} = \{ \alpha \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{=:v_2} : \alpha \in \mathbb{R} \}$

Für Basis $\mathcal{A} = (v_1, v_2),$ Transformationsmatrix $S^{-1} := T_{(e_i)}^{\mathcal{A}} = [v_1, v_2]$ ist

$$S^{-1} = \left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right], \quad S = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

denn $S \cdot S^{-1} = \left[S \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad S \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] = \frac{1}{2} \left[\begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right] = I_2,$ und

$$SAS^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \left[1 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] = \frac{1}{2} \left[1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad 3 \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right] = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

4. $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, P_A(\xi) = (\xi - 2)^2,$

$\text{Eig}(A; 2) = \text{Ker} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_2 = 0 \} = \{ \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \}$ (1-dim!).

Die geometrische Vielfachheit ist nur 1 und echt kleiner als die algebraische Vielfachheit $\mu(P_A; 2) = 2$ der Nullstelle des charakteristischen Polynoms. Nach B 8.4 kann dieses A nicht diagonalisierbar sein.

♡

Es kann zwar, wie im letzten Beispiel, die algebraische Vielfachheit eines Eigenwerts größer als die geometrische Vielfachheit sein, aber nicht umgekehrt.

Lemma 8.14 *Ist λ Eigenwert von $F,$ so gilt $1 \leq \dim \text{Eig}(F; \lambda) \leq \mu(P_F; \lambda).$*

Beweis: s. Vorlesung □

Der Zusammenhang zwischen P_F und der Diagonalisierbarkeit von F stellt sich damit wie folgt dar.

Satz 8.15 *Sei $F \in \text{End}_K(V)$ für einen K -Vektorraum mit $\dim V = n < \infty$, dann sind äquivalent:*

1. F ist diagonalisierbar.
2. (a) Das charakteristische Polynom P_F zerfällt in Linearfaktoren und
(b) $\dim \text{Eig}(F; \lambda) = \mu(P_F; \lambda)$ für alle Eigenwerte λ von F .
3. Für die paarweise verschiedenen Eigenwerte $\lambda_1, \dots, \lambda_k$ von F gilt
 $V = \bigoplus_{i=1}^k \text{Eig}(F; \lambda_i)$.

Beweis: s. Vorlesung □

Falls nun $\dim \text{Eig}(F; \lambda) < \mu(P_F; \lambda)$ für ein $\lambda \in K$, wie sind dann invariante Unterräume mit dem charakteristischen Polynom verbunden? Die Klärung dieser Frage ist relativ aufwendig aber essentiell und folgt in den nächsten Abschnitten.

8.3 Das Minimalpolynom

Betrachtet man für einen Endomorphismus $F \in \text{End}_K(V)$ und einen Vektor $v \in V$ den „Weg“ $v, Fv, F^2v = (F \circ F)v, F^3v, \dots$, den v durch wiederholtes Anwenden von F in V zurücklegt, dann gibt es in V transiente Bereiche, die einen hineinfallenden Anteil von v innerhalb des Bereichs transformieren und zum Teil weiter- aber nie mehr zurückreichen, und so etwas wie kleinste F -invariante Untervektorräume, aus denen ein dort ankommender Anteil von v nicht mehr herauskommt, wenn er denn einmal hineingeraten ist. Die Geraden in Eigenräumen gehören zu diesen kleinsten F -invarianten Untervektorräumen, es kann aber auch höherdimensionale wie etwa bei Drehungen geben. In endlichdimensionalen Vektorräumen landet immer ein Anteil von v nach endlich vielen Anwendungen von F in so einem kleinsten invarianten Untervektorraum, und endliche Potenzen von F reichen aus, um in diesen Untervektorräumen zu enden. Der Weg von v durch V lässt sich abkürzen und besser studieren, wenn man statt der Abbildung F gleich aus Potenzen von F zusammengesetzte Abbildungen betrachtet, also Abbildungen der Form $(a_0F^0 + a_1F^1 + \dots + a_mF^m)(v)$ – diese Abbildungen sind gerade Polynome von F . In der Identifikation der invarianten Untervektorräume spielt das charakteristische Polynom P_F wieder eine zentrale Rolle. Die Information über die kleinsten invarianten Untervektorräume darin enthält das sogenannte Minimalpolynom, das hier betrachtet wird.

(b) Ist U F -invariant, dann ist U auch $f(F)$ -invariant.

Beweis: s. Vorlesung □

Entscheidend ist folgendes Lemma, nach dem teilerfremde Faktoren eines Polynoms f zu einer Zerlegung von $\text{Ker } f(F)$ in eine direkte Summe F -invarianter Untervektorräume führen.

Lemma 8.18 Sei $F \in \text{End}_K(V)$ und $f = \prod_{i=1}^k f_i \in K[\xi]$ mit paarweise teilerfremden Faktoren f_i , dann ist $\text{Ker } f(F) = \bigoplus_{i=1}^k \text{Ker } f_i(F)$ eine Zerlegung von $\text{Ker } f(F)$ in F -invariante Untervektorräume.

Beweis: s. Vorlesung □

Gibt es ein $f \in K[\xi] \setminus \{0\}$ mit $f(F) = 0$ die Nullabbildung, dann ist $V = \text{Ker } f(F)$ und L 8.18 liefert zusammen mit der Primfaktorzerlegung von f (s. S 7.8) eine Zerlegung von V in F -invariante Unterräume.

Für $\dim V = n < \infty$ muss es so ein f geben, denn $\dim K[F] \leq \dim \text{End}_K(V) = \dim K^{n \times n} = n^2$, daher sind F^0, F^1, \dots, F^{n^2} sicher linear abhängig, d. h., es gibt $a_0, \dots, a_{n^2} \in K$ mit

$$a_0 F^0 + \dots + a_{n^2} F^{n^2} = 0 \quad \text{und nicht alle } a_i = 0.$$

Also gibt es ein Polynom vom Grad höchstens n^2 mit $f(F) = 0$. Dass es sogar immer so ein Polynom vom Grad höchstens n gibt, zeigt folgender Satz.

Satz 8.19 (Cayley-Hamilton) Sei $F \in \text{End}_K(V)$ mit $\dim V = n < \infty$, dann ist $P_F(F) = 0$, d. h., F eingesetzt in sein charakteristisches Polynom ist die Nullabbildung.

Beweis: s. Vorlesung □

Im endlichdimensionalen muss es ein normiertes Polynom f kleinsten Grades geben, das $f(F) = 0$ erfüllt. [Die Forderung normiert stellt sicher, dass es nicht das Nullpolynom ist.]

Satz 8.20 Ist $F \in \text{End}_K(V)$ mit $\dim V = n < \infty$, dann gibt es ein eindeutiges normiertes Polynom $m_F \in K[\xi]$ kleinsten Grades mit $m_F(F) = 0$. Für dieses Minimalpolynom von F gilt:

- (a) $f \in K[x] \wedge f(F) = 0 \Rightarrow m_F | f$, insbesondere gilt $m_F | P_F$,
- (b) $\deg(m_F) \leq n$,
- (c) $(F^0, \dots, F^{\deg m_F - 1})$ ist eine Basis von $K[F]$, $\dim K[F] = \deg m_F$,
- (d) $m_F = 1 \Leftrightarrow F^0 = \text{id}_V = 0 \Leftrightarrow V = \{0\}$.

Beweis: s. Vorlesung □

Das Minimalpolynom enthält immer noch die gesamte Information über die Eigenwerte von F .

Beobachtung 8.21 $\lambda \in K$ ist Nullstelle von $m_F \Leftrightarrow \lambda$ ist Eigenwert von F .

Beweis: s. Vorlesung □

Ist U ein F -invarianter Untervektorraum von V , so nennt man das Minimalpolynom von $F|_U$ das Minimalpolynom von F bezüglich U , in Zeichen $m_{F|_U}$. Folgende Rechenregel wird später bei der Bestimmung des Minimalpolynoms sehr hilfreich sein.

Lemma 8.22 Sind U_1, U_2 F -invariant, dann ist

$$m_{F|_{U_1+U_2}} = \text{kgV}(m_{F|_{U_1}}, m_{F|_{U_2}}).$$

Beweis: s. Vorlesung □

8.4 F -zyklische Untervektorräume

Den Schlüssel zur Bestimmung des Minimalpolynoms, und damit zur Bestimmung der Zerlegung von V in reduzierende F -invariante Untervektorräume bilden die Untervektorräume, die für $u \in V$ durch die Vektorenfamilien $(F^i u)_{i \in \mathbb{N}_0}$ (der angesprochene Weg von u durch V bei wiederholter Anwendung von F) aufgespannt werden.

Für $u \in V$ heißt

$$[u]_F = \text{span}(F^i u)_{i \in \mathbb{N}_0} = \text{span}(F^0 u, F^1 u, F^2 u, \dots)$$

der von u erzeugte F -zyklische Untervektorraum (von V).

$[u]_F$ ist F -invariant, da für $v = \sum_{i=0}^k \alpha_i F^i u$ auch $Fv = \sum_{i=0}^k \alpha_i F^{i+1} u \in [u]_F$ gilt. Zu $[u]_F$ lässt sich mit folgender Beobachtung sowohl ein passendes Minimalpolynom als auch eine Basis algorithmisch berechnen.

Beobachtung 8.23 Sei $F \in \text{End}_K(V)$, $\dim V = n < \infty$ und $u \in V$.

- (a) Das gradkleinste normierte Polynom $m_{F,u}$, das $m_{F,u}(F)u = 0$ erfüllt, ist $m_{F,u} = m_{F|_{[u]_F}}$,
- (b) $m_{F,u} = 1 \Leftrightarrow u = 0$,
- (c) $f \in K[\xi] \wedge f(F)u = 0 \Rightarrow m_{F,u} | f$,
- (d) für $n_u = \deg m_{F,u}$ ist $(F^0 u, \dots, F^{n_u-1} u)$ eine Basis von $[u]_F$.

Beweis: s. Vorlesung □

Für ein $0 \neq u \in V$ lässt sich das Minimalpolynom $m_{F,u}$ zu $[u]_F$ bestimmen, indem man schrittweise für $j = 1, 2, \dots$ (bis höchstens n) testet, ob $F^j u \in \text{span}\{u, Fu, \dots, F^{j-1} u\}$ gilt. In Koordinatendarstellung zu einer Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V versucht man mit dem Gaußalgorithmus

$[\Phi_{\mathcal{B}}^{-1}(u), \Phi_{\mathcal{B}}^{-1}(Fu), \dots, \Phi_{\mathcal{B}}^{-1}(F^{j-1}u)]a = \Phi_{\mathcal{B}}^{-1}(F^j u)$ für $a \in K^j$ zu lösen. Gelingt das, ist $m_{F,u} = \xi^j - \sum_{i=0}^{j-1} a_{i+1} \xi^i$ und $n_u = j$, sonst ist $(u, \dots, F^j u)$ noch linear unabhängig und j muss weiter erhöht werden.

Ist zufällig $n_u = n$, hat man gleichzeitig auch das Minimalpolynom m_F bestimmt. Im Fall von $n_u < n$, gibt es ein $u' \in V \setminus [u]_F$, für das sich in gleicher Weise $m_{F,u'}$ berechnen lässt, und L 8.22 besagt, wie daraus $m_{F|[u]_F + [u']_F}$ zu gewinnen ist. Das setzt man solange fort, bis die Summe der zyklischen Untervektorräume ganz V aufspannt, dann ist auch m_F gefunden. Mit Hilfe von L 8.22 lässt sich das Minimalpolynom m_F somit konzeptionell wie folgt bestimmen:

0. Sei eine Basis $\mathcal{B} = \{u_1, \dots, u_n\}$ von V gegeben,
 $k \leftarrow 0$, $U^{(0)} = \{0\}$, $m^{(0)} = 1 \in K[\xi]$.
1. Wähle $i_k \in \{1, \dots, n\}$ mit $u_{i_k} \in V \setminus U^{(k)}$.
2. Finde das kleinste n_k mit $F^{n_k} u_{i_k} = \sum_{j=0}^{n_k-1} a_j^{(k)} F^j u_{i_k}$, dann ist $m_{F, u_{i_k}} = \xi^{n_k} - \sum_{j=0}^{n_k-1} a_j^{(k)} \xi^j$.
3. Setze $m^{(k+1)} \leftarrow \text{kgV}(m^{(k)}, m_{F, u_{i_k}}) = \frac{m^{(k)} \cdot m_{F, u_{i_k}}}{\text{ggT}(m^{(k)}, m_{F, u_{i_k}})}$,
 $U^{(k+1)} \leftarrow U^{(k)} + [u_{i_k}]_F$.
4. Setze $k \leftarrow k + 1$. Ist $U^{(k)} = V$, STOP, sonst gehe zu 1..

Die notwendigen Algorithmen zur Durchführung der einzelnen Schritte sind bekannt, aber auch recht aufwendig. Die Schritte 1., 2., 3. erfordern jeweils die Durchführung des Gaußalgorithmus, Schritt 3. zusätzlich den euklidischen Algorithmus.

Der Algorithmus liefert nur die Darstellung $m_F = \sum_{i=0}^k a_i \xi^i$, aus der i. A. die benötigte faktorisierte Darstellung keineswegs einfach zu bestimmen ist. Für theoretische Zwecke reicht es aber zu wissen, dass es eine Zerlegung in Primfaktoren gibt.

Satz 8.24 (1. Zerlegungssatz) Sei $F \in \text{End}_K(V)$ mit $\dim V = n < \infty$. Ist $m_F = \prod_{i=1}^k p_i^{r_i}$ die Primpotenzdarstellungen des Minimalpolynoms von F , dann gilt $V = \bigoplus_{i=1}^k \text{Ker } p_i^{r_i}(F)$ und für $U_i = \text{Ker } p_i^{r_i}(F)$ ist $U_i \neq \{0\}$ und $m_{F|_{U_i}} = p_i^{r_i}$.

Beweis: s. Vorlesung □

Um die Struktur F -invarianter Untervektorräume zu verstehen, genügt es daher, $F \in \text{End}_K(V)$ mit Minimalpolynom $m_F = p^r$ für ein irreduzibles p und ein $r \in \mathbb{N}$ zu analysieren. Dabei ist nicht zu erwarten, dass $\deg p^r = \dim \text{Ker } p^r = \dim V$ gilt, denn haben zwei Untervektorräume $U_1 \oplus U_2$ die Minimalpolynome $m_{F|_{U_1}} = p^{r_1}$, $m_{F|_{U_2}} = p^{r_2}$ mit $r_1 \leq r_2$, dann ist nach L 8.22 $m_{F|_{U_1+U_2}} = p^{r_2}$. Die Dimension des Gesamtraums ist aus m_F nicht

erkennbar. Man findet aber sogar immer einen F -zyklischen Untervektorraum $[u]_F$, für den $\dim[u]_F = \deg p^r$.

Satz 8.25 Sei \mathcal{B} eine endliche Basis von V und $F \in \text{End}_K(V)$ mit $m_F = p^r$, p irreduzibel, $r \in \mathbb{N}$, dann gibt es ein $u \in \mathcal{B}$ mit $m_{F,u} = p^r$. $[u]_F$ ist dann ein F -zyklischer Untervektorraum von V maximaler Dimension.

Beweis: s. Vorlesung □

$[u]_F$ spannt zwar i. A. noch nicht V auf, aber für $[u]_F$ lässt sich eine „schöne Basis“ konstruieren. Dabei hilft folgendes technische Lemma, das große Freiheit bei dieser Konstruktion gibt.

Lemma 8.26 Sei $F \in \text{End}_K(V)$, $u \in V$. $\deg m_{F,u} = s$ genau dann, wenn für $f_0, \dots, f_{s-1} \in K[\xi]$ mit $\deg f_j = j$ $\mathcal{B} = (f_0(F)u, \dots, f_{s-1}(F)u)$ eine Basis von $[u]_F$ ist.

Beweis: s. Vorlesung □

Satz 8.27 Sei $F \in \text{End}_K(V)$ mit $m_F = p^r$, p irreduzibel, $\deg p = q$, $r \in \mathbb{N}$. Für $u \in V$ ist $m_{F,u} = m_F$ genau dann, wenn $\mathcal{B}_u = (F^i p^j(F)u)_{\substack{i=0, \dots, q-1 \\ j=0, \dots, r-1}}$ eine Basis von $[u]_F$ ist. In diesem Fall gilt

$$\text{Rang } p^{j-1}(F|_{[u]_F}) - \text{Rang } p^j(F|_{[u]_F}) = \begin{cases} \deg p & \text{für } j = 1, \dots, r, \\ 0 & \text{für } j > r. \end{cases}$$

Es lohnt sich, vor dem recht kurzen Beweis die Struktur der Basis und der Rangaussage zu verstehen. Sei dazu $p = \xi^q + a_{q-1}\xi^{q-1} + \dots + a_0\xi^0$. Welche Gestalt hat $A = M_{\mathcal{B}_u}(F|_{[u]_F})$? Die Spalten von A sind die Koordinaten der Bilder der Basisvektoren

$$\mathcal{B}_u = (\underset{u_0}{u}, \underset{u_1}{Fu}, \dots, \underset{u_{q-1}}{F^{q-1}u}, \underset{u_q}{p(F)u}, \underset{u_{q+1}}{Fp(F)u}, \dots, \underset{u_{(r-1)q+q-1}}{F^{q-1}p^{r-1}(F)u}).$$

Basisvektor $u_0 = u$ wird mit F direkt auf den nächsten $u_1 = Fu_0$, dieser wieder auf den nächsten, und so weiter bis u_{q-1} . F bildet nun u_{q-1} auf folgende Linearkombination ab:

$$F(u_{q-1}) = F^q u = p(F)u = \sum_{i=0}^{q-1} a_i F^i u = u_q - a_0 u_0 - a_1 u_1 - \dots - a_{q-1} u_{q-1},$$

$V = \bigoplus_{i=1}^k [u^{(i)}]_F$ mit $m_{F, u^{(i)}} = p^{r_i}$ haben, ohne dass diese $[u^{(i)}]_F$ eine direkte Summe bilden müssen.

Es wird sich zum Glück herausstellen, dass jeweils nach „Abspaltung“ eines F -zyklischen Untervektorraums $[u]_F$ maximaler Dimension wie in S 8.25 sich der Restraum als direkte Summe von $[u]_F$ und einem F -invarianten Untervektorraum W mit $m_{F|_W} = p^{r'}$ mit $r' \leq r$ darstellen lässt, d. h., dieses „Abspalten“ ist auf W rekursiv fortsetzbar.

Wie lässt sich ein Unterraum $U = [u]_F$ heraustrennen?

Hier hilft der Quotientenvektorraum $V/U = \{\bar{v} := v + U : v \in V\}$ mit kanonischer Surjektion ρ .

$$\begin{array}{ccc} V & \xrightarrow{F} & V \\ \rho \downarrow & & \downarrow \rho \\ V/U & \xrightarrow{\bar{F}} & V/U \end{array} \quad \begin{array}{l} \text{Ist } U \text{ } F\text{-invariant, dann ist } F(U) = U \text{ und} \\ \text{wegen } \bar{0} = U \text{ ist } U \subseteq \text{Ker}(\rho \circ F), \text{ daher ist } \bar{F} \\ \text{gerade die Quotientenabbildung zu } \rho \circ F, \text{ also} \\ \text{nach S 4.14 eindeutig festgelegt.} \end{array}$$

Für dieses eindeutige \bar{F} kommutiert das Diagramm dann ganz allgemein, wenn für ein Polynom $f \in K[x]$ die Abbildungen F durch $f(F)$ und \bar{F} durch $f(\bar{F})$ ersetzt werden.

Satz 8.28 Sei $F \in \text{End}_K(V)$ und U ein F -invarianter Untervektorraum von V . Dann gilt für den Quotientenendomorphismus $\bar{F}: V/U \rightarrow V/U$, $\bar{v} \mapsto \overline{Fv} = \rho(Fv)$ mit kanonischer Surjektion ρ und für jedes Polynom $f \in K[x]$

$$\rho \circ f(F) = f(\bar{F}) \circ \rho.$$

Beweis: s. Vorlesung □

Man nennt \bar{F} auch den durch F induzierten Homomorphismus von V/U . Das vorige Resultat lässt schon vermuten, dass das Minimalpolynom von \bar{F} eng mit dem Minimalpolynom von F verknüpft ist.

Satz 8.29 Sei $F \in \text{End}_K(V)$ mit Minimalpolynom m_F , sei $U \subseteq V$ F -invariant mit induziertem $\bar{F} \in \text{End}_K(V/U)$, dann hat \bar{F} ein Minimalpolynom $m_{\bar{F}}$ und es gilt:

- (a) $m_{\bar{F}} | m_F$,
- (b) für $v \in V$: $m_{\bar{F}, \bar{v}} | m_{F, v}$.

Beweis: s. Vorlesung □

Rekursives Abspalten eines zyklischen Unterraums $[\bar{v}]_{\bar{F}}$ im Quotientenraum V/U birgt die Schwierigkeit, dass die Zerlegung ja eigentlich in V gesucht ist. D. h., zu $[\bar{v}]_{\bar{F}}$ möchte man nun wieder ein $v' \in v = v + U$ finden, das $m_{\bar{F}, \bar{v}} = m_{F, v'}$ erfüllt und somit über $[v']_F \subseteq V$ die richtige Entsprechung zu $[\bar{v}]_{\bar{F}}$ liefert.

Warum ist bei der Suche nach einem geeigneten Repräsentanten von \bar{v} die Bedingung $m_{\bar{F},\bar{v}} = m_{F,v'}$ wichtig und richtig? Das Ziel ist eine Zerlegung von V in eine direkte Summe reduzierender F -zyklischer Unterräume und $[u]_F$ soll bereits einer von ihnen sein. Für $s' = \deg m_{\bar{F},\bar{v}} = \deg m_{F,v'}$ ist $(v', Fv', \dots, F^{s'-1}v')$ nach B8.23(d) eine Basis von $[v']_F$ und alle ihre Elemente sind im Bild des Homomorphismus ρ linear unabhängig, also auch linear unabhängig bei beliebiger Linearkombination mit Elementen aus $U = [u]_F$. Für ein $v'' \in \bar{v}$ mit $\deg m_{F,v''} > \deg m_{\bar{F},\bar{v}}$ (kleiner geht wegen S 8.29(b) nicht) enthält die Basis von $[v'']_F$ mehr Elemente als im Bild von ρ linear unabhängig sein können, d. h., gemeinsam sind Basen von $[u]_F$ und $[v'']_F$ immer linear abhängig; mit $[v'']_F$ ist keine direkte Summe möglich, s. S 3.28. Nach B8.23(a) erkennt man ein geeignetes $v' \in \bar{v}$ daran, dass $m_{\bar{F},\bar{v}}(F)v' = 0$. Weil $m_{\bar{F},\bar{v}}(\bar{F})\bar{v} = \bar{0}$, gilt für jedes $v'' \in \bar{v}$ $m_{\bar{F},\bar{v}}(F)v'' \in \bar{0} = U$. Ein eventueller Anteil in $U \setminus \{0\}$ lässt die Basis von $[v'']_F$ in $[u]_F$ hineinragen, und diesen Anteil gilt es im Beweis des folgenden Satzes zu eliminieren.

Satz 8.30 *Sei $F \in \text{End}_K(V)$, $\dim V = n < \infty$, mit $m_F = p^r$ für ein irreduzibles $p \in K[\xi]$ und $r \in \mathbb{N}$. Sei $u \in V$ mit $m_{F,u} = m_F = p^r$ wie in S 8.25 und $U = [u]_F$. Für jedes $\bar{v} = v + U \in V/U$ gibt es ein $v' \in \bar{v}$ mit $m_{\bar{F},\bar{v}} = m_{F,v'}$.*

Beweis: s. Vorlesung □

Nun sind alle Vorbereitungen für die Zerlegung von V in eine direkte Summe F -zyklischer Untervektorräume getroffen.

Satz 8.31 (2. Zerlegungssatz) *Sei $F \in \text{End}_K(V)$, $\dim V = n < \infty$, mit $m_F = p^r$ für ein irreduzibles $p \in K[\xi]$ und ein $r \in \mathbb{N}$. Dann ist V eine direkte Summe F -zyklischer Untervektorräume, $V = \bigoplus_{i=1}^k [u_i]_F$.*

Dabei sind die Polynome $m_{F,u_i} = p^{r_i}$, $1 \leq r_i \leq r$ bis auf die Reihenfolge eindeutig bestimmt und für die Anzahlen $D(j)$, $j \in \mathbb{N}$, der Polynome mit $m_{F,u_i} = p^j$ gilt

$$D(j) + D(j-1) + \dots + D(r) = \frac{1}{\deg p} (\text{Rang } p^{j-1}(F) - \text{Rang } p^j(F)). \quad (\text{Zykr})$$

Beweis: s. Vorlesung □

Für $n = 0$ ist $V = \{0\}$, daher lässt sich mit $V = [0]_F$ für jede Dimension so eine Zerlegung angeben, aber der Fall $n = 0$ hat für den vorigen Satz keine Bedeutung.

8.5 Normalformen

Noch detailreicher als im 2. Zerlegungssatz können Endomorphismen nicht mehr in invariante Untervektorräume zerlegt werden, wie die folgenden

Überlegungen zeigen. Damit ist die Struktur so gut verstanden wie möglich. Sie ist einfach genug, um den Charakter der Abbildung in einer Normalform übersichtlich darzustellen.

Definition 8.32 Sei $F \in \text{End}_K(V)$. Ein Untervektorraum $U \subseteq V$ heißt ***F-irreduzibel***, falls U F -invariant ist und es keine (nicht-triviale) Darstellung $U = U_1 \oplus U_2$ mit U_1, U_2 F -invariant und $U_1 \neq \{0\} \neq U_2$ gibt.

Satz 8.33 Sei $F \in \text{End}_K(V)$, $\dim V = n < \infty$. U ist F -irreduzibel genau dann wenn U F -zyklisch ist und $m_{F|_U} = p^r$ für ein $r \in \mathbb{N}$ und ein irreduzibles $p \in K[\xi]$.

Insbesondere ist U genau dann F -irreduzibel, wenn U F -invariant ist und es eine Basis \mathcal{B}_U von U gibt, sodass

$$M_{\mathcal{B}_U}(F|_U) = M(p^r) := \begin{bmatrix} \boxed{M(p)} & & & & \\ & 1 & & & \\ & & \boxed{M(p)} & & \\ & & & \ddots & \\ & & & & 1 & \\ & & & & & \boxed{M(p)} \end{bmatrix} \begin{matrix} 1 \\ 2 \\ \vdots \\ r \end{matrix}$$

mit $M(p) := \begin{bmatrix} 0 & & & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & \ddots & \vdots \\ & & 0 & -a_{q-2} \\ & & 1 & -a_{q-1} \end{bmatrix}$

Begleitmatrix eines irreduziblen Polynoms

$$p = \sum_{h=0}^q a_h \xi^h \text{ mit } a_q = 1.$$

Beweis: s. Vorlesung □

Der nächste Satz fasst die Ergebnisse zusammen.

Satz 8.34 Sei $F \in \text{End}_K(V)$, $\dim V = n < \infty$ und $m_F = \prod_{i=1}^k p_i^{r_i}$ die Primpotenzdarstellung des Minimalpolynoms.

Dann gibt es $k_i \in \mathbb{N}$, $u_{ij} \in V$, $i = 1, \dots, k$, $j = 1, \dots, k_i$ mit

$$V = \bigoplus_{\substack{i=1, \dots, k \\ j=1, \dots, k_i}} [u_{ij}]_F \quad \text{und} \quad n = \sum_{\substack{i=1, \dots, k \\ j=1, \dots, k_i}} \deg m_{F|_{[u_{ij}]_F}}.$$

Jedes $[u_{ij}]_F$ ist F -irreduzibel, die Polynome erfüllen $m_{F|_{[u_{ij}]_F}} = p_i^{r_{ij}}$ mit $1 \leq r_{ij} \leq r_i$ und werden ***elementare Divisoren*** genannt. Jede weitere derartige Darstellung hat bis auf Umnummerierung die gleichen elementaren Divisoren und F -irreduziblen Untervektorräume.

Beweis: s. Vorlesung □

Vorsicht! Die elementaren Divisoren beinhalten Vielfachheiten! ▽

Die Beziehung zwischen Minimalpolynom und charakteristischem Polynom lässt sich wie folgt beschreiben.

Satz 8.35 Sei $F \in \text{End}_K(V)$, $\dim V = n < \infty$, $m_F = \prod_{i=1}^k p_i^{r_i}$ die Primpotenzdarstellung und d_1, \dots, d_s die elementaren Divisoren.

- (a) Jedes $p_i^{r_i}$ ist in den elementaren Divisoren [aber i. A. nicht umgekehrt].
- (b) $P_F = \prod_{i=1}^s d_i$.
- (c) $P_F | (m_F)^n$.

Beweis: s. Vorlesung □

Für jeden F -irreduziblen Unterraum U_j gibt es also ein $u_j \in U_j$ mit $[u_j]_F = U_j$, einen eigenen elementaren Divisor $d_j = m_{F|_{U_j}} = p_i^{r'_j}$ sowie eine Basis \mathcal{B}_j entsprechend S 8.27, sodass insgesamt eine Blockmatrix mit Blöcken $M(d_j) = M_{\mathcal{B}_j}(F|_{U_j}) = M(p_i^{r'_j})$ wie in S 8.33 entsteht.

Die Blockmatrix

$$\begin{bmatrix} \boxed{M(d_1)} & & & \\ & \boxed{M(d_2)} & & \\ & & \ddots & \\ & & & \boxed{M(d_s)} \end{bmatrix}$$

heißt allgemeine Normalmatrix zu F .

Satz 8.36 Zwei Matrizen sind genau dann ähnlich, wenn sie (bis auf die Reihenfolge) dieselben elementaren Divisoren, bzw. ähnlich zur selben Normalmatrix sind.

Beweis: s. Vorlesung □

Jede Matrix $A = \begin{bmatrix} \lambda & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{bmatrix} \in K^{n \times n}$ ist ähnlich zu $A^\top = \begin{bmatrix} \lambda & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{bmatrix}$ über die Permutationsmatrix $P = \begin{bmatrix} 0 & \cdots & 1 \\ 1 & & 0 \end{bmatrix}$ ($P^{-1} = P$)

$$\begin{bmatrix} \lambda & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 1 \\ 1 & & 0 \end{bmatrix} \begin{bmatrix} \lambda & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{bmatrix} \begin{bmatrix} 0 & \cdots & 1 \\ 1 & & 0 \end{bmatrix},$$

also haben beide dieselben elementaren Divisoren.

Für $d = (\xi - \lambda)^r$ heißt

$$M(d)^T = \begin{bmatrix} \lambda & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{bmatrix} \in K^{r \times r}$$

Jordanblock der Länge r zu Eigenwert λ .

Eine Matrix ist in Jordanscher Normalform, wenn sie eine Blockdiagonalmatrix bestehend aus Jordanblöcken ist. Dabei gruppiert man die Blöcke so, dass Blöcke mit den gleichen Eigenwerten nebeneinander und monoton fallend in der Länge angeordnet sind.

Satz 8.37 (Hauptsatz über die Jordansche Normalform)

- (a) Zwei Jordansche Normalformen sind genau dann ähnlich, wenn sie bis auf die Reihenfolge die gleichen Jordanblöcke haben.
- (b) Eine $n \times n$ Matrix ist genau dann ähnlich zu einer Jordanschen Normalform, wenn ihr charakteristisches Polynom (oder das Minimalpolynom) in Linearfaktoren zerfällt.

Beweis: s. Vorlesung □

Korollar 8.38 Jede Matrix $A \in \mathbb{C}^{n \times n}$ ist ähnlich zu einer Jordanschen Normalform und die Zahl der Jordanblöcke zu einem Eigenwert λ ist die geometrische Vielfachheit des Eigenwerts.

Beweis: s. Vorlesung □

Ist λ ein Eigenwert von F , $r = \mu(m_F; \lambda)$ seine algebraische Vielfachheit im Minimalpolynom, und $A = M_{\mathcal{B}}(F)$ in Jordanscher Normalform, dann ist jeder Jordanblock der Länge r' ($r' \leq r$) zu λ von der Form

$$J = \lambda I + N \text{ mit } N = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix} \in K^{r' \times r'} \text{ und } N^{r'} = 0.$$

Insbesondere ist $(\lambda I - J)^r = 0$ für jeden Jordanblock zu λ . Man nennt

$$\text{Hau}(F; \lambda) := \text{Ker}(\lambda \text{id}_V - F)^r$$

den Hauptraum (oder verallgemeinerten Eigenraum) zu Eigenwert λ .

Ü: Zeige: Hat F k elementare Divisoren $d_i = m_{F|_{[u_i]_F}} = (\xi - \lambda)^{r_i}$, $i = 1, \dots, k$, dann gilt

- (a) $v \in \text{Hau}(F; \lambda) \Leftrightarrow v \in \bigoplus_{i=1}^k [u_i]_F$,
- (b) $\dim \text{Hau}(F; \lambda) = \sum_{i=1}^k \deg d_i = \mu(P_F; \lambda)$.

⊠

Eine Matrix $A \in K^{n \times n}$ mit $A^r = 0$ für ein $r \in \mathbb{N}$ heißt nilpotent, und das kleinste r mit $A^r = 0$ ist der Nilpotenzindex von A . Ebenso heißt ein Endomorphismus F nilpotent mit Nilpotenzindex r , falls $F^r = 0$ aber $F^{r-1} \neq 0$.

Ü: Zeige: Eine quadratische Matrix / ein Endomorphismus F ist genau dann nilpotent zum Nilpotenzindex r , wenn $m_F = \xi^r$. Dies ist genau dann der Fall, falls es eine Darstellung in Jordanscher Normalform gibt mit allen Jordanblöcken der Länge $r' \leq r$ zum Eigenwert 0, wobei mindestens einer die Länge r hat. ⊠

Für $A \in \mathbb{R}^{n \times n}$ muss P_A nicht mehr in Linearfaktoren zerfallen, aber nach S 7.17 sind alle irreduziblen Polynome höchstens vom Grad 2.

Satz 8.39 Für ein $F \in \text{End}(\mathbb{R}^n)$ mit elementaren Divisoren d_i , $i = 1, \dots, s$, gibt es eine Basis \mathcal{B} , so dass $M(d_i)$ entweder ein Jordanblock zu einem Eigenwert λ oder ein Block der Form

$$M(d_i) = \begin{bmatrix} \boxed{B} & & & \\ & 1 & & \\ & & \boxed{B} & \\ & & & 1 & \dots \\ & & & & & \dots \\ & & & & & & 1 & \\ & & & & & & & \boxed{B} \end{bmatrix} \quad \text{mit } B = \begin{bmatrix} 0 & -c_i \\ 1 & -b_i \end{bmatrix}, \quad b_i^2 - 4c_i < 0 \text{ ist.}$$

Beweis: s. Vorlesung

□

Kapitel 9

Räume mit Skalarprodukt

Hier ist der Körper stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, andere Körper werden nicht betrachtet.

9.1 Euklidische und Unitäre Räume

Um geometrische Eigenschaften untersuchen zu können, werden „Längen“ von und „Winkel“ zwischen Vektoren benötigt. Das Skalarprodukt bildet die mathematische Grundlage für beide Begriffe in allgemeinen Vektorräumen.

Definition 9.1 Sei V ein \mathbb{K} -Vektorraum ($\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$). Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ heißt Skalarprodukt oder inneres Produkt, falls für alle $u, v \in V, \alpha \in \mathbb{K}$

- (a) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
- (b) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$,
- (c) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ [für $\mathbb{R} \Rightarrow \langle u, v \rangle = \langle v, u \rangle$],
- (d) $\langle v, v \rangle \in [0, \infty]$ ($\in \mathbb{R}_{\geq 0}$), es ist positiv semidefinit,
- (e) $\langle v, v \rangle = 0 \Leftrightarrow v = 0$, es ist mit (d) positiv definit.

In der Literatur werden oft auch andere Klammertypen verwendet, etwa runde Klammern wie in (u, v) .

Für $\mathbb{K} = \mathbb{R}$ sind die Skalarprodukte gerade die *symmetrischen, positiv definiten Bilinearformen*:

Symmetrie: $\langle u, v \rangle \stackrel{(c)}{=} \langle v, u \rangle$ ($\in \mathbb{R}$).

(L), 1. Argument: $\langle \alpha u + \beta v, w \rangle \stackrel{(a)}{=} \langle \alpha u, w \rangle + \langle \beta v, w \rangle \stackrel{(b)}{=} \alpha \langle u, w \rangle + \beta \langle v, w \rangle$.

(L), 2. Argument: $\langle w, \alpha u + \beta v \rangle \stackrel{(c)}{=} \langle \alpha u + \beta v, w \rangle \stackrel{(a)(b)(c)}{=} \alpha \langle w, u \rangle + \beta \langle w, v \rangle$.

Für $\mathbb{K} = \mathbb{C}$ ist ein Skalarprodukt im 2. Argument nur semilinear:

$$\langle w, \alpha u + \beta v \rangle \stackrel{(c)}{=} \overline{\langle \alpha u + \beta v, w \rangle} \stackrel{(a)(b)}{=} \bar{\alpha} \overline{\langle u, w \rangle} + \bar{\beta} \overline{\langle v, w \rangle} \stackrel{(c)}{=} \bar{\alpha} \langle w, u \rangle + \bar{\beta} \langle w, v \rangle.$$

Wegen der Linearität im 1. und der Semilinearität im 2. Argument nennt man $\langle \cdot, \cdot \rangle$ für \mathbb{C} eine Sesquilinearform ($1\frac{1}{2}$ -fach linear).

Definition 9.2 Ein \mathbb{K} -Vektorraum mit Skalarprodukt heißt für $\mathbb{K} = \mathbb{R}$ Euklidischer (Vektor-)Raum und für $\mathbb{K} = \mathbb{C}$ unitärer (Vektor-)Raum.

Beispiele:

1. $V = \mathbb{R}^n: \langle u, v \rangle = \sum_{i=1}^n u_i v_i = v^\top u$, kanonisches Skalarprodukt im \mathbb{R}^n

(a),(b),(c): folgt aus den Summeneigenschaften,
 (d),(e): $\langle v, v \rangle = \sum_{i=1}^n v_i^2 \geq 0$ und $= 0 \Leftrightarrow v = 0$.

2. $V = \mathbb{C}^n: \langle u, v \rangle = \sum_{i=1}^n u_i \bar{v}_i$, kanonisches Skalarprodukt in \mathbb{C}^n

(a),(b): sind klar, (c): $\overline{\langle v, u \rangle} = \overline{\sum_{i=1}^n v_i \bar{u}_i} = \sum_{i=1}^n \bar{v}_i u_i = \langle u, v \rangle$,
 (d),(e): $\langle v, v \rangle = \sum_{i=1}^n v_i \bar{v}_i = \sum_{i=1}^n |v_i|^2 \geq 0$ und $= 0 \Leftrightarrow v = 0$.

3. $V = \mathbb{R}^n, A = \text{Diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}, \lambda_i > 0$

Ü: $\langle v, w \rangle_A = w^\top A v = \sum_{i=1}^n \lambda_i v_i w_i$ ist Skalarprodukt. Wie geht das für \mathbb{C}^n ? x

4. $V = C_{\mathbb{K}}[a, b] = \{f: [a, b] \rightarrow \mathbb{K}, f \text{ stetig}\}$ mit punktweiser Addition und skalarer Multiplikation ($\dim V = \infty$):

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx \quad \text{ist Skalarprodukt.}$$

(a),(b),(c): aus den Rechenregeln des Integrals (=Summeneigenschaften),
 (d),(e): $\int_a^b f(x) \overline{f(x)} dx = \int_a^b |f(x)|^2 dx \geq 0$ und $= 0 \Leftrightarrow f(x) = 0 \forall x \in [a, b]$.

♡

Eines der hilfreichsten Resultate, keineswegs nur in geometrischen Anwendungen sondern in vielen Abschätzungen, ist die Cauchy-Schwarzsche Ungleichung.

Satz 9.3 (Cauchy-Schwarzsche Ungleichung)

Ist V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, gilt

$$|\langle u, v \rangle|^2 \leq \langle u, u \rangle \langle v, v \rangle \quad \text{für alle } u, v \in V \tag{CS}$$

mit Gleichheit genau dann, wenn u und v linear abhängig sind.

Beweis: s. Vorlesung □

Die (CS)-Ungleichung ermöglicht zu zeigen, dass über $\|v\| := \sqrt{\langle v, v \rangle}$ jedes Skalarprodukt eine sogenannte Norm induziert.

Definition 9.4 Sei V ein \mathbb{K} -Vektorraum. Eine Abbildung $p: V \rightarrow [0, \infty)$ heißt Halbnorm, falls

- (a) $p(\alpha v) = |\alpha|p(v)$ für alle $v \in V$, $\alpha \in \mathbb{K}$,
 (b) $p(u + v) \leq p(u) + p(v)$ für alle $u, v \in V$ (die Dreiecksungleichung).

p heißt Norm, falls zusätzlich gilt

- (c) $p(v) = 0 \Rightarrow v = 0$.

Normen schreibt man meist $\|\cdot\|$ oder $\|\cdot\|_{\text{Name}}$.

Ü: Jede Norm induziert auf eine natürliche Weise eine Metrik (einen Abstand, eine Distanz) auf V über

$$d(u, v) = \|u - v\| \quad \text{für alle } u, v \in V.$$

Definition Sei S eine Menge. Eine Abbildung $d: S \times S \rightarrow [0, \infty)$ heißt Metrik, falls für alle $u, v, w \in S$

- (a) $d(u, v) \geq 0$ und $d(u, v) = 0 \Leftrightarrow u = v$,
 (b) $d(u, v) = d(v, u)$,
 (c) $d(u, v) \leq d(u, w) + d(w, v)$ (Dreiecksungleichung).

Die Dreiecksungleichung bringt zum Ausdruck, dass der Weg von u nach v über einen Zwischenpunkt w mindestens so lange ist, wie der direkte zwischen u und v . ⊗

Satz 9.5 Ist V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, so definiert

$$\|\cdot\|: V \rightarrow [0, \infty), \quad v \mapsto \langle v, v \rangle^{\frac{1}{2}} \quad \text{eine Norm auf } V.$$

Inbesondere gilt für alle $u, v, w \in V$, $\alpha \in \mathbb{K}$

- (a) $\|v\| \geq 0$ und $\|v\| = 0 \Leftrightarrow v = 0$,
 (b) $\|\alpha v\| = |\alpha| \cdot \|v\|$,
 (c) $\|u + v\| \leq \|u\| + \|v\|$ (Dreiecks-/Minkowskische Ungleichung).

Das Skalarprodukt ist aus $\|\cdot\|$ eindeutig über folgende Polarisierungsidentitäten rekonstruierbar:

$$\begin{aligned} \mathbb{K} = \mathbb{R}: \quad \langle u, v \rangle &= \frac{1}{2}(\|u + v\|^2 - \|u\|^2 - \|v\|^2), \\ \mathbb{K} = \mathbb{C}: \quad \langle u, v \rangle &= \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2). \end{aligned}$$

Beweis: s. Vorlesung □

Für $v \neq 0$ hat $\frac{v}{\|v\|}$ stets Norm 1. Ein Vektor der Norm 1 heißt normiert.

Beispiele:

1. $V = \mathbb{R}^n, \|v\| = (\sum_{i=0}^n v_i^2)^{\frac{1}{2}} = \sqrt{v_1^2 + \dots + v_n^2},$

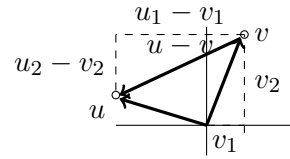
Euklidische Norm oder 2-Norm $\|\cdot\|_2,$

„Länge des Ortsvektors“ vom Ursprung zum Punkt $v \in \mathbb{R}^n, \|v\| = \|v - 0\|.$

$\|u - v\|$ ist der Euklidische Abstand zwischen Punkten u und $v.$

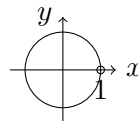
Sphäre vom Radius $r \in \mathbb{R}_{\geq 0}$ um Mittelpunkt $u \in V:$

$$S_r(u) = \{v: \|v - u\| = r\}, \quad \text{z.B. } S_1(0) = \left\{ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} : \sum v_i^2 = 1 \right\}$$



Im $\mathbb{R}^2:$ Einheitskreis

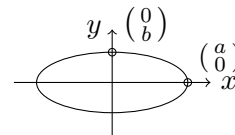
$$S_1(0) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x^2 + y^2 = 1 \right\}$$



2. $V = \mathbb{R}^2, \langle u, v \rangle_A = v^T A u, A = \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix}, \|v\|_A = \sqrt{v^T A v} = \sqrt{\frac{v_1^2}{a^2} + \frac{v_2^2}{b^2}}.$

Im \mathbb{R}^2 ist der Einheitskreis zu dieser Norm

$$S_1^A(0) = \{v: \|v\|_A = 1\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}$$



eine Ellipse mit Halbachsenlängen a und $b.$

Im \mathbb{R}^n wird $S_1^A(0)$ mit $A = \begin{bmatrix} \frac{1}{a_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{a_n^2} \end{bmatrix}$ zu einem Ellipsoid mit Halbachsenlängen $a_1, \dots, a_n.$



Cauchy-Schwarz und Winkel im \mathbb{R}^n

(CS) besagt $|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$ mit Gleichheit $\Leftrightarrow u, v$ linear abhängig.

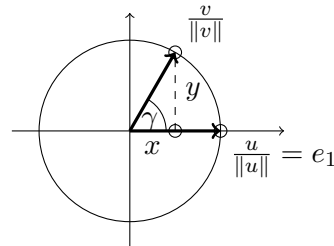
Für $u \neq 0 \neq v$ gilt somit stets $-1 \leq \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|} \leq 1.$

Daher ist
$$\frac{\langle u, v \rangle}{\|u\| \cdot \|v\|} = \left\langle \frac{u}{\|u\|}, \frac{v}{\|v\|} \right\rangle \begin{cases} = 1 & \frac{u}{\|u\|} = \frac{v}{\|v\|}, \\ = -1 & \frac{u}{\|u\|} = -\frac{v}{\|v\|}, \\ \in (-1, 1) & \text{sonst.} \end{cases}$$

Beispiel: Im \mathbb{R}^2 mit kanonischem Skalarprodukt $\langle u, v \rangle = u_1 v_1 + u_2 v_2$

sei $\frac{u}{\|u\|} = e_1, \frac{v}{\|v\|} = \begin{pmatrix} x \\ y \end{pmatrix},$ also $x^2 + y^2 = 1.$

$$\Rightarrow \left\langle \frac{u}{\|u\|}, \frac{v}{\|v\|} \right\rangle = 1 \cdot x + 0 \cdot y = \cos \gamma.$$



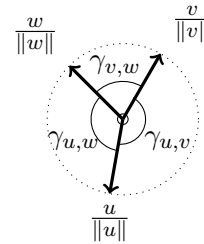
♡

Definition 9.6 Sei V ein Euklidischer Vektorraum $[\mathbb{K} = \mathbb{R}, \text{Skalarprodukt}]$. Der Winkel $\gamma_{u,v} \in [0, \pi]$ zwischen Vektoren $v, w \in V \setminus \{0\}$ ist definiert durch

$$\cos \gamma_{u,v} = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|} \quad (\text{auch } \gamma_{u,v} = \angle v, w).$$

Mit dieser Definition gilt $\langle u, v \rangle = \|u\| \cdot \|v\| \cdot \cos \gamma_{u,v}$,

bzw. $\left\langle \frac{u}{\|u\|}, \frac{v}{\|v\|} \right\rangle = \cos \gamma_{u,v}$.



Vorsicht! Die übliche geometrische Vorstellung vom Winkel stimmt für das kanonische innere Produkt $\langle u, v \rangle = v^\top u$. Für andere innere Produkte wird sie i. A. deutlich davon abweichen. ▽

9.2 Orthogonalität und Orthogonalsysteme

Im reellen entscheidet das Vorzeichen von $\langle u, v \rangle$, ob die Vektoren eher in die gleiche Richtung (Winkel $< \frac{\pi}{2}$) oder in entgegengesetzte Richtung (Winkel $> \frac{\pi}{2}$) zeigen. Ist das innere Produkt 0, der Winkel also genau $\frac{\pi}{2}$, stimmt das für das kanonische innere Produkt mit der üblichen Vorstellung von Vektoren überein, die senkrecht aufeinander stehen. In dem Fall enthält sozusagen kein Vektor einen Anteil der Richtung des anderen, das macht viele Betrachtungen und Berechnungen besonders einfach und robust.

Definition 9.7 In einem \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ heißen zwei Elemente $u, v \in V$ orthogonal (rechtwinklig), falls $\langle u, v \rangle = 0$. Man schreibt auch $u \perp v$ (insbesondere gilt $0 \perp v$ für $v \in V$).

- Eine Teilmenge $\{v_1, \dots, v_r\} \subseteq V$ heißt Orthogonalsystem (OGS), falls $\langle v_i, v_j \rangle = 0$ für $i \neq j$ (die Vektoren sind paarweise orthogonal).
- ein OGS heißt Orthonormalsystem (ONS), falls $\langle v_i, v_j \rangle = \delta_{ij}$ (die Vektoren sind auch normiert, $\|v_i\| = 1$).
- Ein OGS heißt Orthogonalbasis (OGB), falls (v_1, \dots, v_r) eine Basis von V bilden.
- Ein ONS heißt Orthonormalbasis (ONB), falls es eine OGB ist.

Vorsicht! Im Unendlichdimensionalen wird Orthonormalbasis leicht anders definiert und ist meist keine Basis im Vektorraumsinn! [ONB bezeichnet

dort ein maximales, also ein nicht mehr erweiterbares ONS. Die Darstellung eines Vektors benötigt dann meist unendlich viele Nichtnull-Koeffizienten und das führt zu den bekannten Problemen mit der Definition aufgespannter Unterräume.] ∇

Mit orthogonalen Vektoren und Orthonormalsystemen wird vieles sehr einfach.

Satz 9.8 (Pythagoras) Für $u \perp v$ gilt $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

Beweis: s. Vorlesung \square

Lemma 9.9 Die Vektoren eines ONS sind linear unabhängig.

Beweis: s. Vorlesung \square

Jedes ONS (v_1, \dots, v_r) ist daher eine Basis des vom ONS aufgespannten Untervektorraums und für jedes $v = \sum_{i=1}^r \alpha_i v_i \in \text{span}\{v_1, \dots, v_r\}$ ist $\langle v, v_j \rangle = \sum_{i=1}^r \alpha_i \langle v_i, v_j \rangle = \alpha_j$ die j -te Koordinate für diese Basis.

Lemma 9.10 (Gram-Schmidt-Orthogonalisierung) Ist $\{v_1, \dots, v_r\}$ ein ONS in V und $v \in V \setminus \text{span}\{v_1, \dots, v_r\}$, dann ist $u = v - \sum_{i=1}^r \langle v, v_i \rangle v_i \neq 0$ und die Menge $\{v_1, \dots, v_r, \frac{u}{\|u\|}\}$ wieder ein ONS mit $\text{span}\{v_1, \dots, v_r, v\} = \text{span}\{v_1, \dots, v_r, \frac{u}{\|u\|}\}$.

Beweis: s. Vorlesung \square

Korollar 9.11 Jeder endlichdimensionale \mathbb{K} -Vektorraum mit Skalarprodukt hat eine ONB.

Beweis: s. Vorlesung \square

Korollar 9.12 Jedes ONS $\{v_1, \dots, v_r\}$ eines endlichen Vektorraums kann zu einer ONB $(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$ ergänzt werden.

Beweis: s. Vorlesung \square

Beispiel: \mathbb{R}^4 : $w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, w_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

$$v_1 = \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

$$u_2 = w_2 - \langle w_2, v_1 \rangle v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot 0 \cdot v_1, \quad v_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

$$\begin{aligned} u_3 &= w_3 - \langle w_3, v_1 \rangle v_1 - \langle w_3, v_2 \rangle v_2 \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot 1 \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}}(-1) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} =: v_3. \end{aligned}$$

$$u_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{0}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} =: v_4. \quad \heartsuit$$

Korollar 9.13

Sei U ein Untervektorraum eines \mathbb{K} -Vektorraums V mit Skalarprodukt $\langle \cdot, \cdot \rangle$.

Für das orthogonale Komplement $U^\perp := \{v \in V : \forall u \in U \langle u, v \rangle = 0\}$ gilt

- (a) U^\perp ist ein Untervektorraum.
- (b) Ist U endlichdimensional, so ist
 - (b1) $V = U \oplus U^\perp$,
 - (b2) $(U^\perp)^\perp = U$,
 - (b3) für V endlichdimensional $\dim V = \dim U + \dim U^\perp$.

Beweis: s. Vorlesung □

Dass bei unendlichdimensionalem U die Aussage in (b) nicht mehr gelten muss, zeigt das folgende Beispiel.

Beispiel: $V = C^0[0, 1] = \{f: [0, 1] \rightarrow \mathbb{R} : f \text{ stetig}\}$, $\langle f, g \rangle := \int_0^1 f(t)g(t)dt$.
 Betrachte $U := \{f: [0, 1] \rightarrow \mathbb{R} : f \text{ Polynom}\} \subsetneq V$, $\dim U = \infty$.

Für dieses U ist $U^\perp = \{0\}$ (also $(U^\perp)^\perp = V$), wie nun unter Nutzung des Weierstraßschen Approximationssatzes (aus der Analysis) gezeigt wird:

$$\forall f \in V \forall \varepsilon > 0 \exists g \in U \forall t \in [0, 1] |f(t) - g(t)| < \varepsilon. \quad (\text{WAS})$$

[Jedes stetige f lässt sich auf kompakten Intervallen uniform durch Polynome approximieren.]

Es ist zu zeigen, dass für $0 \neq f \in C^0[0, 1]$ es ein $g \in U$ mit $\langle f, g \rangle \neq 0$ gibt. Setze $a := \langle f, f \rangle > 0$ und $M := \max\{|f(t)| : t \in [0, 1]\}$ [max und min wird von stetigen Funktionen auf kompakten Mengen angenommen]. Wähle $\varepsilon := \frac{a}{2M} > 0$ und g entsprechend (WAS), dann ist

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 f(t)g(t)dt = \int_0^1 f(t)[f(t) - (f(t) - g(t))]dt \\ &= \int_0^1 f(t)^2 dt - \int_0^1 f(t)[f(t) - g(t)]dt \\ &\geq \langle f, f \rangle - \int_0^1 \underbrace{|f(t)|}_{\leq M} \cdot \underbrace{|f(t) - g(t)|}_{\leq \varepsilon} dt \\ &\geq \langle f, f \rangle - M\varepsilon \int_0^1 1 dt = a - M \frac{a}{2M} = \frac{a}{2} > 0. \end{aligned}$$

♡

Wie findet man für einen gegebenen endlichdimensionalen Untervektorraum U und ein beliebiges $v \in V$ das $u \in U$ mit kleinstem Abstand $\|v - u\|$? Dieses u heißt dann die Projektion von v auf U . In der Praxis stellt U meist die mit den verfügbaren Mitteln realisierbaren Lösungen dar, und v die Lösung, die man eigentlich gerne umsetzen würde. Die Projektion ist dann bezüglich des von der Norm induzierten Abstandsmaßes die beste Approximation/Annäherung an v innerhalb von U .

Satz 9.14 Sei V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $U \subseteq V$ ein endlichdimensionaler Untervektorraum von V . Der orthogonale Projektor $P_U: V = U \oplus U^\perp \rightarrow U$, $v = u + w \mapsto u$ von V auf U ist linear und erfüllt:
 (a) $\text{Ker } P_U = U^\perp$, $\text{Im } P_U = U$ und P_U ist idempotent: $P_U \circ P_U = P_U$.
 (b) $\|v - P_U(v)\| \leq \|v - u'\|$ für alle $u' \in U$ mit Gleichheit nur für $u' = P_U(v)$.
 (c) Ist $\{u_1, \dots, u_r\}$ eine ONB von U , so ist $P_U(v) = \sum_{i=1}^r \langle v, u_i \rangle u_i$.

Beweis: s. Vorlesung □

Beispiele:

- \mathbb{K}^n , Abstand eines Punktes $v \in \mathbb{K}^n$ zu einer Hyperebene $H \subseteq \mathbb{K}^n$.
 Sei $w \in H$, dann ist $H - w$ ein $n - 1$ dimensionaler Untervektorraum und hat eine ONB $\{u_1, \dots, u_{n-1}\}$. Ergänze diese zu einer ONB $\{u_1, \dots, u_{n-1}, u_n\}$ von \mathbb{K}^n .
Ü: Setze $P_H(v) := P_{H-w}(v - w) + w$. Es gilt:
 (a) $P_H(v) = v + \langle v - v, u_n \rangle u_n$.
 (b) $\|v - P_H(v)\| \leq \|v - u'\|$ für alle $u' \in H$.
 (c) $\|v - P_H(v)\| = |\langle v - w, u_n \rangle|$.
 (d) $H = \{u \in \mathbb{K}^n: \langle u, u_n \rangle = \langle w, u_n \rangle\}$.

⊗

u_n ist (wegen \pm nicht „der“ sondern „ein“) normierter Normalvektor der Hyperebene und die Darstellung von H über u_n ist im reellen als Hessesche Normalform bekannt.

- Im \mathbb{R}^n : Sei $u_n = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, $\|u_n\| = 1$ und $\beta = \langle w, u_n \rangle$, dann ist

$$u_n^\top x = \alpha_1 x_1 + \dots + \alpha_n x_n = \beta$$

eine Darstellung der Hyperebene in Hessescher Normalform.

Im \mathbb{R}^n lässt sich eine Normalrichtung zu linear unabhängigen Vektoren $u^{(1)}, \dots, u^{(n-1)} \in \mathbb{R}^n$ wie folgt über den Laplace-Entwicklungssatz S 6.15 für Determinanten bestimmen:

Berechne formal $\det \begin{bmatrix} x_1 & u_1^{(1)} & & u_1^{(n-1)} \\ \vdots & \vdots & \dots & \vdots \\ x_n & u_n^{(1)} & & u_n^{(n-1)} \end{bmatrix} = \sum_{i=1}^n x_i \cdot \underbrace{(-1)^{i+1} \det U_i}_{=: \gamma_i}$.

U_i bezeichnet dabei die Matrix, die durch Streichung der i -ten Zeile aus $U = [u^{(1)}, \dots, u^{(n-1)}]$ entsteht.

Der Vektor $c = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$ mit $\gamma_i := (-1)^{i+1} \det U_i$ ist Normalvektor zu $\text{span}(u^{(i)})_{i=1}^{n-1}$, denn

$$x \in \text{span}(u^{(i)})_{i=1}^{n-1} \Leftrightarrow \det [x, u^{(1)}, \dots, u^{(n-1)}] = 0 \Leftrightarrow \sum_{i=1}^n \gamma_i x_i = c^\top x = 0.$$

\mathbb{R}^2 : $\det \begin{bmatrix} x_1 & a_1 \\ x_2 & a_2 \end{bmatrix} = x_1 a_2 - x_2 a_1 \Rightarrow c = \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$ ist Normalvektor zu $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$.

\mathbb{R}^3 : $\det \begin{bmatrix} x_1 & a_1 & b_1 \\ x_2 & a_2 & b_2 \\ x_3 & a_3 & b_3 \end{bmatrix} = x_1(a_2 b_3 - b_2 a_3) - x_2(a_1 b_3 - b_1 a_3) + x_3(a_1 b_2 - b_1 a_2)$,
 $\begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} =: a \times b$ heißt Vektorprodukt von a und b (nur im \mathbb{R}^3 !).

Das Vektorprodukt $a \times b$ steht normal auf a und b , $\|a \times b\|$ ist die Fläche des von a und b aufgespannten Parallelogramms (nutze dazu die nun folgende Übung), insbesondere ist $a \times b = 0$ genau dann, wenn a und b linear abhängig sind.

Ü: Zum Vektorprodukt im \mathbb{R}^3 :

1. Zeige $\|a \times b\|^2 = \|a\|^2 \cdot \|b\|^2 - (a^\top b)^2$ (im \mathbb{R}^3 !) und begründe damit die Aussage zur Fläche von oben.
2. Seien $v, w, a, b \in \mathbb{R}^3$ und $G_1 = \{x = v + t \cdot a : t \in \mathbb{R}\}$, $G_2 = \{y = w + t \cdot b : t \in \mathbb{R}\}$ zwei Geraden im \mathbb{R}^3 . Bestimme den „Abstand“ der beiden Geraden $d(G_1, G_2) = \min\{\|x - y\| : x \in G_1, y \in G_2\}$.

⊗

♡

9.3 Biorthogonale Basen und Adjungierte Abbildungen

Jedes $v \in V$ definiert über $\langle \cdot, v \rangle$ eine lineare Abbildung aus V^* und dies ist sogar eine Bijektion.

[Zur Erinnerung: $V^* = \text{Hom}_K(V, K)$ ist der Dualraum aus Abschnitt 4.3 und für $\dim V = n < \infty$ hat jede Basis $\mathcal{B} = (v_1, \dots, v_n)$ eine duale Basis $\mathcal{B}^* = (\varphi_1, \dots, \varphi_n)$ mit $\varphi_i(v_j) = \delta_{ij}$.]

Satz 9.15 (Rieszsche Abbildung)

Sei V ein endlichdimensionaler \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Für $v \in V$ definiert $v^*(u) = \langle u, v \rangle$, $u \in V$, eine lineare Abbildung $v^* \in V^*$. Die Rieszsche Abbildung $R: V \rightarrow V^*$, $v \mapsto v^*$ ist eine semilineare Bijektion, d. h., eine bijektive Abbildungen mit $R(v + w) = R(v) + R(w)$, $R(\alpha v) = \bar{\alpha} R(v)$.

Beweis: s. Vorlesung □

Über das Skalarprodukt wird jedes $v \in V$ direkt mit einem $v^* \in V^*$ verbunden, lateinisch bedeutet *adjungere* „verbinden, unter das gleiche Joch spannen“.

Definition 9.16 Ist V ein endlichdimensionaler \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, identifiziert man V mit V^* über die Rieszsche Abbildung.

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $\mathcal{B}^* = (\varphi_1, \dots, \varphi_n)$ die duale Basis, nennt man $\mathcal{B}^{\text{ad}} = \{u_j = R^{-1}(\varphi_j) : j = 1, \dots, n\}$ die biorthogonale/adjungierte (oder auch nur duale) Basis zu \mathcal{B} (sie erfüllt $\langle v_i, u_j \rangle = \varphi_j(v_i) = \delta_{ij}$).

Satz 9.17 Ist $\mathcal{B} = (v_1, \dots, v_n)$ ONB eines \mathbb{K} -Vektorraums, so gilt $\mathcal{B}^{\text{ad}} = \mathcal{B}$.

Beweis: s. Vorlesung □

Der Annulator U° eines Untervektorraums U lässt sich mit dem nächsten Resultat einfach mit U^\perp identifizieren.

Satz 9.18 Ist V endlichdimensionaler \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $U \subseteq V$ ein Untervektorraum, so gilt $U^\circ = R(U^\perp)$.

Beweis: s. Vorlesung □

Ist insbesondere $\mathcal{B}_{U^\perp} = (v_1, \dots, v_r)$ eine Basis von U^\perp , dann ist $U^\circ = \{\langle \cdot, \sum_{i=1}^r \alpha_i v_i \rangle : \alpha \in \mathbb{K}^r\}$.

In Abschnitt 4.3 wurde zu einer Abbildung $F \in \text{Hom}_K(V, W)$ die duale Abbildung $F^* \in \text{Hom}_K(W^*, V^*)$, $\psi \in W^* \mapsto \psi \circ F \in V^*$ eingeführt und bei endlichen Basen \mathcal{A} von V und \mathcal{B} von W zeigte B 5.8 $M_{\mathcal{B}}^{\mathcal{A}}(F)^\top = M_{\mathcal{A}^*}^{\mathcal{B}^*}(F^*)$.

Ist $\langle \cdot, \cdot \rangle_V$ Skalarprodukt für V , $\langle \cdot, \cdot \rangle_W$ Skalarprodukt für W , bietet es sich an, die Rieszsche Identifikation von V mit V^* und von W mit W^* zu nutzen, um F^* direkt in eine Abbildung von W nach V zu übersetzen:

$$\begin{array}{ccc}
 V^* & \xleftarrow{F^*} & W^* \\
 R_V \uparrow \simeq & & \simeq \uparrow R_W \\
 & (?) & \\
 V & \xleftarrow{F^{\text{ad}}} & W
 \end{array}
 \quad (\simeq \text{ für endlichdimensional}).$$

Definition 9.19 Seien $F \in \text{Hom}_{\mathbb{K}}(V, W)$ und V, W \mathbb{K} -Vektorräume mit Skalarprodukten $\langle \cdot, \cdot \rangle_V, \langle \cdot, \cdot \rangle_W$.

Eine Abbildung $G \in \text{Hom}_{\mathbb{K}}(W, V)$ heißt adjungiert zu F , falls

$$\langle F(v), w \rangle_W = \langle v, G(w) \rangle_V \quad \text{für alle } v \in V, w \in W.$$

Satz 9.20 Sind V, W endlichdimensionale \mathbb{K} -Vektorräume mit Skalarprodukten $\langle \cdot, \cdot \rangle_V, \langle \cdot, \cdot \rangle_W$, dann gibt es zu jedem $F \in \text{Hom}_{\mathbb{K}}(V, W)$ genau ein $F^{\text{ad}} \in \text{Hom}_{\mathbb{K}}(W, V)$, das zu F adjungiert ist. Es gilt

$$F^{\text{ad}} = R_V^{-1} \circ F^* \circ R_W \quad \text{und} \quad (F^{\text{ad}})^{\text{ad}} = F.$$

Beweis: s. Vorlesung □

Es gilt damit immer $\langle F(v), w \rangle_W = \langle v, F^{\text{ad}}(w) \rangle_V$
 und $\langle w, F(v) \rangle_W = \langle F^{\text{ad}}(w), v \rangle_V$.

Man darf daher mit F „die Seite im Skalarprodukt wechseln“, indem man

- durch $^{\text{ad}}$ zur Adjungierten übergeht,
- und das Skalarprodukt wechselt.

Für $F \in \text{End}_{\mathbb{K}}(V)$ bleibt das Skalarprodukt dasselbe.

Satz 9.21 (Matrixdarstellung der adjungierten Abbildung)

Sei $F \in \text{Hom}_{\mathbb{K}}(V, W)$ für endlichdimensionale \mathbb{K} -Vektorräume V, W mit Skalarprodukten und Basen \mathcal{A}, \mathcal{B} . Es gilt:

$$M_{\mathcal{A}^{\text{ad}}}^{\mathcal{B}^{\text{ad}}}(F^{\text{ad}}) = \overline{M_{\mathcal{B}}^{\mathcal{A}}(F)}^{\top} =: M_{\mathcal{B}}^{\mathcal{A}}(F)^*.$$

Sind \mathcal{A}, \mathcal{B} Orthonormalbasen, gilt $M_{\mathcal{A}}^{\mathcal{B}}(F^{\text{ad}}) = M_{\mathcal{B}}^{\mathcal{A}}(F)^*$.

Beweis: s. Vorlesung □

Beachte: Für $a, b \in \mathbb{C}$ ist $\overline{ab} = \bar{a}\bar{b}$, daher ist $(AB)^* = \overline{(AB)}^{\top} = \overline{B}^{\top} \overline{A}^{\top} = B^* A^*$.

Beispiele:

- $\mathbb{K} = \mathbb{R}$: $V = \mathbb{R}^n, \langle u, v \rangle_V = \sum_{i=1}^n v_i u_i = v^{\top} u \quad (= u^{\top} v) \quad u, v \in \mathbb{R}^n,$
 $W = \mathbb{R}^m, \langle p, q \rangle_W = \sum_{i=1}^m q_i p_i = q^{\top} p \quad (= p^{\top} q) \quad p, q \in \mathbb{R}^m.$

$$A \in \mathbb{R}^{m \times n}: A^* = A^{\top},$$

$$\langle Au, q \rangle_W = q^{\top} Au = (A^{\top} q)^{\top} u = \langle u, A^{\top} q \rangle_V = \langle A^{\top} q, u \rangle_V = u^{\top} A^{\top} q.$$

- $\mathbb{K} = \mathbb{C}$: $V = \mathbb{C}^n, \langle u, v \rangle_V = \sum_{i=1}^n \bar{v}_i u_i = v^* u \quad (= \overline{u^* v}) \quad u, v \in \mathbb{C}^n,$
 $W = \mathbb{C}^m, \langle p, q \rangle_W = \sum_{i=1}^m \bar{q}_i p_i = q^* p \quad (= \overline{p^* q}) \quad p, q \in \mathbb{C}^m.$

$$A \in \mathbb{C}^{m \times n}: A^* = \bar{A}^{\top} (!)$$

Vorsicht! A^* bezeichnet traditionell die *adjungierte* und *nicht die duale* Abbildung zu A (die duale ist A^{\top})! ▽

$$\langle Au, q \rangle_W = q^* Au = (A^* q)^* u = \langle u, A^* q \rangle_V = \overline{\langle A^* q, u \rangle_V} = \overline{u^* A^* q}.$$

♡

Ü: Für endlichdimensionale Vektorräume V, W mit Skalarprodukten ist das Bilden der Adjungierten $\cdot^{\text{ad}}: \text{Hom}_{\mathbb{K}}(V, W) \rightarrow \text{Hom}_{\mathbb{K}}(W, V)$ bijektiv und für $\mathbb{K} = \mathbb{R}$ ein Vektorraumisomorphismus. Es gilt:

$$\begin{aligned} (F + G)^{\text{ad}} &= F^{\text{ad}} + G^{\text{ad}}, \\ (\alpha F)^{\text{ad}} &= \bar{\alpha} F^{\text{ad}}, \\ (F \circ G)^{\text{ad}} &= G^{\text{ad}} \circ F^{\text{ad}}. \end{aligned}$$

⊗

9.4 Orthogonale und Unitäre Abbildungen

Von besonderer Bedeutung sind Abbildungen, die den Wert des Skalarprodukts, also auch den Abstand und, im \mathbb{R}^n , die Winkel erhalten.

Definition 9.22 Seien V, W \mathbb{K} -Vektorräume mit Skalarprodukten $\langle \cdot, \cdot \rangle_V$, $\langle \cdot, \cdot \rangle_W$. Ein $F \in \text{Hom}_{\mathbb{K}}(V, W)$ mit

$$\langle F(u), F(v) \rangle_W = \langle u, v \rangle_V \quad \text{für alle } u, v \in V$$

heißt für $\mathbb{K} = \mathbb{R}$ *orthogonal* und für $\mathbb{K} = \mathbb{C}$ *unitär*.

Ein $F \in \text{Hom}_{\mathbb{K}}(V, W)$ heißt *Isometrie*, falls $\|F(v)\|_W = \|v\|_V$ für alle $v \in V$.

Beobachtung 9.23

- (a) Eine lineare Abbildung ist genau dann orthogonal/unitär, wenn sie eine Isometrie ist.
- (b) Jede Isometrie ist injektiv.
- (c) Für jede bijektive Isometrie F ist auch F^{-1} eine Isometrie.
- (d) Sind $G \in \text{Hom}_{\mathbb{K}}(U, V)$ und $F \in \text{Hom}_{\mathbb{K}}(V, W)$ Isometrien, so auch $F \circ G$.

Beweis: s. Vorlesung □

Definition 9.24 Gibt es eine bijektive Isometrie zwischen zwei \mathbb{K} -Vektorräumen mit Skalarprodukt, nennt man sie *isometrisch isomorph*.

Satz 9.25 (Fredholmsche Alternative)

Zwei endlichdimensionale \mathbb{K} -Vektorräume mit Skalarprodukt sind genau dann isometrisch isomorph, wenn sie die gleiche Dimension haben.

Beweis: s. Vorlesung □

Satz 9.26 Für $F \in \text{End}_{\mathbb{K}}(V)$ mit endlichdimensionalem \mathbb{K} -Vektorraum V mit Skalarprodukt sind äquivalent:

- (a) F ist unitär/orthogonal,
- (b) $F^{\text{ad}} \circ F = \text{id}_V$,
- (c) $F \circ F^{\text{ad}} = \text{id}_V$,
- (d) F ist invertierbar und $F^{\text{ad}} = F^{-1}$.

Beweis: s. Vorlesung □

Gemeinsam mit B 9.23(c),(d) zeigt das, dass orthogonale/unitäre Endomorphismen mit der Komposition \circ eine Gruppe bilden. Da Isometrien in keiner Richtung „strecken“ können, haben auch die Eigenwerte und die invarianten Untervektorräume dieser Endomorphismen besonders schöne Eigenschaften.

Satz 9.27 Sei V ein \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$.

- (a) Ist $\dim V < \infty$, bildet
 für $\mathbb{K} = \mathbb{R}$ $O(V) := \{F \in \text{End}_{\mathbb{R}}(V) : F \text{ orthogonal}\}$
 mit der Komposition \circ die orthogonale Gruppe $(O(V), \circ)$,
 für $\mathbb{K} = \mathbb{C}$ $U(V) := \{F \in \text{End}_{\mathbb{C}}(V) : F \text{ unitär}\}$
 mit der Komposition \circ die unitäre Gruppe $(U(V), \circ)$.
- (b) Ist $F \in \text{End}_{\mathbb{K}}(V)$ orthogonal/unitär und $U \subseteq V$ ein endlichdimensionaler F -invarianter Untervektorraum, dann ist auch U^\perp F -invariant.
- (c) Jeder Eigenwert $\lambda \in \mathbb{K}$ eines orthogonalen/unitären $F \in \text{End}_{\mathbb{K}}(V)$ erfüllt $|\lambda| = 1$ und für Eigenwerte $\lambda \neq \mu$ sind zugehörige Eigenvektoren orthogonal.

Beweis: s. Vorlesung □

Orthogonale/Unitäre Matrizen

Die Matrixdarstellungen orthogonaler/unitärer Abbildungen haben besonders schöne Eigenschaften. Sie werden oft als Transformationsmatrizen von einer Orthonormalbasis zu einer anderen genutzt und bilden die Grundlage vieler numerischer Algorithmen.

Definition 9.28 Eine Matrix $A \in M(n \times n, \mathbb{R})$ heißt orthogonal, wenn die Spalten von $A = (a_{ij})$ eine ONB von \mathbb{R}^n bzgl. des kanonischen Skalarprodukts bilden, d. h., wenn

$$\langle A_{\bullet,j}, A_{\bullet,k} \rangle_{\mathbb{R}^n} = \sum_{i=1}^n a_{ij}a_{ik} = \delta_{jk}, \text{ also } A^\top A = I_n.$$

Eine Matrix $A \in M(n \times n, \mathbb{C})$ heißt unitär, wenn die Spalten von $A = (a_{ij})$ eine ONB von \mathbb{C}^n bzgl. des kanonischen Skalarprodukts bilden, d. h., wenn

$$\langle A_{\bullet,j}, A_{\bullet,k} \rangle_{\mathbb{C}^n} = \sum_{i=1}^n a_{ij}\bar{a}_{ik} = \delta_{jk}, \text{ also } A^* A = I_n.$$

Dabei bezeichnet für $B \in \mathbb{C}^{m \times n}$ $B^* = \bar{B}^\top$ die adjungierte Matrix zu B .

Satz 9.29 Folgende Bedingungen sind äquivalent für $A \in \mathbb{K}^{m \times n}$:

- | | |
|--|--|
| $\mathbb{K} = \mathbb{R}$ | $\mathbb{K} = \mathbb{C}$ |
| (a) A ist orthogonal. | (a) A ist unitär. |
| (b) Die Spalten von A bilden eine ONB von \mathbb{R}^n . | (b) Die Spalten von A bilden eine ONB von \mathbb{C}^n . |
| (c) Die Zeilen von A bilden eine ONB von \mathbb{R}^n . | (c) Die Zeilen von A bilden eine ONB von \mathbb{C}^n . |
| (d) $A^\top A = I_n$. | (d) $A^* A = I_n$. |
| (e) $AA^\top = I_n$. | (e) $AA^* = I_n$. |
| (f) A ist regulär und $A^{-1} = A^\top$. | (f) A ist regulär und $A^{-1} = A^*$. |

Beweis: s. Vorlesung □

Korollar 9.30 Für jede orthogonale/unitäre $n \times n$ Matrix A ist $|\det A| = 1$.

Beweis: s. Vorlesung □

Satz 9.31 Sei V \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $\dim V = n < \infty$. Für $F \in \text{End}_{\mathbb{K}}(V)$ sind äquivalent:

- (a) F ist Isometrie.
- (b) F bildet ONBs auf ONBs ab.
- (c) F bildet eine ONB auf eine ONB ab. [Man muss nur eine prüfen!]
- (d) Es existiert eine ONB \mathcal{B} mit $M_{\mathcal{B}}(F)$ orthogonal/unitär.
- (e) Für jede ONB \mathcal{B} ist $M_{\mathcal{B}}(F)$ orthogonal/unitär.

Beweis: s. Vorlesung □

Satz 9.32 Sei $n \in \mathbb{N}$.

- (a) Mit der Matrixmultiplikation bildet
 - $O(n) := \{A \in \mathbb{R}^{n \times n} : A \text{ orthogonal}\}$ die orthogonale Gruppe,
 - $U(n) := \{A \in \mathbb{C}^{n \times n} : A \text{ unitär}\}$ die unitäre Gruppe.
- (b) Jeder Eigenwert $\lambda \in \mathbb{R}$ ($\in \mathbb{C}$) einer orthogonalen (unitären) $n \times n$ Matrix erfüllt $|\lambda| = 1$ und für Eigenwerte $\lambda \neq \mu$ sind zugehörige Eigenvektoren orthogonal.

Beweis: s. Vorlesung □

Eine wesentliche Anwendung orthogonaler Matrizen besteht darin, dass sich jede Matrix $A \in \mathbb{K}^{m \times n}$ als Produkt einer orthogonalen Matrix Q und einer oberen Dreiecksmatrix R schreiben lässt, es also eine „Zerlegung“ $A = Q \cdot R$ gibt. Die Konstruktion dieser Zerlegung bietet auch eine numerisch gute Variante, Gleichungssysteme zu lösen: Die Inverse zu Q ist Q^{\top} (Q^*), daher gilt $Ax = QRx = b \Leftrightarrow Rx = Q^{\top}b$ und das zweite System ist wegen der Dreiecksgestalt direkt lösbar.

Satz (QR-Zerlegung) Für jede Matrix $A \in \mathbb{K}^{m \times n}$ gibt es eine orthogonale (unitäre) Matrix $Q \in O(m)$ ($\in U(m)$), sodass $A = Q \cdot R$ mit $R \in \mathbb{K}^{m \times n}$ in Zeilenstufenform (obere Dreiecksmatrix).

Beweis: Verwende Gram-Schmidt-Orthogonalisierung L9.10 auf den Spalten von A :

Algorithmus (v_1, \dots, v_m wird die ONB in den Spalten von Q)

1. $\hat{i} \leftarrow 0$.
2. Für $j = 1, \dots, n$ setze $u \leftarrow A_{\bullet, j}$.
 - (2a) Für $i = 1, \dots, \hat{i}$: $R_{ij} := v_i^* u$, $u \leftarrow u - R_{ij} \cdot v_i$.

(2b) Falls $u \neq 0$, setze $\hat{i} \leftarrow \hat{i} + 1$, $v_i \leftarrow \frac{u}{\|u\|}$, $R_{i,j} := \|u\| = v_i^* u$.

[Nun ist $u \leftarrow u - R_{i,j} v_i = u - \|u\| \frac{u}{\|u\|} = 0$.]

(2c) Für $i = \hat{i} + 1, \dots, m$: $R_{i,j} \leftarrow 0$.

3. Ist $\hat{i} < m$, ergänze (v_1, \dots, v_i) zu einer ONB von \mathbb{R}^m , $Q = [v_1, \dots, v_m]$.

Beachte: Wegen $v_i^* v_j = \delta_{ij}$ ist

$$R_{ij} = v_i^* \cdot \left(A_{\bullet,j} - \sum_{\substack{h=1 \\ h \neq i}}^n R_{h,j} v_h \right) = v_i^* A_{\bullet,j} \quad \text{für } i = 1, \dots, m.$$

Daher ist $Q^T A = R$ und R in Zeilenstufenform nach Konstruktion. \square

R beschreibt A , wenn im Bildraum die ONB $Q = [v_1, \dots, v_m]$ verwendet wird, und Q^T beschreibt die Transformationsmatrix des Basiswechsels zu Q ($Q^T v = \begin{pmatrix} v_1^* v \\ \vdots \\ v_m^* v \end{pmatrix}$ sind die neuen Koordinaten von v zur Basis (v_1, \dots, v_m)).

Numerisch gibt es wieder bessere Verfahren zur Bestimmung einer QR-Zerlegung, denn wenn u nach der Ausführung von Schritt (2a) im Schritt (2b) sehr klein aber nicht Null ist, ist es nicht gut als nächster Basisvektor geeignet (durch „Stellenauslöschung“ ist es oft nicht mehr zu den früheren orthogonal).

Orthogonale 2x2 Matrizen, $O(2)$

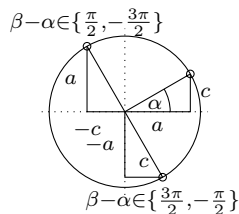
Die Struktur orthogonaler 2×2 Matrizen ist besonders einfach, hilft aber später auch ein gutes Verständnis für den n -dimensionalen Fall zu entwickeln.

Sei $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, dann ist

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in O(2) \Leftrightarrow \begin{cases} \|A_{\bullet,1}\| = 1 & \Leftrightarrow a^2 + c^2 = 1, \\ \wedge \|A_{\bullet,2}\| = 1 & \Leftrightarrow b^2 + d^2 = 1, \\ \wedge (A_{\bullet,2})^T A_{\bullet,1} = 0 & \Leftrightarrow ab + cd = 0. \end{cases}$$

Nun ist

$$\begin{aligned} a^2 + c^2 = 1 & \Leftrightarrow a = \cos \alpha, c = \sin \alpha \text{ für ein } \alpha \in [0, 2\pi), \\ b^2 + d^2 = 1 & \Leftrightarrow b = \cos \beta, d = \sin \beta \text{ für ein } \beta \in [0, 2\pi), \\ ab + cd = 0 & \stackrel{\text{Winkelsumme}}{\Leftrightarrow} \cos(\beta - \alpha) = 0 \Leftrightarrow (\beta - \alpha) \in \left\{ \pm \frac{\pi}{2}, \pm \frac{3\pi}{2} \right\}. \end{aligned}$$



$$\beta - \alpha \in \left\{ \frac{\pi}{2}, -\frac{3\pi}{2} \right\}: \cos \beta = -\sin \alpha, \sin \beta = \cos \alpha,$$

$$\beta - \alpha \in \left\{ \frac{3\pi}{2}, -\frac{\pi}{2} \right\}: \cos \beta = \sin \alpha, \sin \beta = -\cos \alpha.$$

1. Fall: $A = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$, $\det A = (\cos \alpha)^2 + (\sin \alpha)^2 = 1$.

A dreht Vektoren um den Winkel α und ist orientierungstreu.

Das charakteristische Polynom lautet $(\lambda - \cos \alpha)^2 + \sin(\alpha)^2$.

Eigenwerte: $\alpha = 0$: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, zweimal Eigenwert 1,

$\alpha = \pi$: $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, zweimal Eigenwert -1,

sonst: keine reelle Lösungen, keine Eigenwerte.

2. Fall: $A = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix}$, $\det A = -(\cos \alpha)^2 - (\sin \alpha)^2 = -1$.

A spiegelt Vektoren entlang der Achse $\{\rho \cdot \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix} : \rho \in \mathbb{R}\}$ und ist nicht orientierungstreu.

Das charakteristische Polynom lautet $\lambda^2 - (\cos \alpha)^2 - (\sin \alpha)^2$.

Die Eigenwerte sind stets 1 und -1 mit den Eigenvektoren

$$\text{zu } +1: \quad v_1 = \begin{pmatrix} \cos \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} \end{pmatrix}, \quad Av_1 = \begin{pmatrix} \cos \alpha \cos \frac{\alpha}{2} + \sin \alpha \sin \frac{\alpha}{2} \\ \sin \alpha \cos \frac{\alpha}{2} - \cos \alpha \sin \frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \cos(\alpha - \frac{\alpha}{2}) \\ \sin(\alpha - \frac{\alpha}{2}) \end{pmatrix},$$

$$\text{zu } -1: \quad v_{-1} = \begin{pmatrix} \sin \frac{\alpha}{2} \\ -\cos \frac{\alpha}{2} \end{pmatrix}, \quad Av_{-1} = \begin{pmatrix} \cos \alpha \sin \frac{\alpha}{2} - \sin \alpha \cos \frac{\alpha}{2} \\ \sin \alpha \cos \frac{\alpha}{2} + \cos \alpha \sin \frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} -\sin(\alpha - \frac{\alpha}{2}) \\ \cos(\alpha - \frac{\alpha}{2}) \end{pmatrix}.$$

Sei $v = \xi_1 v_1 + \xi_2 v_{-1}$, dann sind $\xi_1 = \langle v, v_1 \rangle = v_1^\top v$, $\xi_2 = \langle v, v_{-1} \rangle = v_{-1}^\top v$ die Koordinaten von v bezüglich der Basis aus Eigenvektoren. Die Matrix

$P = [v_1, v_{-1}] = \begin{bmatrix} \cos \frac{\alpha}{2} & \sin \frac{\alpha}{2} \\ \sin \frac{\alpha}{2} & -\cos \frac{\alpha}{2} \end{bmatrix}$ ist orthogonal, $P^\top P = I_n$ und

$$Av = 1 \cdot \xi_1 v_1 + (-1) \cdot \xi_2 v_{-1} = P \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = P \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot P^\top v.$$

Ähnlich zur alternierenden Gruppe, die mit allen Permutationen mit positivem Signum eine Untergruppe der Permutationsgruppe bildet, bilden auch die orientierungstreuen orthogonalen Abbildungen (mit Determinante 1) eine Untergruppe der orthogonalen Abbildungen. Die Definition über die Determinante lässt sich auch auf unitäre Abbildungen übertragen.

Satz 9.33 Sei $n \in \mathbb{N}$.

- $SO(n) = \{A \in \mathbb{R}^{n \times n} : A \text{ orthogonal und } \det A = 1\}$ bildet mit der Matrixmultiplikation die spezielle orthogonale Gruppe als Untergruppe von $O(n)$.
- $SU(n) = \{A \in \mathbb{C}^{n \times n} : A \text{ unitär und } \det A = 1\}$ bildet mit der Matrixmultiplikation die spezielle unitäre Gruppe als Untergruppe von $U(n)$.

Beweis: s. Vorlesung □

Wie im Fall der Spiegelungen lässt sich für jede orthogonale Matrix A eine ONB $P = [v_1, \dots, v_n]$ mit $P^\top P = I_n$ finden, sodass sich aus $P^\top A P$ direkt ablesen lässt, welche Richtungen gespiegelt und welche Unterräume gedreht werden. Dass für jede Isomorphie F über eine passend gewählte ONB eine

- $y = Bx$ betrifft nur einzelne Koordinaten oder Koordinatenpaare:
 $B_{ii} = 1$: $y_i = x_i$ (A ändert nichts in Richtung v_i),
 $B_{ii} = -1$: $y_i = -x_i$ (A spiegelt die Richtung v_i),
 $B_{\{i,i+1\}^2} = C(\alpha_h)$: $\begin{pmatrix} y_i \\ y_{i+1} \end{pmatrix} = \begin{bmatrix} \cos \alpha_h & \sin \alpha_h \\ -\sin \alpha_h & \cos \alpha_h \end{bmatrix} \begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix}$ dreht um α_h ,
 (A dreht Untervektorraum $\text{span}(v_i, v_{i+1})$ um α_h).
- Py transformiert die Koordinaten bezüglich \mathcal{B} wieder in die Koordinaten bezüglich der kanonischen Basis.

Die Vektoren in \mathcal{B} sind also die Richtungen/Untervektorräume, bezüglich derer die Wirkung von A geometrisch direkt ablesbar ist. Jede orthogonale Transformation ist eine direkte Summe von eindimensionalen unveränderten oder gespiegelten Unterräumen und zweidimensionalen gedrehten Untervektorräumen.

An der Darstellung erkennt man gut, dass für jede orthogonale/unitäre Matrix A stets $|\det A| = 1$ gilt: Sei $B = PAP^*$ die (block)diagonalisierte Form, dann ist $\det B$ das Produkt der Determinanten der Diagonalelemente/blöcke, deren Betrag jeweils 1 ist, also ist $|\det B| = 1$ und

$$1 = |\det B| = |\det(PAP^*)| = \underbrace{|\det P \det P^*|}_{=\det I_n=1} |\det A| = |\det A|.$$

Dies stimmt auch mit der geometrischen Interpretation der Determinante als Vorzeichenvolumen überein, denn die Spalten von A beschreiben ein ONB und spannen daher einen Einheitswürfel auf.

9.5 Normale Endomorphismen

Für endlichdimensionale unitäre Vektorräume ($\mathbb{K} = \mathbb{C}$) zeigt S 9.34(b), dass man zu jedem unitären Endomorphismus eine ONB aus Eigenvektoren finden kann. Es stellt sich nun die Frage, für welche Abbildungen zwischen *unitären* Vektorräumen (also derzeit nur für $\mathbb{K} = \mathbb{C}$) das allgemein möglich ist. Es wird sich herausstellen, dass dafür die folgende Eigenschaft unitärer Endomorphismen zur Charakterisierung ausreicht.

Definition 9.35 Ein $F \in \text{End}_{\mathbb{C}}(V)$ heißt *normal*, falls $F \circ F^{\text{ad}} = F^{\text{ad}} \circ F$.

Unitäre F sind normal, für diese ist nach S 9.26 $F \circ F^{\text{ad}} = F^{\text{ad}} \circ F = \text{id}_V$.

Matrizen mit $F = F^{\text{ad}}$ sind normal: $F \cdot F^{\text{ad}} = F \cdot F = F^{\text{ad}} \cdot F$.

Diagonalmatrizen $D \in \mathbb{C}^{n \times n}$ sind normal: $D \cdot D^* = D^* \cdot D$.

[Zur Erinnerung: Für F, G diagonalisierbar ist $F \circ G = G \circ F \Leftrightarrow F, G$ sind simultan diagonalisierbar über eine Basis \mathcal{B} , $F \circ G = \Phi_{\mathcal{B}} \circ D_F \circ D_G \circ \Phi_{\mathcal{B}}^{-1}$.]

Lemma 9.36 Sei $F \in \text{End}_{\mathbb{C}}(V)$ normal, dann gilt:

- (a) $\text{Ker } F = \text{Ker } F^{\text{ad}}$.

(b) für $\lambda \in \mathbb{C}$ ist $\text{Eig}(F, \lambda) = \text{Eig}(F^{\text{ad}}, \bar{\lambda})$,

Beweis: s. Vorlesung □

Satz 9.37 Für $F \in \text{End}_{\mathbb{C}}(V)$ mit $\dim V = n < \infty$ sind äquivalent:

(a) F ist normal.

(b) Es gibt eine ONB von V bestehend aus Eigenvektoren von F .

Beweis: s. Vorlesung □

Ein besonders wichtiger Spezialfall normaler Endomorphismen sind diejenigen, die $F = F^{\text{ad}}$ erfüllen, also zu sich selbst adjungiert sind.

9.6 Selbstadjungierte Endomorphismen

In diesem Abschnitt steht V durchgehend für einen endlichdimensionalen \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$.

Definition 9.38 Ein $F \in \text{End}_{\mathbb{K}}(V)$ heißt selbstadjungiert, wenn

$$\forall u, v \in V \quad \langle F(v), u \rangle = \langle v, F(u) \rangle \quad [\text{also } F = F^{\text{ad}}].$$

Ist $\mathbb{K} = \mathbb{R}$, nennt man F symmetrisch (symmetrischer Operator).

Ist $\mathbb{K} = \mathbb{C}$, nennt man F hermitesch (hermitescher Operator).

Eine Matrix $A \in M(n \times n, \mathbb{K})$ heißt selbstadjungiert, falls $A = A^*$ [$a_{ij} = \overline{a_{ji}}$].

Für $\mathbb{K} = \mathbb{R}$ heißt A dann auch symmetrische Matrix ($A = A^{\top}$).

Für $\mathbb{K} = \mathbb{C}$ heißt A dann auch hermitesche Matrix ($A = A^*$, $a_{ii} = \overline{a_{ii}} \in \mathbb{R}$).

Satz 9.39 Für $F \in \text{End}_{\mathbb{K}}(V)$ sind äquivalent:

(a) F ist selbstadjungiert.

(b) Es existiert eine ONB \mathcal{B} , sodass $M_{\mathcal{B}}(F)$ selbstadjungiert.

(c) $M_{\mathcal{B}}(F)$ ist für jede ONB \mathcal{B} selbstadjungiert.

Beweis: s. Vorlesung □

Lemma 9.40 Ist $F \in \text{End}_{\mathbb{K}}(V)$ selbstadjungiert und U ein F -invarianter Untervektorraum, so ist auch U^{\perp} F -invariant.

Beweis: s. Vorlesung □

Satz 9.41 (Spektralzerlegung selbstadj./symm./herm. Abb.)

Für jedes selbstadjungierte $F \in \text{End}_{\mathbb{K}}(V)$ gibt es eine ONB $\mathcal{B} = (v_1, \dots, v_n)$ mit

$$M_{\mathcal{B}}(F) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad \lambda_i \in \mathbb{R}(!), i = 1, \dots, n.$$

(Die λ_i sind Eigenwerte und \mathcal{B} ist eine ONB aus Eigenvektoren.)

$\mathbb{K} = \mathbb{C}$: Zu jeder hermiteschen Matrix $A \in \mathbb{C}^{n \times n}$ gibt es eine unitäre Matrix P , sodass $\Lambda_A = P^*AP$ eine reelle Diagonalmatrix ist.

$\mathbb{K} = \mathbb{R}$: Zu jeder symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$ gibt es eine orthogonale Matrix P , sodass $\Lambda_A = P^\top AP$ eine reelle Diagonalmatrix ist.

Die Spalten von (dem i. A. nicht eindeutigen) P bilden eine ONB bestehend aus Eigenvektoren und es gilt $A = P\Lambda_AP^* = \sum_{i=1}^n \lambda_i v_i v_i^*$.

Beweis: s. Vorlesung □

P ist immer dann nicht eindeutig (über die Reihenfolge der λ_i hinaus), wenn mindestens ein Eigenwert λ Vielfachheit größer als 1 hat; für so ein λ funktioniert jede ONB des Eigenraums $\text{Ker}(\lambda I - A)$.

Beispiel: $A = I = P^*P$ für jedes unitäre P . [Deshalb ist es auch bei der JNF egal, mit welchem maximalen F-zyklischen Untervektorraum man beginnt.]
♡

Ü: Zeige: Ist $A \in \mathbb{C}^{n \times n}$ hermitesch mit $\Lambda_A = P^*AP$, $P = [v_1, \dots, v_n]$ orthogonal und $\Lambda_A = \text{Diag}(\lambda_1, \dots, \lambda_n)$, dann ist $A = P\Lambda_AP^* = \sum_{i=1}^n \lambda_i v_i v_i^*$.

Zur geometrischen Interpretation betrachte $A \cdot v = P\Lambda_AP^*v$. $x = P^*v$ ist die Koordinatendarstellung von v bezüglich der ONB aus Eigenvektoren mit $x_i = v_i^*v = \langle v, v_i \rangle$. Bezüglich dieser Basis streckt/skaliert die Abbildung $\Lambda_A x$ jede Koordinate x_i mit λ_i . Nach der Skalierung erfolgt über $P \cdot (\lambda_1 x_1, \dots, \lambda_n x_n)^\top = \sum_{i=1}^n v_i (\lambda_i x_i)$ die Rücktransformation in das ursprüngliche Koordinatensystem. ⊗

Korollar 9.42 *Ist A eine symmetrische oder hermitesche Matrix, so gilt:*

- (a) $P_A(\xi) = \prod_{i=1}^k (\xi - \lambda_i)^{r_i}$ mit $\lambda_i \neq \lambda_j$ für $i \neq j$, $\sum_{i=1}^k r_i = n$.
- (b) $r_i = \mu(P_A; \lambda_i) = \dim \text{Eig}(A; \lambda_i)$ für $i = 1, \dots, k$.

Beweis: s. Vorlesung □

Aus der Spektralzerlegung $A = P\Lambda_AP^*$ für Eigenwerte $\lambda_i(A)$ lässt sich gut

$$\det A = \det(P\Lambda_AP^*) = \underbrace{\det P \det P^*}_{=\det I_n=1} \det \Lambda_A = \det \Lambda_A = \prod_{i=1}^n \lambda_i(A)$$

ableiten. Das charakteristische Polynom zeigt

$$\text{Spur} A = \sum_{i=1}^n A_{ii} = \sum_{i=1}^n \lambda_i(A).$$

Die Summe der Diagonalelemente einer selbstadjungierten Matrix ist also stets gleich der Summe der Eigenwerte.

Für diagonalisierbare Matrizen A und B gibt es genau dann eine gemeinsame Basis aus Eigenvektoren (sie sind simultan diagonalisierbar), wenn sie kommutieren ($AB = BA$). Im selbstadjungierten Fall gibt es sogar eine gemeinsame Orthonormalbasis.

Satz 9.43 Sind $F, G \in \text{End}_{\mathbb{K}}(V)$ selbstadjungiert, so sind äquivalent:

- (a) $F \circ G = G \circ F$ (die Abbildungen kommutieren).
- (b) Es existiert eine ONB bestehend aus Eigenvektoren von F und G .

Beweis: s. Vorlesung □

Für selbstadjungierte Matrizen gibt es hilfreiche Charakterisierungen der Eigenwerte über Optimierungsaufgaben. Diese werden in der Numerischen Mathematik zur Berechnung eingesetzt.

Für $A = A^*$ (bzw. für selbstadjungiertes $F \in \text{End}_{\mathbb{K}}(V)$) ist für beliebiges $x \in \mathbb{K}^n$ ($v \in V$)

$$(x^*Ax)^* = x^*Ax \in \mathbb{R} \quad (\text{bzw. } \langle v, F(v) \rangle = \langle F(v), v \rangle = \overline{\langle v, F(v) \rangle} \in \mathbb{R}),$$

daher ist für $x, y \in \mathbb{K}^n$ ein Größenvergleich wie $x^*Ax \leq y^*Ay$ wohldefiniert.

Obwohl die nächsten Ergebnisse nur für selbstadjungierte Matrizen formuliert sind, gelten sie in gleicher Weise für selbstadjungierte Endomorphismen in endlichdimensionalen Vektorräumen mit Skalarprodukt.

Satz 9.44 (Rayleigh-Ritz) Sei $A \in \mathbb{K}^{n \times n}$ selbstadjungiert mit Eigenwerten $\lambda_1 \leq \dots \leq \lambda_n$. Dann gilt $\lambda_1 x^*x \leq x^*Ax \leq \lambda_n x^*x$ für alle $x \in \mathbb{K}^n$.

Insbesondere ist $\lambda_1 = \min_{x \neq 0} \frac{x^*Ax}{x^*x} = \min_{\|x\|=1} x^*Ax$
 und $\lambda_n = \max_{x \neq 0} \frac{x^*Ax}{x^*x} = \max_{\|x\|=1} x^*Ax$.

Beweis: s. Vorlesung □

Schränkt man nun die Abbildung auf das orthogonale Komplement des Untervektorraums zu den kleinsten (größten) Eigenvektoren ein, lassen sich die nächstkleineren (nächstgrößeren) Eigenwerte bestimmen.

Beobachtung 9.45 Sei $A \in \mathbb{K}^{n \times n}$ selbstadjungiert mit Eigenwerten $\lambda_1 \leq \dots \leq \lambda_n$ und sei (v_1, \dots, v_n) eine ONB zugehöriger Eigenvektoren. Dann gilt für $k \in \{1, \dots, n\}$

$$\max_{\substack{x \neq 0 \\ x \perp v_n, \dots, v_{n-k+1}}} \frac{x^*Ax}{x^*x} = \max_{\substack{\|x\|=1 \\ x \perp v_n, \dots, v_{n-k+1}}} x^*Ax = \lambda_{n-k}, \quad \min_{\substack{\|x\|=1 \\ x \perp v_1, \dots, v_{k-1}}} x^*Ax = \lambda_k.$$

Beweis: s. Vorlesung □

Auch ohne Kenntnis einer ONB aus Eigenvektoren lässt sich der k -kleinste (k -größte) Eigenwert so charakterisieren, dass er der größte (kleinste) Eigenwert auf dem orthogonalen Komplement eines $n - k$ dimensionalen Untervektorraums ist, den ein Gegenspieler so wählt, dass dieser Wert möglichst klein (groß) wird.

Satz 9.46 (Courant-Fischer) Sei $A \in \mathbb{K}^{n \times n}$ selbstadjungiert mit Eigenwerten $\lambda_1 \leq \dots \leq \lambda_n$ und $k \in \{1, \dots, n\}$. Dann gilt:

$$\lambda_k = \min_{w_1, \dots, w_{n-k} \in \mathbb{K}^n} \max_{0 \neq x \perp w_1, \dots, w_{n-k}} \frac{x^* Ax}{x^* x},$$

$$\lambda_k = \max_{w_1, \dots, w_{k-1} \in \mathbb{K}^n} \min_{0 \neq x \perp w_1, \dots, w_{k-1}} \frac{x^* Ax}{x^* x}.$$

Dabei wird für $k = n$ der obere Ausdruck als $\max_{0 \neq x \in \mathbb{K}^n} \frac{x^* Ax}{x^* x}$ und für $k = 1$ der untere Ausdruck als $\min_{0 \neq x \in \mathbb{K}^n} \frac{x^* Ax}{x^* x}$ interpretiert.

Beweis: s. Vorlesung □

Die w_1, \dots, w_{n-k} legen den Untervektorraum fest, auf den x orthogonal steht,

$$\lambda_k = \min_{\substack{\text{UVR } U \subseteq \mathbb{K}^n \\ \dim U = n - k}} \max_{0 \neq x \in U^\perp} \frac{x^* Ax}{x^* x}.$$

Ein bestes U wird durch die (nicht notwendig eindeutigen) Eigenvektoren v_n, \dots, v_{k+1} aufgespannt. Auf dem (dann A -invarianten) orthogonalen Komplement U^\perp ist v_k ein Eigenvektor zum größten Eigenwert.

Wie stark können sich die Eigenwerte ändern, wenn ein Diagonalelement mit dazupassender Zeile und Spalte gelöscht (oder hinzugefügt) wird?

Satz 9.47 (von Cauchy über die Verschränkung der Eigenwerte)

Sei $\hat{A} = \begin{bmatrix} A & a \\ a^* & \alpha \end{bmatrix} \in \mathbb{K}^{(n+1) \times (n+1)}$ selbstadjungiert mit $a \in \mathbb{K}^n$ und seien $\lambda_1 \leq \dots \leq \lambda_n$ die Eigenwerte von A sowie $\hat{\lambda}_1 \leq \dots \leq \hat{\lambda}_{n+1}$ die Eigenwerte von \hat{A} . Es gilt $\hat{\lambda}_1 \leq \lambda_1 \leq \hat{\lambda}_2 \leq \dots \leq \hat{\lambda}_n \leq \lambda_n \leq \hat{\lambda}_{n+1}$.

Beweis: s. Vorlesung □

Ü: Zeige, dass die Aussage auch dann stimmt, wenn für beliebiges $i \in \{1, \dots, n + 1\}$ die i -te Zeile und Spalte gestrichen wird. Wie lässt sich die Aussage auf Abbildungen und Untervektorräume übertragen? ⊗

9.7 Positiv semidefinite Endomorphismen

In der Analysis haben symmetrische Matrizen schon wegen der Hessematrix (Werte der 2. Ableitung zweimal stetig differenzierbarer Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}$) besondere Bedeutung. Dort und in der Optimierung ist wesentlich, ob diese Matrix $x^T Ax > 0$ für alle $x \neq 0$ erfüllt, denn dann ist die Funktion lokal konvex (sie krümmt sich nach oben).

Definition 9.48 Ein $F \in \text{End}_{\mathbb{K}}(V)$ heißt positiv (semidefinit), falls F selbstadjungiert ist und $\langle v, F(v) \rangle \geq 0$ für alle $v \in V$ erfüllt.

Gilt sogar $\langle v, F(v) \rangle > 0$ für alle $v \in V \setminus \{0\}$, heißt F positiv definit.

Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt

- positiv semidefinit ($A \succeq 0$), falls $A = A^*$ und $v^*Av \geq 0 \ \forall v \in \mathbb{K}^n$,
- positiv definit ($A \succ 0$), falls $A = A^*$ und $v^*Av > 0 \ \forall v \in \mathbb{K}^n \setminus \{0\}$.

Beobachtung 9.49

- (a) Ist $A \in \mathbb{K}^{n \times n}$ positiv (semi-)definit, so ist auch jede Hauptuntermatrix $A_{\mathcal{J}, \mathcal{J}}$ mit $\mathcal{J} \subseteq \{1, \dots, n\}$ positiv (semi-)definit.
- (b) Eine Blockdiagonalmatrix $A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{bmatrix}$ mit $A_i \in \mathbb{K}^{n_i \times n_i}$ ist genau dann positiv (semi-)definit, wenn alle A_i positiv (semi-)definit sind.
- (c) Ist $A \in \mathbb{K}^{n \times n}$ und $B \in \text{GL}(n; \mathbb{K})$, so ist A genau dann positiv (semi-)definit, wenn auch B^*AB positiv (semi-)definit ist.
- (d) Ist $C \in \mathbb{K}^{m \times n}$, dann gilt
 - C^*C ist positiv semidefinit, $\text{Ker}(C) = \text{Ker}(C^*C)$,
 - CC^* ist positiv semidefinit, $\text{Ker}(C^*) = \text{Ker}(CC^*)$,
 - $\text{Rang}(C^*C) = \text{Rang}(CC^*) = \text{Rang}(C)$.

Beweis: s. Vorlesung □

Ü: Ist F positiv (semi-)definit, dann ist für jeden Untervektorraum U $F|_U$ positiv (semi-)definit. Wo spiegelt sich das in B 9.49 wieder? ⊗

Satz 9.50 (Charakterisierung positiv (semi-)definiter Matrizen)

Für $A \in \mathbb{K}^{n \times n}$ mit $A = A^*$ sind äquivalent:

- | | |
|---|---|
| <ul style="list-style-type: none"> (a) A ist positiv semidefinit, (b) $\exists m \in \mathbb{N}, C \in \mathbb{K}^{m, n}: A = C^*C$, (c) $\exists B \in \mathbb{K}^{n \times n}, B^* = B: A = B^2$, (d) Für λ EW von A gilt $\lambda \geq 0$, (e) $\exists P \in \mathbb{K}^{n \times n}, P^*P = I:$
 $P^*AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \lambda_i \geq 0.$ | <ul style="list-style-type: none"> (a) A ist positiv definit, (b) $\exists C \in \text{GL}(n, \mathbb{K}): A = C^*C$, (c) $\exists B \in \text{GL}(n; \mathbb{K}), B^* = B: A = B^2$, (d) Für λ EW von A gilt $\lambda > 0$, (e) $\exists P \in \mathbb{K}^{n \times n}, P^*P = I:$
 $P^*AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \lambda_i > 0.$ |
|---|---|

Beweis: s. Vorlesung □

Insbesondere ist $A \succ 0$ genau dann wenn $A^{-1} = P\Lambda_A^{-1}P^* \succ 0!$

Korollar 9.51 Für jedes positiv (semi-)definite $A \in \mathbb{K}^{n \times n}$ und $k \in \mathbb{N}$ gibt es eine eindeutig bestimmte positiv (semi-)definite Matrix $B \in \mathbb{K}^{n \times n}$ mit $A = B^k$, die k -te Wurzel $A^{\frac{1}{k}}$ aus A .

Beweis: s. Vorlesung □

Ein weiteres nützliches Hilfsmittel in der Arbeit mit positiv semidefiniten Matrizen ist das Schurkomplement.

Satz 9.52 (Schurkomplement) Sei $M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \in \mathbb{K}^{n \times n}$ mit $M^* = M$ und positiv definitem $A \in \mathbb{K}^{m \times m}$, dann gilt:

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \succeq 0 \Leftrightarrow C - B^*A^{-1}B \succeq 0 \quad \text{und} \quad \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \succ 0 \Leftrightarrow C - B^*A^{-1}B \succ 0.$$

Beweis: s. Vorlesung □

Damit lässt sich eine Matrix leicht algorithmisch auf positive Definitheit testen.

Satz (Cholesky Faktorisierung)

Zu positiv definitem $A \in \mathbb{K}^{n \times n}$ gibt es eine eindeutig bestimmte linke untere Dreiecksmatrix $L \in GL(n; \mathbb{K})$ mit $A = LL^*$ und positiver Diagonale.

Beweis: Nutze $A_n = \begin{bmatrix} \alpha_n & b_n^* \\ b_n & C_n \end{bmatrix} \succ 0 \stackrel{S9.52}{\Leftrightarrow} (\alpha_n > 0 \wedge \underbrace{C_n - \frac{1}{\alpha_n}b_nb_n^*}_{=:A_{n-1}} \succ 0)$.

Für $u_n = \begin{bmatrix} \sqrt{\alpha_n} \\ \frac{1}{\sqrt{\alpha_n}}b_n \end{bmatrix}$ ist dann $A_n = u_nu_n^* + \begin{bmatrix} 0 & 0 \\ 0 & A_{n-1} \end{bmatrix}$.

Algorithmus (Cholesky Faktorisierung)

1. Setze $A_n = A$.
2. Für $k = n, n - 1, \dots, 1$
 - Sei $A_k = \begin{bmatrix} \alpha_k & b_k^* \\ b_k & C_k \end{bmatrix}$.
 - Ist $\alpha_k \leq 0$ STOP; A ist nicht positiv definit.
 - Setze $u_k = \begin{bmatrix} 0_{(n-k) \times 1} \\ \sqrt{\alpha_k} \\ \frac{1}{\sqrt{\alpha_k}}b_k \end{bmatrix}$ und, falls $k > 1$, $A_{k-1} = C_k - \frac{1}{\alpha_k}b_kb_k^*$.

[Es gilt $A = \sum_{i=k}^n u_iu_i^* + \begin{bmatrix} 0_{(n-k+1) \times (n-k+1)} & 0 \\ 0 & A_{k-1} \end{bmatrix}$.]
3. Setze $L = [u_n, u_{n-1}, \dots, u_1]$.
 [$LL^* = \sum_{i=1}^m u_iu_i^* = A$ und L ist linke untere Dreiecksmatrix.]

Eindeutigkeit: Sei L linke untere Dreiecksmatrix mit positiver Diagonale und $A = LL^*$, dann gilt $A_{11} = L_{11}^2$, daher ist $L_{11} = \sqrt{A_{11}}$ und aus der Matrixgleichung folgt nun $L_{\bullet,1} = \begin{bmatrix} \sqrt{A_{11}} \\ \frac{1}{\sqrt{A_{11}}}b_n \end{bmatrix}$. Setze wie im Algorithmus rekursiv fort. □

Für positiv definites A lässt sich $Ax = b$ gut lösen: Berechne $LL^* = A$, löse zuerst $Ly = b$ unter Nutzung der Dreiecksgestalt und darauf $L^*x = y$, dann ist $Ax = LL^*x = Ly = b$.

Für positiv definite Matrizen gibt es eine weitere, für theoretische Untersuchungen sehr hilfreiche Charakterisierung.

Satz 9.53 Eine Matrix $A = A^* \in \mathbb{K}^{n \times n}$ ist genau dann positiv definit, wenn für eine beliebige Folge $\emptyset \subsetneq \mathcal{J}_1 \subsetneq \mathcal{J}_2 \subsetneq \dots \subsetneq \mathcal{J}_n = \{1, \dots, n\}$ die Determinanten der Hauptuntermatrizen zu \mathcal{J}_i streng positiv sind, also wenn $\det A_{\mathcal{J}_i, \mathcal{J}_i} > 0$ für $i = 1, \dots, n$.

Beweis: s. Vorlesung □

Vorsicht! Für $A \succeq 0$ folgt zwar $\det(A_{\mathcal{J}, \mathcal{J}}) \geq 0$ für $\mathcal{J} \subseteq \{1, \dots, n\}$, aber auch für $A \not\succeq 0$ kann es $\mathcal{J}_1, \dots, \mathcal{J}_n$ geben mit $\det(A_{\mathcal{J}_i, \mathcal{J}_i}) \geq 0$, zum Beispiel

$$A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathcal{J}_1 = \{1\}, \quad \det A_{\mathcal{J}_1, \mathcal{J}_1} = 0, \quad \mathcal{J}_2 = \{1, 2\}, \quad \det A = 0.$$

▽

Über positiv (semi-)definite Matrizen lassen sich Zerlegungen allgemeiner Matrizen/Abbildungen konstruieren, die für Theorie und Praxis große Bedeutung haben.

Im Vergleich zur Normalform von K 5.4, die für $F \in \text{Hom}_K(V, W)$ die Existenz von Basen \mathcal{A}, \mathcal{B} mit $M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ zusicherte, garantiert die nun folgende Singulärwertzerlegung sogar die Existenz von Orthonormalbasen P, Q sodass die Identität I_r nur durch eine positive Diagonalmatrix Σ ersetzt werden muss.

Satz 9.54 (Singulärwertzerlegung)

Für $C \in \mathbb{K}^{m \times n}$ und $r = \text{Rang}(C)$ gibt es eindeutig bestimmte Singulärwerte $\sigma_1 \geq \dots \geq \sigma_r > 0$ und (i.A. nicht eindeutige) Isometrien $Q \in \mathbb{K}^{m \times m}, P \in \mathbb{K}^{n \times n}$ mit $C = Q\Sigma P^*$, wobei $\Sigma = \begin{bmatrix} \text{Diag}(\sigma) & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{K}^{m \times n}$.

Beweis: s. Vorlesung □

Die Singulärwerte sind gerade die Wurzeln aus den Eigenwerten von C^*C oder CC^* (wie gerade gezeigt, haben beide dieselben positiven Eigenwerte aber i. A. unterschiedlich viele 0-Eigenwerte). Gelegentlich werden die 0-Eigenwerte $\sigma_i = 0, i = r + 1, \dots, \min\{n, m\}$ als Singulärwerte mitgezählt. Ebenso verzichtet man oft in Σ auf die Nullzeilen und -spalten. Hat C Rang r , wird die Singulärwertzerlegung in diesem Fall direkt mit $\Sigma = \text{Diag}(\sigma) \in \mathbb{K}^{r \times r}$ angegeben, $Q \in \mathbb{K}^{m \times r}$ und $P \in \mathbb{K}^{n \times r}$ enthalten dann jeweils nur die ersten r orthonormalen Spalten.

Eine weitere oft verwendete Zerlegung in eine Matrix mit orthonormalen Spalten und eine positiv semidefinite Matrix ist die Polardarstellung.

Satz 9.55 (Polardarstellung) Für $C \in \mathbb{K}^{m \times n}$ mit $m \geq n$ gibt es eine Matrix $U \in \mathbb{K}^{m \times n}$ mit orthonormalen Spalten und eine eindeutige positiv semidefinite Matrix $A \in \mathbb{K}^{n \times n}$ mit $C = UA$.

Hat C vollen Spaltenrang, ist auch U eindeutig.

Beweis: s. Vorlesung □

Zum Abschluss dieses Kapitels sollen als ausführliches Anwendungsbeispiel Matrixnormen besprochen werden.

Definition 9.56 Eine Norm $\|\cdot\|$ auf $\mathbb{K}^{n \times n}$ heißt Matrixnorm, falls sie zusätzlich submultiplikativ ist, d. h., $\|A \cdot B\| \leq \|A\| \cdot \|B\|$

Beispiel: Die Euklidische Matrixnorm, meist als Frobenius-Norm $\|A\|_F := \sqrt{\sum_{i,j=1,\dots,n} |a_{ij}|^2}$ (auch Schur- oder Hilbert-Schmidt-Norm) bezeichnet, ist eine Matrixnorm für $\mathbb{K}^{n \times n}$, denn

$$\begin{aligned} \|AB\|_F^2 &= \sum_{i,j=1,\dots,n} \left| \sum_{k=1}^n a_{ik} b_{kj} \right|^2 \\ &\stackrel{\text{(CS)}}{\leq} \sum_{i,j=1,\dots,n} \left(\sum_{k=1}^n |a_{ik}|^2 \right) \left(\sum_{k=1}^n |b_{kj}|^2 \right) \\ &= \sum_{i,j=1,\dots,n} |a_{ij}|^2 \sum_{i,j=1,\dots,n} |b_{ij}|^2 = \|A\|_F^2 \cdot \|B\|_F^2. \end{aligned}$$

♡

Satz 9.57 Sei $\|\cdot\|$ eine Norm auf \mathbb{K}^n . Dann definiert $\|A\| := \max_{\|x\|=1} \|Ax\|$ eine Matrixnorm, die induzierte/Operator-/least upper bound-Norm auf $\mathbb{K}^{n \times n}$. Sie erfüllt $\|A \cdot x\| \leq \|A\| \cdot \|x\|$ für alle $A \in \mathbb{K}^{n \times n}$, $x \in \mathbb{K}^n$ und $\|I\| = 1$.

Beweis: Norm: $\|\cdot\| \geq 0$ ist klar.

$$\|A\| = 0 \Leftrightarrow Ax = 0 \text{ für alle } x \text{ mit } \|x\| = 1 \Leftrightarrow A = 0.$$

$$\|\alpha A\| = \max_{\|x\|=1} \|\alpha Ax\| = |\alpha| \max_{\|x\|=1} \|Ax\| = |\alpha| \|A\|.$$

$$\begin{aligned} \|A + B\| &= \max_{\|x\|=1} \|Ax + Bx\| \\ &\stackrel{\text{S 9.5(c)}}{\leq} \max_{\|x\|=1} (\|Ax\| + \|Bx\|) \\ &\stackrel{\Delta\text{-Unglg.}}{\leq} \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \|A\| + \|B\|. \end{aligned}$$

Submultiplikativität: Ist $B = 0$, gilt $0 = \|B\| = \|A\| \cdot \|B\| \leq \|A \cdot B\|$, sonst

$$\begin{aligned} \|A \cdot B\| &= \max_{\|x\|=1} \|ABx\| = \max_{x \neq 0} \frac{\|ABx\|}{\|x\|} = \max_{Bx \neq 0} \frac{\|ABx\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|} \\ &\leq \max_{y \neq 0} \frac{\|Ay\|}{\|y\|} \cdot \max_{x \neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| \cdot \|B\|. \end{aligned}$$

$\|A \cdot x\| \leq \|A\| \cdot \|x\|$: Dies gilt für $x = 0$, sonst ist

$$\|Ax\| = \|x\| \cdot \left\| A \frac{x}{\|x\|} \right\| \leq \|x\| \max_{\|y\|=1} \|Ay\| = \|x\| \cdot \|A\|$$

Für die Identität gilt $\|I\| = \max_{\|x\|=1} \|Ix\| = \max_{\|x\|=1} \|x\| = 1$. □

Beispiele:

(a) Für die 1-Norm $\|x\|_1 = \sum_{i=1}^n |x_i|$ ist die induzierte Matrixnorm die

Spaltensummennorm $\|A\|_1 = \max_{j=1, \dots, n} \|A_{\bullet, j}\|_1$ (Ü)

(b) Für die ∞-Norm $\|x\|_\infty = \max_{i=1, \dots, n} |x_i|$ ist die induzierte Matrixnorm die

Zeilensummennorm $\|A\|_\infty = \max_{i=1, \dots, n} \|A_{i, \bullet}\|_1$ (Ü)

(c) Für die (Euklidische) 2-Norm $\|x\|_2 = \sqrt{x^*x}$ ist die induzierte Matrixnorm die

Spektralnrm $\|A\|_2 = \max\{\sqrt{\lambda} : \lambda \text{ ist EW von } A^*A\}$

denn: $\|A\|_2 = \max_{\|x\|=1} \|Ax\| = \max_{\|x\|=1} (x^*A^*Ax)^{\frac{1}{2}} = \left(\max_{\|x\|=1} x^*A^*Ax \right)^{\frac{1}{2}} \stackrel{\text{S 9.44}}{=} \sqrt{\lambda_{\max}(A^*A)}$. □

♡

Ist A invertierbar, dann ist $\|A^{-1}\|_2 = \frac{1}{\sigma_{\min}(A)}$, denn $A^*A \succ 0$, $AA^* \succ 0$ und

$$\lambda_{\max}((A^{-1})^*A^{-1}) = \lambda_{\max}((AA^*)^{-1}) = \frac{1}{\lambda_{\min}(AA^*)} = \frac{1}{\lambda_{\min}(A^*A)}$$

Man nennt $\kappa(A) := \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$ die Konditionszahl von A . Sie gibt eine Schranke dafür an, wie empfindlich die Lösung des Gleichungssystems $Ax = b$ gegenüber Störungen ist, wie die folgenden Überlegungen zeigen.

Sei b um ein (kleines) $\delta \in \mathbb{K}^n$ gestört, dann ist

$$A(x + \varepsilon) = b + \delta \text{ mit } A\varepsilon = \delta \text{ bzw. } \varepsilon = A^{-1}\delta, \text{ also}$$

$$\|\varepsilon\|_2 = \|A^{-1}\delta\| \leq \|A^{-1}\|_2 \|\delta\|.$$

Bei großem $\|A^{-1}\|_2$ können selbst kleine Störungen in b die Lösung x völlig verfälschen. Allerdings wird $\|A^{-1}\|_2$ immer groß, wenn man das ganze System mit einer kleinen Zahl multipliziert,

für $10^{-6}Ax = 10^{-6}b$ ist $\|(10^{-6}A)^{-1}\|_2 = 10^6 \|A^{-1}\|_2$.

In der Konditionszahl kürzen sich diese Skalierungsfaktoren wieder heraus,

$$\| \|10^{-6}A\|_2 \| \| (10^{-6}A)^{-1} \|_2 = \| \|A\|_2 \| \|A^{-1}\|_2 = \kappa(A).$$

Ein System mit kleinem (großem) $\kappa(A)$ heißt gut (schlecht) konditioniert. Im besten Fall ist $\kappa(A) = 1$, das ist nur für $A = \alpha P$ mit orthogonalem/unitärem P der Fall (**Ü**).

Bei einem schlecht konditionierten System sucht man ein „einfaches“ $B \in \text{GL}(n; \mathbb{K})$ zur Vorkonditionierung so, dass das System $BAx = Bb$ besser konditioniert ist. Ideal wäre dafür $B = A^{-1}$, denn dann ist $\kappa(BA) = \kappa(I) = 1$, aber dieses B zu bestimmen wäre aufwendiger als das System selbst zu lösen. Man nützt daher meist problemspezifische Vorabinformationen, um ein geeignetes B zu konstruieren; entsprechende Ansätze dazu werden in der Numerischen Mathematik thematisiert.

Kapitel 10

Bilinearformen und quadratische Formen

10.1 Bilinearformen

Zur Wiederholung sei daran erinnert, dass für einen K -Vektorraum V eine Abbildung $s: V \times V \rightarrow K$, $(v, w) \mapsto s(v, w)$ Bilinearform heißt, falls s in beiden Argumenten linear ist, also für alle $\alpha, \beta \in K$, $u, v, w \in V$

$$(B1) \quad s(\alpha u + \beta v, w) = \alpha s(u, w) + \beta s(v, w),$$

$$(B2) \quad s(u, \alpha v + \beta w) = \alpha s(u, v) + \beta s(u, w).$$

Eine Bilinearform s heißt symmetrisch, falls $s(v, w) = s(w, v)$ für alle $v, w \in V$.

Für $K = \mathbb{C}$ heißt s Sesquilinearform, falls (B1) gilt und im zweiten Argument

$$(\overline{B2}) \quad s(u, \alpha v + \beta w) = \bar{\alpha} s(u, v) + \bar{\beta} s(u, w) \text{ für alle } \alpha, \beta \in \mathbb{C}, u, v, w \in V.$$

Eine Sesquilinearform s heißt hermitesch, falls $s(v, w) = \overline{s(w, v)}$ für alle $v, w \in V$.

Zur Vereinfachung der Sprechweise werden wir im Folgenden nur von Bilinearformen sprechen, aber für $K = \mathbb{C}$ dabei Sesquilinearformen meinen.

Ist für V eine Basis $\mathcal{B} = (v_i)_{i \in \mathcal{J}}$ bekannt, ist wegen der Linearität jede Multilinearform schon durch die Werte auf den Kombinationen der Basisvektoren eindeutig festgelegt.

Satz 10.1 Sei V ein K -Vektorraum mit K beliebig und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis. Die Abbildung $s \mapsto M_{\mathcal{B}}(s) \in K^{n \times n}$ mit $[M_{\mathcal{B}}(s)]_{ij} = s(v_i, v_j)$ von den Bilinearformen auf V in $M(n \times n; K)$ ist bijektiv.

Sei das Koordinatensystem $\Phi_{\mathcal{B}}$ und $A = M_{\mathcal{B}}(s)$, dann gilt für $x, y \in K^n$ mit $v = \Phi_{\mathcal{B}}(x), w = \Phi_{\mathcal{B}}(y) \in V$

$$s(v, w) = x^{\top} A y \quad (\text{sesquilinear } x^{\top} A \bar{y})$$

und s ist genau dann symmetrisch (hermitesch), wenn A es ist.

Beweis: Sesquilinearer Fall: Für $v = \sum_{i=1}^n x_i v_i, w = \sum_{i=1}^n y_i v_i$ ist

$$s(v, w) = \sum_{i=1}^n x_i s(v_i, w) = \sum_{i=1}^n \sum_{j=1}^n x_i \bar{y}_j s(v_i, v_j) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i \bar{y}_j = x^\top A \bar{y}.$$

Dies zeigt die Bijektivität der Zuordnung zwischen s und A . Es ist s hermitesch $\Leftrightarrow s(v_i, v_j) = \overline{s(v_j, v_i)} \forall i, j \Leftrightarrow e_i^\top A e_j = \overline{e_j^\top A e_i} \forall i, j \Leftrightarrow A = A^*$. \square

Man nennt $M_{\mathcal{B}}(s)$ auch die **Gramsche Matrix** von s in der Basis \mathcal{B} . Ist \mathcal{A} eine andere Basis und $v = \Phi_{\mathcal{A}}(x_{\mathcal{A}}) = \Phi_{\mathcal{B}}(x_{\mathcal{B}})$ mit $T_{\mathcal{B}}^{\mathcal{A}} = \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}}$, dann ist

$$s(v, w) = x_{\mathcal{A}}^\top M_{\mathcal{A}}(s) \bar{y}_{\mathcal{A}} = (T_{\mathcal{B}}^{\mathcal{A}} x_{\mathcal{A}})^\top M_{\mathcal{B}}(s) \overline{(T_{\mathcal{B}}^{\mathcal{A}} y_{\mathcal{A}})}.$$

Das zeigt

Satz 10.2 (Transformationsformel)

Sei V ein K -Vektorraum mit endlichen Basen \mathcal{A} und \mathcal{B} und sei $T_{\mathcal{B}}^{\mathcal{A}}$ die Transformationsmatrix von \mathcal{A} nach \mathcal{B} , dann gilt

$$M_{\mathcal{A}}(s) = T_{\mathcal{B}}^{\mathcal{A}\top} M_{\mathcal{B}}(s) T_{\mathcal{B}}^{\mathcal{A}} \quad (\text{sesquilinear } T_{\mathcal{B}}^{\mathcal{A}\top} M_{\mathcal{B}}(s) \overline{T_{\mathcal{B}}^{\mathcal{A}}}).$$

Vorsicht! Die Struktur der Basistransformation $x = Cy$ mit $C \in \text{GL}(n; K)$ ist $C^{-1}AC$ für Endomorphismen $A \in K^{n \times n}$,
 $C^\top AC$ für Bilinearformen $A \in K^{n \times n}$ ($C^\top AC$ sesquilinear). ∇

Definition 10.3 Zwei Matrizen $A, B \in K^{n \times n}$ heißen **kongruent**, wenn es ein $C \in \text{GL}(n; K)$ gibt mit $B = C^* A C$.

Seien s_V, s_W Bilinearformen auf K -Vektorräumen V bzw. W . Man nennt (V, s_V) und (W, s_W) **isometrisch isomorph**, wenn es einen Isomorphismus $F \in \text{Hom}_K(V, W)$ gibt mit $s_W(Fu, Fv) = s_V(u, v)$ für alle $u, v \in V$.

Satz 10.4 Seien V, W K -Vektorräume mit endlichen Basen \mathcal{A} bzw. \mathcal{B} und Bilinearformen s_V bzw. s_W . (V, s_V) und (W, s_W) sind genau dann isometrisch isomorph, wenn gilt

- (a) $\dim V = \dim W$ und
- (b) $\exists C \in \text{GL}(n; K)$ mit $M_{\mathcal{A}}(s_V) = C^\top M_{\mathcal{B}}(s_W) C$ ($C^\top M_{\mathcal{B}}(s_W) \bar{C}$).

Beweis: \Rightarrow : Sei $F \in \text{Hom}_K(V, W)$ der Isomorphismus, dann folgt (a) sofort und für $C = M_{\mathcal{B}}^{\mathcal{A}}(F) = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}$, $\mathcal{A} = (v_1, \dots, v_n)$ mit $\Phi_{\mathcal{A}}(e_i) = v_i$ gilt

$$\begin{aligned} M_{\mathcal{A}}(s_V)_{ij} &\stackrel{\text{S 10.1}}{=} s_V(v_i, v_j) = s_W(Fv_i, Fv_j) \\ &\stackrel{\text{S 10.1}}{=} \Phi_{\mathcal{B}}^{-1}(Fv_i)^\top M_{\mathcal{B}}(s_W) \overline{\Phi_{\mathcal{B}}^{-1}(Fv_j)} \\ &= \underbrace{[(\Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}})(e_i)]^\top}_{=: C} M_{\mathcal{B}}(s_W) \bar{C} e_j \\ &= e_i^\top \underbrace{C^\top}_{=: C} M_{\mathcal{B}}(s_W) \bar{C} e_j = [C^\top M_{\mathcal{B}}(s_W) \bar{C}]_{ij}. \end{aligned}$$

$$\Leftarrow: F = \Phi_B \circ C \circ \Phi_A^{-1}. \quad \square$$

10.2 Symmetrische/Hermitesche Bilinearformen

Nach S 10.1 ist s_V genau dann symmetrisch (hermitesch), wenn $M_B(s_V)$ symmetrisch (hermitesch) ist. Für symmetrisches A erklärt

$$x^\top Ax = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_i x_j$$

ein homogenes quadratisches Polynom in den Unbestimmten x_1, \dots, x_n . Ein multivariates Polynom $p \in K[x_1, \dots, x_n]$ hat $n \geq 2$ Unbestimmte, jeder Term $x_1^{k_1} \cdots x_n^{k_n}$ mit $k_1, \dots, k_n \in \mathbb{N}_0$ heißt Monom, $\sum_{i=1}^n k_i$ ist der Grad des Monoms. Ein Polynom heißt homogen, wenn alle Monome mit Koeffizienten ungleich 0 den gleichen Grad haben. In einem homogenen quadratischen Polynom haben alle Monome Grad 2. Im Vergleich dazu haben etwa im homogenen linearen Gleichungssystem alle Monome Grad 1 und es gibt keinen nicht-trivialen konstanten Term vom Grad 0.

Definition 10.5 Sei s eine symmetrische Bilinearform auf dem K -Vektorraum V . Die Abbildung $q: V \rightarrow K$, $v \mapsto q(v) = s(v, v)$ heißt die zu s gehörige quadratische Form.

Weil s bilinear ist, ist $q(\lambda v) = \lambda^2 q(v)$ ($|\lambda|^2 q(v)$) und für $V = K^n$ mit $A = M(s)$ ist $q(x) = x^\top Ax$ ($x^\top A\bar{x} = (x^\top A\bar{x})^\top = x^* Ax$).

q und s bestimmen sich gegenseitig eindeutig.

Satz 10.6 (Polarisierung) Ist $\text{char}(K) \neq 2$, so gilt für jede symmetrische Bilinearform und die zugehörige quadratische Form q auf V

$$s(v, w) = \frac{1}{2}[q(v+w) - q(v) - q(w)]$$

$$(\text{sesquilinear } s(v, w) = \frac{1}{4}[q(v+w) - q(v-w) + iq(v+iw) - iq(v-iw)]).$$

Beweis: $q(v) = s(v, v)$ einsetzen und ausrechnen. □

Die Bedingung $\text{char}(K) > 2$ wird benötigt, weil im Falle von $\text{char}(K) = 2$ der Ausdruck $2^{-1} = (1+1)^{-1} = 0^{-1}$ nicht definiert ist bzw. für $s(v, w) = s(w, v)$ immer $q(v+w) = q(v) + q(w)$ gilt.

Definition 10.7 Eine symmetrische Bilinearform s auf einem \mathbb{K} -Vektorraum V (mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$) mit

$$s(v, v) > 0 \text{ für alle } 0 \neq v \in V \quad \text{heißt} \quad \underline{\text{positiv definit}},$$

$$s(v, v) < 0 \quad \underline{\text{negativ definit}},$$

$$\text{sonst} \quad \underline{\text{indefinit}}.$$

Bestimme Basis \mathcal{B} über $T_{\mathcal{A}}^{\mathcal{B}} = \overline{P} \cdot \hat{\Lambda}^{-1} \in \text{GL}(n; \mathbb{K})$, dann ist nach der Transformationsformel S 10.2

$$M_{\mathcal{B}}(s) = T_{\mathcal{A}}^{\mathcal{B}\top} M_{\mathcal{A}}(s) \overline{T_{\mathcal{A}}^{\mathcal{B}}} = \hat{\Lambda}^{-1} \underbrace{P^* A P}_{=\Lambda_A} \hat{\Lambda}^{-1} = \begin{bmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{bmatrix}.$$

(b): folgt mit (a) aus S 10.4.

Eindeutigkeit des Tripels: Folgende Annahme wird zum Widerspruch geführt:

Es gebe Basen $\mathcal{A} = (v_1, \dots, v_n)$, $\mathcal{B} = (w_1, \dots, w_n)$ mit $M_{\mathcal{A}}(s) = \begin{bmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{bmatrix}$,
 $M_{\mathcal{B}}(s) = \begin{bmatrix} I_{p'} & & \\ & -I_{q'} & \\ & & 0_{r'} \end{bmatrix}$ und $p < p'$ (analog zeigt man die Fälle $q < q'$ bzw. $r < r'$).

Setze $V_{\mathcal{A}}^+ = \text{span}(v_1, \dots, v_p)$, $V_{\mathcal{B}}^+ = \text{span}(w_1, \dots, w_{p'})$. Weil $\dim V_{\mathcal{A}}^+ = p < p' = \dim V_{\mathcal{B}}^+$, gibt es ein $0 \neq v \in V_{\mathcal{B}}^+ \cap \text{span}(v_{p+1}, \dots, v_n)$ ($v \notin V_{\mathcal{A}}^+$), also $v = \sum_{i=p+1}^n \alpha_i v_i = \sum_{i=1}^{p'} \beta_i w_i$. Für v gilt

$$s(v, v) = \sum_{i=p+1}^n \sum_{j=p+1}^n \alpha_i \bar{\alpha}_j \underbrace{s(v_i, v_j)}_{=M_{\mathcal{A}}(s)_{ij}} = \sum_{i=p+1}^n |\alpha_i|^2 M_{\mathcal{A}}(s)_{ii} \leq 0$$

$$\text{und } s(v, v) = \sum_{i=1}^{p'} |\beta_i|^2 M_{\mathcal{B}}(s)_{ii} > 0. \quad \zeta$$

□

Definition 10.10 Das Tripel (p, q, r) aus S 10.9 heißt Signatur der Metrik s bzw. der symmetrischen/hermiteschen Matrix A .

Korollar 10.11

- Zwei endlichdimensionale Räume mit Metrik sind genau dann isometrisch isomorph, wenn ihre Metriken die gleiche Signatur haben.
- Ist $A \in \mathbb{K}^{n \times n}$ symmetrisch/hermitesch und $C \in \text{GL}(n; \mathbb{K})$, dann haben A und $C^* A C$ mit Vielfachheit gezählt die gleichen Anzahlen positiver und negativer Eigenwerte.

Jede symmetrische Bilinearform im \mathbb{R}^n mit Signatur $(n, 0, 0)$ ist isometrisch isomorph zum Euklidischen Raum \mathbb{R}^n mit kanonischem Skalarprodukt, d. h., wählt man ein anderes Skalarprodukt $\langle x, y \rangle_A = x^\top A y$ mit beliebigem positiv definitem A , gibt es einen Isomorphismus, der $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_A)$ isometrisch (gleiche Winkel und Längen) auf $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{I_n})$ abbildet. Die Matrix $T_{\mathcal{A}}^{\mathcal{B}}$ aus dem Beweis zu S 10.4 zeigt, wie dies geschehen kann. Geometrisch sind also all diese Räume gleich.

Jede Sesquilinearform der Signatur $(n, 0, 0)$ ist isometrisch isomorph zum unitären Raum \mathbb{C}^n mit kanonischem Skalarprodukt.

Der Raum (\mathbb{R}^4, s) mit Signatur $(3, 1, 0)$ heißt Minkowskiraum und hat besondere Bedeutung in der speziellen Relativitätstheorie.

10.3 Quadriken (im \mathbb{R}^n)

Eine Hyperfläche m -ter Ordnung im \mathbb{R}^n ist durch eine Gleichung der Form

$$\sum_{\substack{k_1 + \dots + k_n \leq m \\ k_1, \dots, k_n \in \mathbb{N}_0}} a_{k_1, \dots, k_n} x_1^{k_1} \dots x_n^{k_n} = 0$$

mit Koeffizienten $a_{k_1, \dots, k_n} \in \mathbb{R}$ gegeben. Sie ist die Menge der Punkte im \mathbb{R}^n , die Nullstellen des gegebenen multivariaten Polynoms vom Grad m sind.

Beispiel: Eine Hyperfläche 1. Ordnung ist $\sum_{i=1}^n a_i x_i + a_0 = 0$. Diese Hyperfläche ist für $(a_1, \dots, a_n) \neq 0$ eine Hyperebene bzw. ein $n - 1$ dimensionaler affiner Unterraum. ♡

Quadriken sind Hyperflächen 2. Ordnung.

Definition 10.12 Eine Teilmenge $Q \subseteq K^n$ heißt Quadrik, falls es ein Polynom $p \in K[x_1, \dots, x_n]$ vom Grad 2 gibt mit

$$Q = \{x \in K^n : p(x_1, \dots, x_n) = 0\}.$$

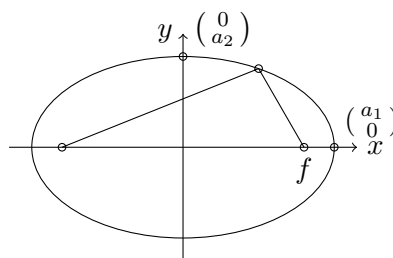
Auf Grund der Polarisierung S10.6 sind im \mathbb{R}^n die homogenen Polynome vom Grad 2 genau die quadratischen Formen $x^\top Ax$ für symmetrische Matrizen A .

Jedes (auch ein inhomogenes) Polynom vom Grad 2 lässt sich in der Form $x^\top Ax + b^\top x + \gamma$ schreiben. Die dazupassende Quadrik ist

$$Q = \{x \in \mathbb{R}^n : x^\top Ax + b^\top x + \gamma = 0\}.$$

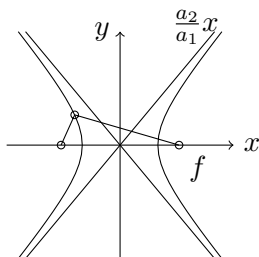
Beispiele:

- $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, b = 0, \gamma = -1: x^2 + y^2 = 1$ beschreibt den Einheitskreis.
- $A = \begin{bmatrix} \frac{1}{a_1^2} & 0 \\ 0 & \frac{1}{a_2^2} \end{bmatrix}, b = 0, \gamma = -1: \frac{x^2}{a_1^2} + \frac{y^2}{a_2^2} = 1$ beschreibt eine Ellipse mit Halbachsen $a_1, a_2 > 0$.



Für $a_1 \geq a_2$ und $f = \sqrt{a_1^2 - a_2^2}$ ist sie die Menge aller Punkte mit $\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} f \\ 0 \end{pmatrix} \| + \| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} -f \\ 0 \end{pmatrix} \| = 2a_1$, für die also die Summe der Abstände zu den Brennpunkten $\begin{pmatrix} f \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -f \\ 0 \end{pmatrix}$ konstant $2a_1$ ist (Ü).

- $A = \begin{bmatrix} \frac{1}{a_1^2} & 0 \\ 0 & -\frac{1}{a_2^2} \end{bmatrix}$, $b = 0$, $\gamma = -1$: $\frac{x^2}{a_1^2} - \frac{y^2}{a_2^2} = 1$ beschreibt eine Hyperbel.

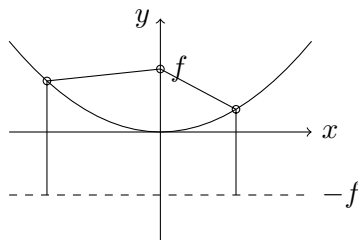


Für $a_1 > 0$ und $f = \sqrt{a_1^2 + a_2^2}$ enthält sie die Punkte mit

$$\left| \left\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} f \\ 0 \end{pmatrix} \right\| - \left\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} -f \\ 0 \end{pmatrix} \right\| \right| = 2a_1,$$

d. h., der Absolutbetrag der Differenz der Abstände zu den Brennpunkten $\begin{pmatrix} f \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -f \\ 0 \end{pmatrix}$ ist konstant $2a_1$ (Ü).

- $A = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\gamma = 0$: $y = ax^2$ beschreibt eine Parabel.



Für $a > 0$ und $f = \frac{1}{4a}$ enthält sie die Punkte mit

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 0 \\ f \end{pmatrix} \right\| = \left\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x \\ -f \end{pmatrix} \right\|.$$

Diese Punkte haben den gleichen Abstand zum Brennpunkt $\begin{pmatrix} 0 \\ f \end{pmatrix}$ und zur Geraden $\{ \begin{pmatrix} x \\ -f \end{pmatrix} : x \in \mathbb{R} \}$ (Ü).

♡

Es wird sich herausstellen, dass – bis auf Spezialfälle – alle Quadriken im \mathbb{R}^2 kongruent zu diesen Grundformen sind, d. h., durch Verschiebung und Rotation/Spiegelung zur Deckung gebracht werden können. Selbst in höheren Dimensionen bleiben diese Grundformen bestimmend.

Definition 10.13 Eine (allgemeine) Abb. $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ heißt Kongruenz (für $\| \cdot \|$), falls $\|v - w\| = \|f(v) - f(w)\|$ für alle $v, w \in \mathbb{K}^n$.

Beobachtung 10.14 Sei V ein n -dimensionaler \mathbb{K} -Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und induzierter Norm $\| \cdot \|$. Ein $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ ist genau dann eine Kongruenz, wenn $F(v) := f(v) - f(0)$, $v \in V$ eine Isometrie (also orthogonal/unitär) ist.

Beweis: Wir zeigen nur den Fall $\mathbb{K} = \mathbb{R}$.

\Rightarrow : Wegen $\|F(v)\| = \|f(v) - f(0)\| = \|v - 0\| = \|v\|$ und B 9.23(a) ist nur zu zeigen, dass F linear ist.

Aus der Polarisierung folgt für das innere Produkt von $u, v \in V$

$$\begin{aligned} 2 \langle F(u), F(v) \rangle &= \|F(u)\|^2 + \|F(v)\|^2 - \|F(u) - F(v)\|^2 \\ &= \|u\|^2 + \|v\|^2 - \|f(u) - f(v)\|^2 \\ &= \|u\|^2 + \|v\|^2 - \|u - v\|^2 = 2 \langle u, v \rangle. \end{aligned}$$

Es genügt, für $w = \alpha u + \beta v$ mit $\alpha, \beta \in \mathbb{R}$ nun $\|F(w) - \alpha F(u) - \beta F(v)\|^2 = 0$ zu zeigen, denn aus der Definitheit der Norm folgt dann $F(w) = \alpha F(u) + \beta F(v)$.

$$\begin{aligned} \|Fw - \alpha Fu + \beta Fv\|^2 &= \|Fw\|^2 - 2 \langle Fw, \alpha Fu + \beta Fv \rangle + \|\alpha Fu + \beta Fv\|^2 \\ &= \|w\|^2 + \alpha^2 \|u\|^2 + \beta^2 \|v\|^2 - 2\alpha \langle w, u \rangle - 2\beta \langle w, v \rangle + 2\alpha\beta \langle u, v \rangle \\ &= \|w - \alpha u - \beta v\|^2 = 0. \end{aligned}$$

\Leftarrow : Ist $F(v) = f(v) - f(0)$ eine Isometrie auf \mathbb{R}^n , also $F \in O(V)$ nach S 9.26/S 9.27, gilt $\|f(v) - f(w)\| = \|F(v) - F(w)\| = \|F(v - w)\| = \|v - w\|$. \square

Im Euklidischen Raum \mathbb{R}^n entsprechen daher mit S 9.31 die Kongruenzabbildungen genau den Abbildungen $f(x) = Px + a$ mit P orthogonal (Drehung/Spiegelung) und $a \in \mathbb{R}^n$ (Translation).

Die Quadriken sollen nun so gedreht und verschoben werden, dass die Koordinatenachsen gerade den Hauptachsen der Ellipsen/Hyperbeln/Parabeln entsprechen.

Satz 10.15 (Hauptachsentransformation)

(a) Sei V ein n -dimensionaler \mathbb{K} -Vektorraum mit Skalarprodukt und $q: V \rightarrow \mathbb{R}$ eine quadratische Form zu einer symmetrischen/hermiteschen Bilinearform. Dann gibt es eine ONB \mathcal{B} (bzgl. des Skalarprodukts) und reelle Zahlen $\lambda_1, \dots, \lambda_n$ mit

$$q(v) = \Phi_{\mathcal{B}}^{-1}(v)^* \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \Phi_{\mathcal{B}}^{-1}(v) = \sum_{i=1}^n \lambda_i |x_i|^2 \text{ für } v \in V, x = \Phi_{\mathcal{B}}^{-1}(v).$$

(b) Sei V ein n -dimensionaler \mathbb{K} -Vektorraum und $q: V \rightarrow \mathbb{R}$ eine quadratische Form (zu einer symmetrischen/hermiteschen Bilinearform). Dann gibt es eine Basis \mathcal{B} in V und $k \leq m \leq n$, sodass für $v \in V, x = \Phi_{\mathcal{B}}^{-1}(v)$

$$q(v) = \Phi_{\mathcal{B}}^{-1}(v)^* \begin{bmatrix} I_k & & \\ & -I_{m-k} & \\ & & 0_r \end{bmatrix} \Phi_{\mathcal{B}}^{-1}(v) = \sum_{i=1}^k |x_i|^2 - \sum_{i=k+1}^m |x_i|^2.$$

Beweis: (a) folgt aus der Spektralzerlegung S9.41, (b) aus dem Trägheitssatz von Sylvester S 10.9. \square

Für die Quadriken soll nun dargestellt werden, wie die Drehungen/Spiegelungen der Hauptachsentransformation zu einer anschaulichen Normalform führen.

Sei $p(x) = x^\top Ax + b^\top x + \gamma$ und die Eigenwertzerlegung von A sei $A = P\Lambda P^\top$ mit $P^\top P = I_n$, $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_n)$ und $\lambda_1 \geq \dots \geq \lambda_k > 0 > \lambda_{k+1} \geq \dots \geq \lambda_m$, $0 = \lambda_{m+1} = \dots = \lambda_n$. O. B. d. A. sei dabei P so gewählt, dass $(P^\top b)_i = 0$ für $i = m + 2, \dots, n$.

Setze $x = P\hat{x}$, $\hat{b} = P^\top b$, dann ist

$$p(x) = p(P\hat{x}) = \hat{x}^\top P^\top P\Lambda P^\top P\hat{x} + b^\top P\hat{x} = \hat{x}\Lambda\hat{x} + \hat{b}^\top \hat{x} + \gamma =: \hat{p}(\hat{x}).$$

P ist Isometrie, also nur eine Kombination von Spiegelungen und Rotationen. P ändert nichts an der relativen geometrischen Lage der Punkte der Quadrik. Der Einfluss der Komponenten \hat{x}_i in \hat{p} vereinfacht sich mit $\hat{p}_i(\hat{x}_i) = \lambda_i \hat{x}_i^2 + \hat{b}_i \hat{x}_i$ zu

$$\hat{p}(\hat{x}) = \sum_{i=1}^m \hat{p}_i(\hat{x}_i) + \hat{b}_{m+1} x_{m+1} + \gamma.$$

Durch die Hauptachsentransformation ist die Wirkung jeder Koordinate \hat{x}_i in \hat{p} unabhängig von den Werten der anderen, \hat{p} wird dadurch separabel (man kann es koordinatenweise untersuchen und verstehen).

Für jede Komponente \hat{x}_i mit $\lambda_i \neq 0$ lässt sich der lineare Term $\hat{b}_i \hat{x}_i$ durch eine Translation (Verschiebung) beseitigen: Für $\hat{x}_i = \tilde{x}_i - \frac{\hat{b}_i}{2\lambda_i}$ wird

$$\hat{p}_i(\tilde{x}_i - \frac{\hat{b}_i}{2\lambda_i}) = \lambda_i \tilde{x}_i^2 - 2\lambda_i \tilde{x}_i \frac{\hat{b}_i}{2\lambda_i} + \frac{\hat{b}_i^2}{4\lambda_i} + \hat{b}_i \tilde{x}_i - \frac{\hat{b}_i^2}{2\lambda_i} = \lambda_i \tilde{x}_i^2 - \frac{\hat{b}_i^2}{4\lambda_i} = \tilde{p}_i(\tilde{x}_i).$$

Setze $\tilde{b}_i = \frac{\hat{b}_i}{2\lambda_i}$ für $i = 1, \dots, m$ und 0 sonst, dann ist für $x = P(\tilde{x} - \tilde{b})$ (Translation mit $-\tilde{b}$ und anschließend Rotation/Spiegelung mit P)

$$p(P(\tilde{x} - \tilde{b})) = \sum_{i=1}^m \lambda_i \tilde{x}_i^2 + \hat{b}_{m+1} \tilde{x}_{m+1} + \tilde{\gamma}.$$

Ist $\hat{b}_{m+1} \neq 0$, lässt sich auch noch $\tilde{\gamma}$ durch eine Translation in \tilde{x}_{m+1} entfernen, aber das wird unübersichtlich. Angenehmer ist es, wenn man quadratische, lineare und konstante Terme in einem gemeinsamen Matrixausdruck behandeln kann. Dazu dient die

Homogenisierung:

$$x^\top Ax + b^\top x + \gamma = \begin{pmatrix} 1 \\ x \end{pmatrix} \underbrace{\begin{bmatrix} \gamma & \frac{1}{2}b^\top \\ \frac{1}{2}b & A \end{bmatrix}}_{=: A'} \underbrace{\begin{pmatrix} 1 \\ x \end{pmatrix}}_{=: x'}.$$

Statt im \mathbb{R}^n arbeitet man im \mathbb{R}^{n+1} mit $x' = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$. Damit ist

$$Q = \{x \in \mathbb{R}^n : x' = \begin{pmatrix} 1 \\ x \end{pmatrix}, x'^T A' x' = 0\}$$

Haben in einem multivariaten Polynom nicht alle Monome den gleichen Grad, wird in der Homogenisierung in jedem Monom die fehlende Potenz k durch eine neue Unbestimmte x_0^k ergänzt, die man gedanklich auf den Wert $x_0 = 1$ setzt. Hier wird damit $x^T A x + b^T x \cdot x_0 + \gamma x_0^2$ ein homogenes multivariates Polynom vom Grad 2.

Um den Zusammenhang zwischen Quadriken und Kegelschnitten zu erkennen, betrachte man die Menge der Nullstellen der quadratischen Form A' ,

$$\mathcal{K} = \{x' \in \mathbb{R}^{n+1} : x'^T A' x' = 0\}.$$

Diese ist geometrisch ein **Doppelkegel**, denn für jedes $x' \in \mathcal{K}$ ist auch die Gerade durch x' und 0 in \mathcal{K} , also für $\lambda \in \mathbb{R}$ stets $\lambda x' \in \mathcal{K}$, weil $(\lambda x')^T A' (\lambda x') = \lambda^2 (x'^T A' x') = 0$. Die Quadrik Q ist die Projektion des Schnitts des Doppelkegels mit der Hyberebene $x'_0 = 1$ auf die ursprünglichen Koordinaten,

$$Q = \{x \in \mathbb{R}^n : \begin{pmatrix} x_0 \\ x \end{pmatrix} \in \mathcal{K} \cap \{x' \in \mathbb{R}^{n+1} : x'_0 = 1\}\}.$$

In homogenisierter Form lassen sich Kongruenzen besonders elegant beschreiben.

Lemma 10.16 Setze $\mathcal{Z}_{n+1} = \{ \begin{bmatrix} 1 & 0 \\ a & P \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)} : P \in O(n), a \in \mathbb{R}^n \}$.

(a) $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist Kongruenz \Leftrightarrow es gibt $Z \in \mathcal{Z}_{n+1}$ mit $\begin{pmatrix} 1 \\ f(v) \end{pmatrix} = Z \begin{pmatrix} 1 \\ v \end{pmatrix}$ für alle $v \in V$.

(b) $(\mathcal{Z}_{n+1}, \cdot)$ ist Untergruppe von $GL(n+1; \mathbb{R}^{n+1})$.

(c) Ist eine Quadrik $Q = \{y \in \mathbb{R}^n : \begin{pmatrix} 1 \\ y \end{pmatrix}^T A' \begin{pmatrix} 1 \\ y \end{pmatrix} = 0\}$ und eine Kongruenz f durch $Z \in \mathcal{Z}_{n+1}$ gegeben, so ist $f(Q)$ ebenfalls eine Quadrik,

$$f(Q) := \{f(y) : y \in Q\} = \{x \in \mathbb{R}^n : \begin{pmatrix} 1 \\ x \end{pmatrix}^T Z^{-T} A' Z^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix} = 0\}.$$

Beweis: (a): Mit $a = f(0)$ und $F = f(v) - f(0)$ ist nach B 10.14 f genau dann Kongruenz, wenn F orthogonal ist, also mit S 9.31 genau dann, wenn $F(v) = P v$ für ein $P \in O(n)$.

(b): $O(n)$ ist Gruppe nach S 9.32. Sind $P, Q \in O(n)$, $a, b \in \mathbb{R}^n$, so sind das Produkt $\begin{bmatrix} 1 & 0 \\ a & P \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ b & Q \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a+Pb & PQ \end{bmatrix}$, das Inverse $\begin{bmatrix} 1 & 0 \\ a & P \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -P^T a & P^T \end{bmatrix}$ und das neutrale Element $\begin{bmatrix} 1 & 0 \\ 0 & I_n \end{bmatrix}$ in \mathcal{Z}_{n+1} .

(c): Für $\begin{pmatrix} 1 \\ x \end{pmatrix} = Z \begin{pmatrix} 1 \\ y \end{pmatrix}$ ist

$$0 = \begin{pmatrix} 1 \\ y \end{pmatrix}^T A' \begin{pmatrix} 1 \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix}^T Z^T Z^{-T} A' Z^{-1} Z \begin{pmatrix} 1 \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix}^T Z^{-T} A' Z^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix}. \quad \square$$

Kongruenztransformationen von Quadriken ergeben daher wieder Quadriken und sind einfach durch Kongruenztransformation von A' mit $Z \in \mathcal{Z}_{n+1}$ darstellbar.

Satz 10.17 (Klassifikation der Quadriken im \mathbb{R}^n bis auf Kongruenz)

Sei $Q = \{x \in \mathbb{R}^n : \begin{pmatrix} 1 \\ x \end{pmatrix}^\top A' \begin{pmatrix} 1 \\ x \end{pmatrix} = 0\}$ eine Quadrik mit $A' = \begin{bmatrix} \gamma & \frac{1}{2}b^\top \\ \frac{1}{2}b & A \end{bmatrix}$. Dann gibt es eine Kongruenz $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, natürliche Zahlen $k \leq m \leq n$ und reelle Zahlen $\beta_i > 0$, $1 \leq i \leq m$, sodass die Quadrik $f(Q)$ durch folgende Gleichung beschrieben wird:

- (a) $\sum_{i=1}^k \frac{x_i^2}{\beta_i^2} - \sum_{i=k+1}^m \frac{x_i^2}{\beta_i^2} = 0$, $k \geq m - k$, falls $\text{Rang } A' = \text{Rang } A$,
- (b) $\sum_{i=1}^k \frac{x_i^2}{\beta_i^2} - \sum_{i=k+1}^m \frac{x_i^2}{\beta_i^2} = 1$, falls $\text{Rang } A' = \text{Rang } A + 1$,
- (c) $\sum_{i=1}^k \frac{x_i^2}{\beta_i^2} - \sum_{i=k+1}^m \frac{x_i^2}{\beta_i^2} = x_{m+1}$, $k \geq m - k$, $m < n$, falls $\text{Rang } A' = \text{Rang } A + 2$.

Beweis: 1. Hauptachsentransformation (S 10.15): Es gibt $P \in O(n)$, sodass $P^\top A P = \Lambda = \text{Diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_1 \geq \dots \geq \lambda_k > 0 > \lambda_{k+1} \geq \dots \geq \lambda_m$ und $\lambda_{m+1} = \dots = \lambda_n = 0$, wobei im Fall $m < n$ und $b \notin \text{span}(P_{\bullet,1}, \dots, P_{\bullet,m})$ mit $P_{\bullet,m+1} = -\frac{b - P_{\bullet,(1,\dots,m)} P_{\bullet,(1,\dots,m)}^\top b}{\|b - P_{\bullet,(1,\dots,m)} P_{\bullet,(1,\dots,m)}^\top b\|}$ die Spalte $m+1$ so gewählt wird, dass $\hat{b} = P^\top b = (\hat{b}_1, \dots, \hat{b}_m, \delta < 0, 0, \dots, 0)^\top$. Dann ist $P' = \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix} \in \mathcal{Z}_{n+1}$ und

$$P'^\top A' P' = \begin{bmatrix} \gamma & \frac{\hat{b}_1}{2} & \dots & \frac{\hat{b}_m}{2} & \frac{\delta}{2} & 0 \\ \frac{\hat{b}_1}{2} & \lambda_1 & & 0 & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots & \\ \frac{\hat{b}_m}{2} & 0 & & \lambda_m & 0 & 0 \\ \frac{\delta}{2} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} =: B'.$$

2. Translation: um $\tilde{b}_i = \begin{cases} -\frac{\hat{b}_i}{2\lambda_i}, & i \leq m, \\ -(\gamma - \sum_{i=1}^m \frac{\hat{b}_i^2}{4\lambda_i^2})/\delta, & i = m+1 \wedge \delta < 0, \\ 0, & \text{sonst,} \end{cases}$ durch

$T' = \begin{bmatrix} 1 & 0 \\ \tilde{b} & I_n \end{bmatrix} \in \mathcal{Z}_{n+1}$. Dann ist

$$T'^\top B' T' = \begin{bmatrix} \tilde{\gamma} & 0 & \dots & 0 & \frac{\delta}{2} & 0 \\ 0 & \lambda_1 & & 0 & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots & \\ 0 & 0 & & \lambda_m & 0 & 0 \\ \frac{\delta}{2} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} =: C' \quad \text{mit} \quad \delta \neq 0 \Rightarrow \tilde{\gamma} = 0.$$

3. Fallunterscheidung:

(a): $\tilde{\gamma} = \delta = 0 \Leftrightarrow \text{Rang } A' = \text{Rang } A = m$. O.E. sei $k \geq m - k$, denn $Q = \{x \in \mathbb{R}^n : \begin{pmatrix} 1 \\ x \end{pmatrix}^\top A' \begin{pmatrix} 1 \\ x \end{pmatrix} = 0 = \begin{pmatrix} 1 \\ x \end{pmatrix}^\top (-A') \begin{pmatrix} 1 \\ x \end{pmatrix}\}$. Setze $\beta_i = \frac{1}{\sqrt{|\lambda_i|}}$, $i = 1, \dots, m$, dann ist $\begin{pmatrix} 1 \\ x \end{pmatrix}^\top C' \begin{pmatrix} 1 \\ x \end{pmatrix} = \sum_{i=1}^k \frac{x_i^2}{\beta_i^2} - \sum_{i=k+1}^m \frac{x_i^2}{\beta_i^2}$.

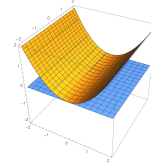
(b): $\tilde{\gamma} \neq 0, \delta = 0 \Leftrightarrow \text{Rang } A' = \text{Rang } A + 1 = m + 1$. Dann liefert $\beta_i = \frac{1}{\sqrt{|\lambda_i|}}$, $i = 1, \dots, m$ die gesuchte Formel.

(c): $\delta \neq 0 \Leftrightarrow \text{Rang } A' = \text{Rang } A + 2 = m + 2$. O. E. sei wieder $k \geq m - k$, nun liefert $\beta_i = \sqrt{\frac{|\delta|}{|\lambda_i|}}$ das gewünschte Ergebnis. \square

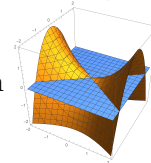
Quadriken im \mathbb{R}^2 (Q entspricht den Nullstellen eines quadratischen Polynoms)

(a) $\text{Rang } A' = \text{Rang } A = m$: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^\top \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$

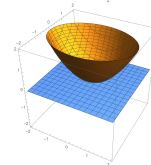
$m = 1, k = 1$: $x_1^2 = 0$ eine Gerade



$m = 2, k = 1$: $\frac{x_1^2}{\beta_1^2} - x_2^2 = 0$ zwei sich schneidende Geraden

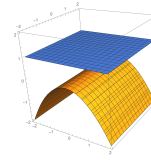


$k = 2$: $\frac{x_1^2}{\beta_1^2} + x_2^2 = 0$ ein Punkt $(0, 0)$

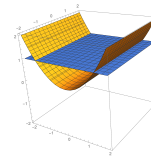


(b) $\text{Rang } A' = \text{Rang } A + 1 = m + 1$: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^\top \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \gamma = 0$

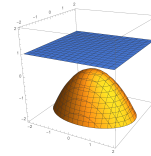
$m = 1, k = 0$: $-\frac{x_1^2}{\beta_1^2} = 1$ leere Menge



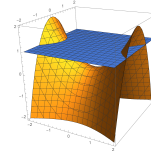
$k = 1$: $\frac{x_1^2}{\beta_1^2} = 1$ zwei parallele Geraden



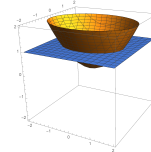
$m = 2, k = 0: -\frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$ leere Menge



$k = 1: \frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$ Hyperbel

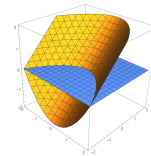


$k = 2: \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} = 1$ Ellipse

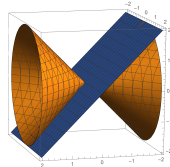


(c) Rang $A' = \text{Rang } A + 2 = m + 2$: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^\top \begin{bmatrix} \lambda_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \delta x_2 = 0$

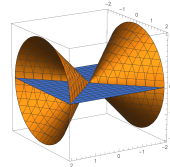
$m = 1, k = 1: \frac{x_1^2}{\beta_1^2} = x_2$ Parabel



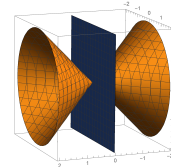
Bis auf die Fälle der leeren Menge und der parallelen Geraden bezeichnet man die Quadriken der Ebene als Kegelschnitte.



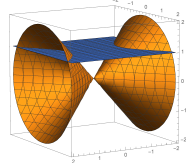
Gerade



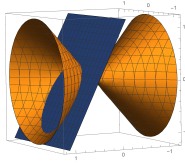
schneidende Geraden



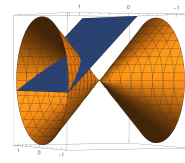
Punkt



Hyperbel



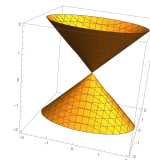
Ellipse



Parabel

Beispiele von Quadriken in \mathbb{R}^3

(a) $m = 3, k = 2: \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} - \frac{x_3^2}{\beta_3^2} = 0$ Ellipsenkegel



(b) $m = 2, k = 1$:	$\frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = 1$	Hyperbelzylinder	
$k = 2$:	$\frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} = 1$	Ellipsenzylinder	
$m = 3, k = 1$:	$\frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} - \frac{x_3^2}{\beta_3^2} = 1$	zweischaliges Hyperboloid	
$k = 2$:	$\frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} - \frac{x_3^2}{\beta_3^2} = 1$	einschaliges Hyperboloid	
$k = 3$:	$\frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} + \frac{x_3^2}{\beta_3^2} = 1$	Ellipsoid	
(c) $m = 1, k = 1$:	$\frac{x_1^2}{\beta_1^2} = x_2$	Parabelzylinder	
$m = 2, k = 1$:	$\frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = x_3$	hyperbolisches Paraboloid	
$k = 2$:	$\frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} = x_3$	elliptisches Paraboloid	

Wenn man neben Translationen nicht nur orthogonale sondern allgemeine lineare Abbildungen zulässt, spricht man von affinen Transformationen.

Definition 10.18 Eine Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt affin, falls es ein $b \in \mathbb{R}^n$ und ein $A \in \mathbb{R}^{n \times n}$ gibt mit $f(x) = Ax + b$ [eine „verschobene“ lineare Abbildung].

Zwei Quadriken $Q_1, Q_2 \subset \mathbb{R}^n$ heißen affin äquivalent, falls es eine bijektive

affine Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ gibt mit $f(Q_1) = Q_2$.

Wie Kongruenzen kann man affine Abbildungen homogenisieren. Für $f(x) = Cx + d$ mit $C \in \text{GL}(n; \mathbb{R})$, $C' = \begin{bmatrix} 1 & 0 \\ d & C \end{bmatrix}$ und $q(x) = x^\top Ax + b^\top x + \gamma = \begin{pmatrix} 1 \\ x \end{pmatrix}^\top \begin{bmatrix} \gamma & \frac{1}{2}b^\top \\ \frac{1}{2}b & A \end{bmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}$ ist

$$f(Q) = \left\{ x \in \mathbb{R}^n : \begin{pmatrix} 1 \\ x \end{pmatrix}^\top C'^{-\top} \begin{bmatrix} \gamma & \frac{1}{2}b^\top \\ \frac{1}{2}b & A \end{bmatrix} C'^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix} = 0 \right\}.$$

Korollar 10.19 Jede Quadrik im \mathbb{R}^n ist affin äquivalent zu einer Quadrik, die durch eine der folgenden Gleichungen beschrieben wird ($k \leq m \leq n$):

- (a) $\sum_{i=1}^k x_i^2 - \sum_{i=k+1}^m x_i^2 = 0$,
- (b) $\sum_{i=1}^k x_i^2 - \sum_{i=k+1}^m x_i^2 = 1$,
- (c) $\sum_{i=1}^k x_i^2 - \sum_{i=k+1}^m x_i^2 = x_{m+1}$, $m < n$.

Bis auf affine Äquivalenz gibt es also nur endlich viele Quadriken in \mathbb{R}^n .

Beweis: Ausgehend von S 10.17 ist

$$q(x) = \begin{pmatrix} 1 \\ x \end{pmatrix}^\top \begin{bmatrix} \tilde{\gamma} & 0 & & 0 & & \frac{\delta}{2} & 0 \\ & \beta_1^2 & & & & 0 & 0 \\ & & \ddots & & & \vdots & \vdots \\ & & & \beta_k^2 & & & \\ & & & & -\beta_{k+1}^2 & & \\ & 0 & & & & \ddots & \\ & & & & & & -\beta_m^2 \\ \frac{\delta}{2} & 0 & & 0 & & 0 & 0 \\ 0 & 0 & & 0 & & 0 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}$$

mit $\begin{pmatrix} \gamma \\ \delta \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -\frac{1}{2} \end{pmatrix} \right\}$, $\beta_i > 0, i = 1, \dots, m$. Setze

$$C' = \begin{bmatrix} 1 & 0 & & 0 \\ & \beta_1^{-1} & & \\ & & \ddots & \\ & & & \beta_m^{-1} \\ & 0 & & & I_{n-m} \end{bmatrix},$$

dann ist $C'^{-\top} A' C'^{-1}$ von der gewünschten Form. □

Index

- Abbildung, 7
- abelsch, 23
- abgeschlossen, 32
- Absolutbetrag, 42
- abzählbar, 9
- abzählbar unendlich, 9
- Addition, 38
- adjungiert, 152
- adjunkte Matrix, 111
- affin, 184
- affin äquivalent, 184
- affin unabhängig, 66
- affine Hülle, 65
- affiner Unterraum, 64
- Affinkombination, 66
- ähnlich, 96
- algebraisch abgeschlossen, 121
- Allquantor, 3
- alternierend
 - h -fach lineare Abb., 107
- alternierende Gruppe, 36
- Annulator, 81
- antisymmetrisch, 16
- äquivalent, 96
- Äquivalenzklasse, 17
- Äquivalenzrelation, 17
- Argument, 42
- Assoziativgesetz, 23
- Assoziativität, 6, 10
- assoziiert, 118
- aufgespannt, 55
- Aussage, 1
- Aussageform, 1
- aussagenlogisch wahr, 2
- Automorphismengruppe, 33
- Automorphismus, 32
- Axiome, 4
- Basis, 59
 - adjungierte, 152
 - biorthogonale, 152
 - duale, 80
 - kanonische/Standard-, 59
 - Orthogonal-, 147
 - Orthonormal-, 147
- Begleitmatrix des Polynoms p , 136
- bijektiv, 8
- Bikonditional, 2
- Bild, 32, 73
- Bild von f , 7
- Bild von x , 7
- Bild(menge), 7
- bilinear, 73
- Bilinearform, 171
 - symmetrische, 171
- Charakteristik, 41
- charakteristische Funktion, 8
- charakteristisches Polynom
 - einer Abbildung, 128
 - einer Matrix, 127
- Definition, 4
- Definitionsbereich, 7
- Determinante, 108
 - des Endomorphismus, 111
- diagonalisierbar, 125
- Diagonalmatrix, 125
- Diagramm
 - kommutierendes, 33
- Differenzmenge, 6
- Dimension, 61
- direkte Summe, 63

- disjunkt, 6, 35
- Disjunktion, 2
- Diskriminante, 121
- Distributivgesetze, 37
- Distributivität, 6
- Doppelkegel, 180
- Dreiecksmatrix
 - (linke) untere, 101
 - (rechte) obere, 101
- Dreiecksungleichung, 42, 145
- duale Abbildung, 84
- duale Basis, 80
- Dualraum, 73

- Ebene, 65
- Eigenraum, 126
 - verallgemeinerter, 141
- Eigenvektor, 125
- Eigenwert, 125
- Einschränkung, 69
- Einselement, 24, 37
- Einsetzungshomomorphismus, 131
- elementare Divisoren, 139
- Elementarmatrizen, 99
- elementarsymmetrischen Funktionen, 122
- Elemente, 5
- endlich, 9
- endlich erzeugt, 59
- endliche Menge, 9
- Endomorphismus, 32
 - positiv (semidefinit), 165
 - positiv definit, 165
 - selbstadjungiert, 161
- Epimorphismus, 32
- Erzeugendensystem, 59
- Erzeuger, 27
- Existenzquantor, 3

- Faktorgruppe, 29
- faktorierte, 100
- Faktormenge, 18
- falsch, 1
- Familie, 11

- Faser, 7
- Faserung, 15
- Fixpunkt, 35
- freie Variable, 44
- führender Koeffizient, 47
- Fundamentalsystem, 104
- Funktion
 - affine, 70

- ganze Zahlen, 5
- gebundene Variable, 44
- gemeinsamer Teiler, 118
- geometrische Vielfachheit, 126
- Gerade, 65
- gleich, 5, 6, 12
- gleichmächtig, 9
- gleichorientiert, 115
- Gleichung
 - homogene, 53
- Gleichungssystem
 - homogenes, 54, 102
 - lineares, 101
- Grad von f , 47
- Gramsche Matrix, 172
- Graph einer Abbildung, 7
- größter gemeinsamer Teiler, 118
- Gruppe, 25
 - allgemeine lineare, 91
 - alternierende, 36
 - Automorphismen-, 33
 - Faktor-, 29
 - orthogonale, 155
 - Quotienten-, 29
 - Restklassen-, 29
 - spezielle orthogonale, 158
 - spezielle unitäre, 158
 - symmetrische, 26
 - unitäre, 155
 - zyklische, 27

- Halbgruppe, 23
- Halbnorm, 145
- Hauptraum, 141
- hermitesch

- Endomorphismus, 161
- Matrix, 161
- Sesquilinearform, 171
- hinreichend, 4
- homogen, 173
- homogene Gleichung, 53
- homogenes Gleichungssystem, 54
- Homomorphismus
 - durch F induzierter bezüglich V/U , 137
 - Gruppen-, 30
 - Ring-, 40
 - Vektorraum-, 68
- Homothetie, 69
- Hyperebene, 65
- Hyperfläche m -ter Ordnung, 176
- idempotent, 150
- identifiziert, 151
- Identität, 8, 39
- imaginäre Einheit, 42
- Imaginärteil, 42
- indefinit
 - symmetrische Bilinearform, 173
- Index, 28
- Indexmenge, 11
- Indirekter Beweis, 2
- induziert, 29
- inhomogen, 102
- injektiv, 8
- inneres Produkt, 143
- Integritätsbereich, 40
- invariant
 - F -invariant, 123
- invarianter Untervektorraum, 123
- inverses Element, 24
- Invertierbar
 - Matrix, 91
- invertierbar, 24, 117
- irreduzibel
 - F -irreduzibel, 139
 - Polynom, 118
- Isometrie, 154
- isometrisch isomorph, 154, 172
- isomorph, 34, 71
- Isomorphismus, 32
- Jordanblock, 140
- Jordanscher Normalform, 140
- kanonische/Standardbasis, 59
- kanonisches Skalarprodukt im \mathbb{R}^n , 144
- kanonisches Skalarprodukt in \mathbb{C}^n , 144
- Kardinalität/Mächtigkeit, 9
- Kartesische Produkt, 6
- Kern, 32, 73
- Klasseneinteilung, 17
- kleiner oder gleich, 9
- kleinstes gemeinsames Vielfaches (kgV), 119
- Koeffizient
 - führender, 47
- Koeffizienten, 14, 43, 47
- Koeffizientenmatrix, 101
- Körper, 41
- kommutativ, 23, 37
- Kommutativgesetz, 23
- Kommutativität, 6
- Komplement, 6
- komplementäre Matrix, 111
- komplexe Zahlen, 42
- Komposition, 9
- Konditional, 2
- Konditionszahl, 169
- kongruent, 172
- Kongruenz, 177
- Kongruenz modulo m , 16
- konjugiert komplexe Zahl, 42
- Konjunktion, 2
- konstant, 47
- Koordinaten, 60
- Koordinatensystem, 71
- Kroneckersymbol, 39
- Kürzungsregel, 48
- l.a., 57
- l.u., 57
- Länge, 59
- linear, 68

- h -fach, 107
- linear abhängig, 57
 - von einer Familie, 58
- linear unabhängig, 57
 - von einer Familie, 58
- lineare Hülle, 54
- linearer Operator, 68
- lineares Funktional, 73
- lineares Gleichungssystem, 43
- Linearfaktor, 119
- Linearform, 73
 - h -fache, 108
- Linearisierung, 70
- Linearkombination, 55
- Linksnebenklasse, 27
- Linkstranslation, 26
- lösbar
 - eindeutig, 104
 - universell, 104
- Lösung
 - triviale, 104
- Lösungsmenge, 43
- Lösungsraum, 102
- Matrix, 13
 - adjungierte, 155
 - adjunkte, 111
 - erweiterte, 102
 - hermitesche, 161
 - invertierbare, 91
 - komplementäre, 111
 - positiv definite, 165
 - positiv semidefinite, 165
 - reguläre, 91
 - singuläre, 91
 - symmetrische, 161
- Matrizelemente, 13
- Matrixnorm, 168
 - Spektral-, 169
- mehrfache Nullstelle, 120
- Menge, 9
- Menge der Permutationen, 10
- Metrik, 145
- Minimalpolynom, 132
 - von F bezüglich U , 133
- Minor
 - h -reihiger, 114
- Modul, 50
- Monoid, 23
- Monom, 173
- Monomorphismus, 32
- multilinear, 107
- Multilinearform, 108
- Multiplikation, 38
- multivariates Polynom, 173
- natürliche Zahlen, 5
- Negation, 2
- negativ definit
 - symmetrische Bilinearform, 173
- neutrales Element,, 23
- nilpotent, 141
- nilpotent mit Nilpotenzindex r , 141
- Nilpotenzindex, 141
- Norm, 145
 - ∞ -Norm, 169
 - 1-Norm, 169
 - 2-Norm, 146
 - Euklidische, 146
 - Frobenius-, 168
 - induzierte, 168
 - least upper bound-, 168
 - Operator-, 168
 - Spektral-, 169
- normal, 160
- normale Untergruppe, 28
- Normalformen, 97
- Normalmatrix
 - allgemeine, 140
- Normalteiler, 28
- normiert, 47, 145
- notwendig, 4
- notwendig und hinreichend, 4
- Nullelement, 24, 37
- Nullmatrix, 39
- Nullpolynom, 47
- Nullstelle, 119
 - Vielfachheit, 120

- nullteilerfrei, 40
- Operation, 21
- Ordnung, 28
- orientierungstreu, 115
- orientierungsuntreu, 115
- orthogonal, 147, 154, 155
- Orthogonalbasis, 147
- orthogonale Gruppe, 155
 - spezielle, 158
- orthogonaler Projektor, 150
- orthogonales Komplement, 149
- Orthogonalsystem, 147
- Orthonormalbasis, 147
- Orthonormalsystem, 147
- Paar, geordnetes, 6
- Partition, 17
- Permutation, 10
- Permutationsgruppe, 26
- Permutationsmatrix, 99
- Polarisierungsidentitäten, 145
- Polynom, 47
 - charakteristisches, 127
 - multivariates, 173
- positiv (semidefinit), 165
- positiv definit, 143
 - Endomorphismus, 165
 - Matrix, 165
 - symmetrische Bilinearform, 173
- positiv semidefinit, 143
 - Endomorphismus, 165
 - Matrix, 165
- Potenzmenge, 6
- Primfaktorardarstellung, 119
- Primfaktorzerlegung, 119
- Projektion, 149
- Punkt, 65
- quadratische, 39
- quadratische Form, 173
- Quadrik, 176
- Quantoren, 3
- Quotientengruppe, 29
- Quotientenmenge, 18
- Quotientenvektorraum von V nach N , 66
- Rang
 - einer Abbildung, 74
 - einer Matrix, 92
 - Spalten-, 92
 - Zeilen-, 92
- Raum mit Metrik, 174
- Realteil, 42
- rechte Seite, 43, 101
- Rechtsnebenklasse, 27
- Rechtstranslation, 26
- reduzibel
 - Polynom, 118
- reduzierender Untervektorraum, 124
- reflexiv, 16
- regulär
 - Matrix, 91
- Relation, 15
 - antisymmetrische, 16
 - reflexive, 16
 - symmetrische, 16
 - transitive, 16
- Repräsentant, 18
- Restklasse, 28
- Restklassen/Kongruenzklassen, 19
- Restklassengruppe, 29
- Rieszsche Abbildung, 151
- Ring, 36
- Ring mit Eins, 37
- Schnitt, 6, 77
- selbstadjungiert, 161
 - Endomorphismus, 161
- semilinear, 143
- separabel, 179
- Sesquilinearform, 144, 171
 - hermitische, 171
- Signatur, 175
- Signum, 11
- singulär
 - Matrix, 91
- Singulärwert, 167

- Skalare, 50
- Skalarprodukt, 143
 - kanonisches im \mathbb{C}^n , 144
 - kanonisches im \mathbb{R}^n , 144
- Spaltenrang, 92
- Spaltensummennorm, 169
- Spektralnorm, 169
- Spektrum von F , 126
- spezielle orthogonale Gruppe, 158
- spezielle unitäre Gruppe, 158
- Spur, 127
- Streichungsmatrix, 112
- submultiplikativ, 168
- Summe, 62
 - direkte, 63
- surjektiv, 8
- symmetrisch, 16
 - h -fach lineare Abb., 107
 - Endomorphismus, 161
 - Matrix, 161
- symmetrische Differenz, 6
- symmetrische Gruppe, 26
- Tautologie, 2
- Teiler, 117
- teilerfremd, 118
- Teilmatrix
 - h -reihige, 114
- Teilmenge, 5
 - echte, 6
- Theorie, 4
- Transformationsmatrix, 93
- transitiv, 16
- Transponierte, 91
- Transposition, 35
- triviale Lösung, 104
- Tupel
 - n -Tupel, 11
 - \mathcal{J} -Tupel, 11
- überabzählbar, 9
- Umkehrabbildung, 8
- Umkehrschluss, 2
- Unbestimmte, 46
- unitär, 154, 155
- unitäre Gruppe, 155
 - spezielle, 158
- universelle Eigenschaft, 78
- Untergruppe, 27
 - von a erzeugte zyklische, 27
- Untermatrix
 - h -reihige, 114
- Unterring, 40
- Untervektorraum, 52
 - F -invarianter, 123
 - F -zyklischer, 133
 - orthogonales Komplement, 149
 - reduzierender, 124
- Urbild, 7
- Urbildmenge, 7
- Vektoren, 50
- Vektorprodukt von a und b , 151
- Vektorraum, 50
 - aufgespannter/erzeugter, 55
 - Euklidischer, 144
 - unitärer, 144
- Vereinigung, 6
- Verknüpfung, 21
- Vielfachheit der Nullstelle, 120
- voller Zeilenrang, 104
- Vorkonditionierung, 170
- wahr, 1
- Wertevorrat/Bildraum, 7
- Winkel, 147
- wohldefiniert, 18
- Wurzel, 119
 - k -te (semi-)definite, 165
- Zahlen
 - ganze, 5
 - komplexe, 42
 - natürliche, 5
- Zeilenrang, 44, 92
 - voller, 104
- Zeilenstufenform, 44
- Zeilensummennorm, 169
- zyklische Gruppe, 27

der Ordnung m , 29
unendliche, 29
Zyklus, 35