Lineare Algebra
Skript* zur Vorlesung LA10

Manuel Bodirsky
Institut für Algebra,
manuel.bodirsky@tu-dresden.de

4. April 2024

*Wintersemester 15/16; aufbauend auf einem handschriftlichen Skript von Reinhard Pöschel, mit Änderungen von Andreas Thom im Wintersemester 16/17. Vollständige Überarbeitung fürs Wintersemester 23/24. Ich freue mich über Emails mit Kommentaren und Verbesserungswünschen.

Inhaltsverzeichnis

1 Mengen, Relationen, Abbildungen **9**

1.1 Mengen . . . 9

1.1.1 Beschreibung von Mengen durch Eigenschaften . . . 10

1.1.2 Mengentheoretische Operationen und Bezeichnungen . . . 10

1.1.3 Rechenregeln . . . 12

1.1.4 Kardinalitäten . . . 12

1.1.5 Das Russellsche Paradoxon . . . 12

1.1.6 Zermelo-Fraenkel Mengenlehre . . . 13

1.2 Relationen . . . 13

1.2.1 Äquivalenzrelationen . . . 14

1.2.2 Abbildungen (Funktionen) . . . 15

1.2.3 Spezielle Eigenschaften von Funktionen . . . 15

1.2.4 Umkehrabbildung . . . 16

1.2.5 Operationen . . . 16

1.2.6 Komposition von Abbildungen . . . 16

1.2.7 Gleichmächtige Mengen . . . 16

1.2.8 Das Auswahlaxiom . . . 17

1.2.9 Die natürlichen Zahlen . . . 18

1.2.10 Restklassen modulo n . . . 19

1.3 Beweisprinzipien . . . 19

1.3.1 Logische Konnektoren . . . 19

1.3.2 Abkürzungen . . . 20

1.3.3 Aussagenlogik . . . 20

1.3.4 Mengengleichheit . . . 20

1.3.5 Vollständige Induktion . . . 21

2 Gruppen, Körper, Vektorräume **23**

2.1 Gruppen . . . 23

2.1.1 Erste Folgerungen . . . 24

2.1.2 Beispiel: die symmetrische Gruppe . . . 24

2.1.3 Untergruppen . . . 25

2.2 Körper . . . 25
2.2.1 Der Körper mit zwei Elementen . . . 26
2.2.2 Weitere endliche Körper . . . 26
2.2.3 Der Körper der komplexen Zahlen . . . 27
2.2.4 Weitere Begriffsbildungen . . . 28
2.3 Vektorräume . . . 28
2.3.1 Beispiele . . . 29
2.3.2 Erste Folgerungen . . . 30
2.3.3 Untervektorräume . . . 31
2.4 Basen und Dimension . . . 32
2.4.1 Linearkombinationen . . . 32
2.4.2 Lineare Unabhängigkeit . . . 33
2.4.3 Basen . . . 34
2.4.4 Austauschsatz . . . 36
2.4.5 Dimension . . . 38

3 Lineare Abbildungen, Gleichungssysteme, Matrizen 43
3.1 Lineare Abbildungen I . . . 43
3.2 Matrizen . . . 44
3.2.1 Matrizenmultiplikation . . . 45
3.2.2 Rang . . . 47
3.2.3 Zeilenumformungen . . . 49
3.2.4 Algorithmus zur Umwandlung einer Matrix in Stufenform . . . 51
3.2.5 Bestimmung von Dimension und Basen . . . 52
3.2.6 Invertierbarkeitskriterium . . . 53
3.2.7 Konstruktion der inversen Matrix . . . 54
3.3 Lineare Gleichungssysteme . . . 56
3.3.1 Definitionen . . . 56
3.3.2 Lösbarkeitskriterium . . . 57
3.3.3 Bild und Kern . . . 59
3.3.4 Der Gaußsche Algorithmus . . . 60
3.3.5 Bestimmung des Kerns . . . 62
3.3.6 Bestimmung des Bilds . . . 63
3.3.7 Unlösbarkeitskriterium . . . 65
3.4 Lineare Abbildungen II . . . 66
3.4.1 Beispiele . . . 66
3.4.2 Beschreibung linearer Abbildungen . . . 67
3.4.3 Kern, Bild, Rang, Defekt . . . 67
3.4.4 Faktorräume . . . 70
3.4.5 Lineare Abbildungen und Matrizen . . . 75
3.4.6 Basiswechsel und Koordinatentransformation . . . 78
3.4.7 Transformationsformel für Matrizen einer linearen Abbildung . . . 79
3.4.8 Äquivalenz von Matrizen . . . 80

3.4.9 Homogene Gleichungssysteme und Untervektorräume 82
3.4.10 Gleichungssysteme und affine Unterräume 83

4 Determinanten, Polynome, Diagonalisierbarkeit 87
4.1 Determinanten 87
4.1.1 Permutationen 87
4.1.2 Determinantenfunktionen 89
4.1.3 Eigeschaften von Determinantenfunktionen 90
4.1.4 Die Leibnizsche Formel 92
4.1.5 Berechnung der Determinante 94
4.1.6 Die Determinante von linearen Abbildungen 99
4.1.7 Lösung linearer Gleichungssysteme mittels Determinanten 99
4.1.8 Invertieren einer Matrix mittels Determinanten 104
4.2 Polynomringe 105
4.2.1 Ringe 105
4.2.2 Polynome über $\mathbb{K}$ 106
4.2.3 Der Polynomring $R[X]$ 108
4.2.4 Der Grad eines Polynoms 108
4.2.5 Polynomfunktionen 108
4.2.6 Der Auswertungshomomorphismus 109
4.2.7 Polynomdivision 110
4.3 Eigenwerte, Eigenvektoren, Diagonalisierbarkeit 111
4.3.1 Anwendung: Pagerank 114
4.3.2 Berechnung von Eigenwerten und das charakteristische Polynom . 115
4.3.3 Diagonalmatrizen 118
4.3.4 Wie diagonalisiert man eine Matrix? 122
4.3.5 Anwendung: Lineares Wachstum 125
4.3.6 Trigonalisierbarkeit 127
4.3.7 Anwendung: Stochastische Matrizen 129

Organisatorisches

Für die Student:innen

- Übungen
- *aktive* Teilnahme.
- Zusammenarbeit zur Lösungsfindung: empfohlen!
- Aufschreiben: jeder selbst.
- Gegenseitige Kontrolle: gerne in kleinen Gruppen!
- In der Vorlesung: mitschreiben!

Zum Skript

- Blau markiert: Kommentare, Hervorhebungen, Literaturverweise.
- Rot markiert: nachträglich geändert oder hinzugefügt, bzw. Hyperlink.
- *Kursiv gedruckt:* Begriff wird definiert, oder sonst herausgehoben.
- Das Symbol $\square$ markiert das Ende eines Beweises und das Symbol $\triangle$ markiert das Ende eines Beispiels.

Literatur

- *Linear Algebra*, von Peter Petersen, Springer Verlag, 2012. Auf Englisch.
- *Lineare Algebra: Eine Einführung für Studienanfänger*, von Gerd Fischer, Springer Verlag, 18te Auflage, 2013.
- *Lineare Algebra: Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen*, von Albrecht Beutelspacher, Vieweg+Teubner Verlag, 6te Auflage, 2003.

Vorbemerkungen

- Großer Fortschritt in der Geometrie: Einführung von Koordinaten.
- Übersicht LA10:
  - Gruppen und Körper
  - Vektorräume und lineare Abbildungen
  - Matrizen, Determinanten, Gleichungssysteme
  - Eigenwerte und Eigenvektoren, Diagonalisierbarkeit
- Zunächst jedoch: grundlegendes zur Sprache der Mathematik.

Kapitel 1

Mengen, Relationen, Abbildungen

Die Mengenlehre ist die Basis der modernen Mathematik. Nahezu alle Teilgebiete der Mathematik lassen sich in der Sprache der Mengenlehre formalisieren. Um die Mengenlehre streng formal aufzubauen, benötigt man Begriffe aus der Prädikatenlogik (auch *Logik erster Stufe* genannt). Dies ist nicht Gegenstand dieser Vorlesung. Da aber die Grundbegriffe der Mengenlehre (wie *Mengen*, *Relationen*, und *Abbildungen*) für die lineare Algebra und für das Mathematikstudium allgemein praktisch sind, beginnen wir die Vorlesung mit einer kurzen informellen Einführung. Auf potentielle Probleme mit der naiven Mengenlehre und die Notwendigkeit eines streng formalen Aufbaus der Mengenlehre kommen wir ebenfalls kurz zu sprechen. Mehr dazu erfährt man aber erst in anderen Vorlesungen (wie z.B. [1]).

1.1 Mengen

Mengen bestehen aus *Elementen*. Schreibweise:

$e \in M$	e ist Element der Menge M
$e \notin M$	e ist *nicht* Element der Menge M

Eine Menge kann beschrieben werden, indem man alle Elemente der Menge angibt. Zum Beispiel bedeutet die Schreibweise

$$M = \{5, 7, 11\}$$

dass M die Menge ist, die genau die Elemente 5, 7 und 11 hat. Mengen selbst können auch Element anderer Mengen sein: beispielsweise ist $\{1, \{2, 3\}\}$ die Menge mit genau den Elementen 1 und $\{2, 3\}$. Sonderfall: die Menge $\{\}$ ohne Elemente heißt *leere Menge* und wird mit $\varnothing$ bezeichnet. Mit der *Pünktchen-Methode* kann man auch unendliche Mengen angeben:

$\mathbb{N} := \{0, 1, 2, 3, \dots\}$	Die Menge der natürlichen Zahlen
$\mathbb{N}_+ := \{1, 2, 3, \dots\}$	Die Menge der positiven natürlichen Zahlen
$\mathbb{Z} := \{0, 1, -1, 2, -2, 3, -3, \dots\}$	Die Menge der ganzen Zahlen

Weitere besondere Mengen, die in dieser Vorlesung von Bedeutung sind:

$$\begin{aligned} \mathbb{Q} &:= \text{ Menge der rationalen Zahlen} \\ \mathbb{R} &:= \text{ Menge der reellen Zahlen} \\ \mathbb{C} &:= \text{ Menge der komplexen Zahlen} \end{aligned}$$

Gleichheit von Mengen: Zwei Mengen A, B sind gleich, wenn sie genau die gleichen Elemente enthalten. Wir schreiben dann $A = B$.

1.1.1 Beschreibung von Mengen durch Eigenschaften

Der Ausdruck

$$\{x \in \mathbb{N} \mid \underbrace{\exists y \in \mathbb{N}:}_{\text{es existiert ein } y} \quad x = 3y + 1\}$$

bezeichnet die Menge aller natürlichen Zahlen, für die ein $y \in \mathbb{N}$ existiert, so dass $x = 3y+1$. Also alle natürlichen Zahlen, die bei Division durch 3 den Rest 1 lassen. Allgemein verwenden wir Ausdrücke der Gestalt

$$\{x \in M \mid x \text{ hat Eigenschaft } E\}$$

für eine Eigenschaft E.

1.1.2 Mengentheoretische Operationen und Bezeichnungen

Die *Enthaltenseinsbeziehung* (*Inklusion*)[1]:

$A \subseteq B$	A ist *Teilmenge* von B
	d.h., jedes Element von A ist auch Element von B
$A \subset B$	A ist *echte* Teilmenge von B:
	$A \subseteq B$ und $A \neq B$

Der Unterschied von $A \subseteq B$ und $A \subset B$ ist also, dass es in letzterem Fall ein Element $x \in B$ gibt mit $x \notin A$. Zum Beispiel:

$$\mathbb{N}_+ \subset \mathbb{N} \subset \mathbb{Z}$$

Durchschnitt (*Schnitt*):

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

[1] In manchen Gebieten der Mathematik, z.B. in der Analysis, wird für $\subseteq$ meist $\subset$ verwendet; für $\subset$ wird dann z.B. $\subsetneq$ verwendet. Die Meinungen gehen hier unversöhnlich auseinander. Über kurz oder lang werden Ihnen noch andere terminologische Konflikte in der Literatur begegnen, es ist also gut, sich bereits früh daran zu gewöhnen. Innerhalb dieser Vorlesung aber werde ich mich um Konsistenz bemühen.

Man sagt, dass A und B *disjunkt* sind falls $A \cap B = \emptyset$.

Vereinigung:

$$A \cup B := \{x \mid x \in A \text{ oder } x \in B\}$$

Verallgemeinerung: sei I eine Menge und für jedes $i \in I$ sei M_i eine Menge. Dann definieren wir

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

Beispiel 1.1.1. Für $i \in \mathbb{N}$ sei $M_i := \{1, 2, \ldots, n\}$; also $M_0 = \emptyset$, $M_1 = \{0\}$, $M_2 = \{0, 1\}$, $M_3 = \{0, 1, 2\}$, und so weiter. Dann ist

$$\bigcup_{i \in \mathbb{N}} M_i = \mathbb{N}$$
$$\bigcap_{i \in \mathbb{N}} M_i = \emptyset. \qquad \triangle$$

Differenz:

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

Komplement: falls klar ist, dass wir Teilmengen einer Menge M betrachten, und $A \subseteq M$, dann steht $\overline{A}$ für $M \setminus A$, das *Komplement* von A (in M). Die Komplementmenge hängt also auch von M ab; da aber M oft aus dem Kontext klar ist, fließt diese Information nicht in die Notation ein.

Potenzmenge:

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

Kartesisches Produkt:

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

Was ist ein Paar (a, b)? Es gelte $(a, b) = (a', b')$ genau dann wenn $a = a'$ und $b = b'$. Die Reihenfolge ist wichtig! Es gilt: $(1, 2) \neq (2, 1)$ aber $\{1, 2\} = \{2, 1\}$.

Als Mengen fassen wir das Paar (a, b) daher (zum Beispiel) als $\{\{a\}, \{a, b\}\}$ auf.
Verallgemeinerung: *n-Tupel* $(a_1, \ldots, a_n)$. Das Element a_i, für $i \in \{1, \ldots, n\}$, wird der *i-te Eintrag des Tupels* genannt.

$$A_1 \times \cdots \times A_n := \big\{(a_1, \ldots, a_n) \mid \forall i \in \{1, \ldots, n\} \text{ gilt } a_i \in A_i\big\}$$

Hier steht "$\forall i \in M$" für: "für alle $i \in M$".
Analog: "$\exists i \in M$" steht für "es gibt (mindestens) ein $i \in M$".

Weitere Abkürzung:

$$A^n := A \times \cdots \times A$$

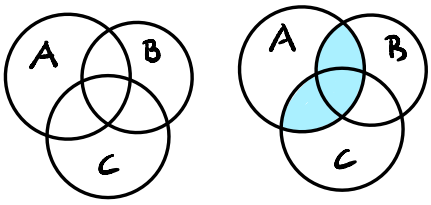
Beispiel: $\mathbb{R}^3$.

1.1.3 Rechenregeln

Es gibt folgende Rechenregeln für Mengenoperationen:

$A \cap A = A$	Schnitt ist *idempotent*
$A \cup A = A$	Vereinigung ist *idempotent*
$A \cap B = B \cap A$	Schnitt ist *kommutativ*
$A \cup B = B \cup A$	Vereinigung ist *kommutativ*
$A \cap (B \cap C) = (A \cap B) \cap C$	Schnitte sind *assoziativ*
$A \cup (B \cup C) = (A \cup B) \cup C$	Vereinigungen sind *assoziativ*
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	Schnitt ist *distributiv* über Vereinigung
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Vereinigung ist *distributiv* über Schnitt

Diese Rechenregeln können besonders einfach mit sogenannten Venn-Diagrammen verdeutlicht werden.



Für die Regel $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ beispielsweise sieht man, dass die Ausdrücke auf beiden Seiten dieselbe farbige Fläche im Diagramm rechts beschreiben.

1.1.4 Kardinalitäten

$|A|$ bezeichnet die Anzahl der Elemente (die *Mächtigkeit*) einer Menge A.

$$|\varnothing| = 0 \qquad |\{\varnothing\}| = 1 \qquad |\{2, 4, 4\}| = 2$$

Es gilt folgendes.

$$\begin{aligned} |\mathcal{P}(A)| &= 2^{|A|} \\ |A \times B| &= |A| \cdot |B| \\ |A \cup B| &= |A| + |B| - |A \cap B| \end{aligned}$$

1.1.5 Das Russellsche Paradoxon

$$M := \{x \mid x \notin x\}$$

Gilt $M \in M$? Gilt $M \notin M$?

Notwendigkeit einer streng formalen, *axiomatischen* Mengenlehre.

1.1.6 Zermelo-Fraenkel Mengenlehre

ZF: weitverbreitete axiomatische Mengenlehre.

1. *Leere Menge:* Es gibt eine leere Menge.
2. *Extensionalität:* Wenn zwei Mengen die gleichen Elemente haben, dann sind sie gleich.
3. *Paarmenge:* Für alle Mengen A und B gibt es eine Menge $\{A, B\}$ mit der Eigenschaft dass $C \in \{A, B\}$ genau dann wenn $C = A$ oder $C = B$.
4. *Vereinigung:* Für alle Mengen M existiert die Menge $\bigcup M$, die gleich der Vereinigung aller Mengen in M ist, soll heissen,
$$\bigcup M := \{x \mid \text{ es gibt ein } e \in M \text{ so dass } x \in e\}.$$
5. *Unendliche Mengen:* Es gibt eine Menge M, die die leere Menge und die Menge $\{e\}$ für jedes $e \in M$ enthält.
6. *Potenzmengen:* Für jede Menge M gibt es eine Menge, die genau alle Teilmengen von M enthält.
7. *Ersetzungsschema:* Informell: Bilder von Mengen unter definierbaren Funktionen sind selbst wieder Mengen; eine Formalisierung des Funktionsbegriffs folgt in Kapitel 1.2.2. Für eine formale Definition des Begriffs *definierbar* verweisen wir auf die Vorlesung *Einführung in die mathematische Logik* [1].
8. *Fundierung:* Jede Menge $M \neq \varnothing$ enthält ein e, so dass $e \cap M = \varnothing$. Insbesondere: Mengen enthalten sich nicht selbst.

1.2 Relationen

Eine (zweistellige, oder binäre) *Relation* R zwischen A und B ist eine Teilmenge von $A \times B$. Schreibweise: statt $(a, b) \in R$ auch $R(a, b)$.

Bemerkung 1.2.1. Praktische Visualisierungen:

- Falls $A \cap B = \varnothing$: Darstellung durch *Graphen*, mit Pfeil von a nach b falls $(a, b) \in R$.
- Spezialfall $A = B$: man spricht von einer *Relation auf* A. Darstellung durch *gerichtete Graphen*: Pfeil von a nach b falls $(a, b) \in R$.

Eine Relation $R \subseteq A^2$ heißt

- *reflexiv* wenn für alle $a \in A$ gilt: $(a, a) \in R$.
- *symmetrisch* wenn für alle $a, b \in A$ gilt: falls $(a, b) \in R$, dann auch $(b, a) \in R$.

- *antisymmetrisch* wenn für alle $a, b \in A$ gilt: falls $(a, b) \in R$ und $(b, a) \in R$, dann ist $a = b$.
- *transitiv* wenn für alle $a, b, c \in A$ mit $(a, b) \in R$ und $(b, c) \in R$ auch gilt $(a, c) \in R$.

Beispiele: '<' ist eine binäre Relation auf $\mathbb{N}$, und ist transitiv, aber nicht reflexiv und nicht symmetrisch. Die binäre Relation $\leq$ auf $\mathbb{N}$, definiert durch $n \leq m$ falls $n < m$ oder $n = m$, ist ebenfalls transitiv, zusätzlich reflexiv, und antisymmetrisch.

1.2.1 Äquivalenzrelationen

Eine Relation $R \subseteq A^2$ ist eine *Äquivalenzrelation (auf A)* falls R reflexiv, symmetrisch, und transitiv ist. Motivation: Verallgemeinerung von Gleichheit. Klassenbildung.

$$[x]_R := \{y \in A \mid R(y, x)\}$$

heißt die *Äquivalenzklasse* von x bezüglich R. Wir schreiben A/R für die Menge aller Äquivalenzklassen von A bezüglich R, die *Faktormenge* von A nach R.

Lemma 1.2.2. [2] *Sei R eine Äquivalenzrelation auf einer Menge A. Dann sind zwei Elemente aus A genau dann äquivalent, wenn sie die gleiche Aquivalenzklasse haben:*

$$R(x, y) \text{ gilt genau dann wenn } [x]_R = [y]_R$$

Beweis. Seien $x, y \in A$ so dass $R(x, y)$. Zeigen zuerst $[x]_R \subseteq [y]_R$. Sei $z \in [x]_R$, d.h. $R(z, x)$. Wegen $R(x, y)$ und Transitivität folgt $R(z, y)$, also $z \in [y]_R$. Zur Inklusion $[y]_R \subseteq [x]_R$: wir haben $R(y, x)$ mit Symmetrie, und verwenden das soeben bewiesene.

Umgekehrt: nehme an, dass $[x]_R = [y]_R$. Da $y \in [y]_R$ wegen Reflexivität gilt also $y \in [x]_R$, und damit $R(x, y)$. □

Definition 1.2.3 (Partition)**.** Eine *Partition* einer Menge A ist eine Menge $\mathcal{P}$ nicht leerer Teilmengen von A die paarweise disjunkt sind und deren Vereinigung gleich A ist.

Man nennt die Elemente von $\mathcal{P}$ die *Klassen* der Partition $\mathcal{P}$.

Proposition 1.2.4 (Äquivalenz und Partition)**.** [3] *Die Faktormenge A/R einer Äquivalenzrelation R auf einer Menge A ist stets eine Partition. Umgekehrt gilt: ist $\mathcal{P}$ eine Partition von A, dann ist $R_{\mathcal{P}} := \bigcup_{A_i \in \mathcal{P}} A_i \times A_i$ eine Äquivalenzrelation. Es gilt $R = R_{A/R}$ und $A/R_{\mathcal{P}} = \mathcal{P}$.*

Übung 1. Beweisen Sie Proposition 1.2.4.

[2] Ein *Lemma* (altgriechisch für „das Angenommene"; Mehrzahl Lemmata) ist eine Hilfsaussage, die praktisch ist in Beweisen von anderen Aussagen. Das konkret vorliegende Lemma zum Beispiel ist beim Beweis von Proposition 1.2.4 weiter unten hilfreich.

[3] Eine *Proposition* bezeichnet in der Mathematik wie das Wort *Satz* eine wahre Aussage, allerdings eine, die vielleicht weniger bedeutend ist, und meist keinen Namen trägt. Die Unterteilung in Satz, Proposition, und Lemma ist bisweilen nicht eindeutig und hängt auch von den Vorlieben der Autor:innen ab.

1.2.2 Abbildungen (Funktionen)

Schreibweise für *Funktion* f von einer Menge A *(Definitionsbereich)* in eine Menge B *(Wertebereich)*:

$$f\colon A \to B$$

Jedem $x \in A$ wird genau ein Element aus B zugeordnet, das mit $f(x)$ bezeichnet wird. Formal ist eine Funktion ein Paar bestehend aus

1. einer Relation $G_f \subseteq A \times B$ — dem *Graph* der Funktion, und
2. dem Wertebereich B,

mit folgenden Eigenschaften:

1. f ist überall auf A definiert: d.h., für alle $a \in A$ gibt es ein $b \in B$ mit $(a, b) \in G_f$.
2. Eindeutigkeit: für alle $a \in A$ und für alle $b, b' \in B$ mit $(a, b) \in G_f$ und $(a, b') \in G_f$ gilt $b = b'$.

Schreibweise: $b = f(a)$ falls $(a, b) \in G_f$. Nennen $f(a)$ das *Bild von a unter f*. Weitere häufige Schreibweise: $x \mapsto f(x)$.

1.2.3 Spezielle Eigenschaften von Funktionen

- $f\colon A \to B$ heißt *surjektiv* falls für alle $b \in B$ (mindestens) ein $a \in A$ existiert mit $(a, b) \in G_f$. In anderen Worten, zu jedem $b \in B$ gibt es einen Pfeil.
- $f\colon A \to B$ heißt *injektiv* falls für alle $a, a' \in A$ und $b \in B$ gilt: falls $f(a) = f(a')$ dann auch $a = a'$. In anderen Worten, zu jedem $b \in B$ gibt es höchstens einen Pfeil.
- $f\colon A \to B$ heißt genau dann *bijektiv* wenn f injektiv und surjektiv ist. In anderen Worten, zu jedem $b \in B$ gibt es *genau* einen Pfeil.

Weitere Bezeichnungen. Sei $f\colon A \to B$ und $A' \subseteq A$. Dann definieren wir das *Bild* von A' unter f als

$$f[A'] := \{f(a) \mid a \in A'\}\,.$$

Die Abbildung $g\colon A' \to B$ definiert durch $f \mapsto f(a)$ heißt *Einschränkung* von f auf A', und wird mit $f|_{A'}$ bezeichnet.

Für $B' \subseteq B$ definieren wir die *Urbildmenge* von B' unter f als

$$f^{-1}[B'] := \{a \in A\colon f(a) \in B'\}$$

Der *Kern* von f ist die folgende Äquivalenzrelation auf A

$$\{(a, a') \in A^2 \mid f(a) = f(a')\}\,.$$

Beispiel 1.2.5. Wir untersuchen einige Beispiele von konkreten Funktionen auf die Eigenschaften *injektiv*, *surjektiv*, und *bijektiv*.

- $f\colon \mathbb{Z} \to \mathbb{N} : x \mapsto x^2$ ist weder injektiv noch surjektiv.
- Die Additionsfunktion $\mathbb{R}\times\mathbb{R} \to \mathbb{R} : (x,y) \mapsto x+y$ ist nicht injektiv, aber surjektiv.
- $\mathrm{id}\colon A \to A : x \mapsto x$ heißt die *identische Funktion* oder *Identität* auf A (ist bijektiv). Bezeichnung häufig id_A.
- Für das direkte Produkt $A\times B$ heißen $\pi_1\colon A\times B \to A : (a,b) \mapsto a$ und $\pi_2\colon A\times B \to B : (a,b) \mapsto b$ *Projektionen* auf ersten beziehungsweise zweiten Faktor. △

1.2.4 Umkehrabbildung

Wenn $f\colon A \to B$ eine Funktion ist, dann definiert $(G_f)^{-1}$ genau dann eine Funktion von B nach A wenn f bijektiv ist. Falls f zumindest injektiv ist, dann definiert $(G_f)^{-1}$ eine Funktion von $f[A]$ nach A. Diese Funktion wird dann die *Umkehrfunktion* von f genannt, und mit f^{-1} bezeichnet.

1.2.5 Operationen

Eine *n-stellige Operation* auf einer Menge A ist eine Abbildung $f\colon A^n \to A$.

Beispiel 1.2.6. Die Addition und Multiplikation von natürlichen Zahlen sind 2-stellige Operationen auf $\mathbb{N}$. △

Beispiel 1.2.7. Für alle Mengen A sind Schnitt und Vereinigung zweistellige Operationen auf $\mathcal{P}(A)$. △

1.2.6 Komposition von Abbildungen

$$f\colon A \to B, \quad g\colon B \to C$$

Durch $h(x) := g(f(x))$ für alle $x \in A$ wird eine Abbildung

$$h\colon A \to C$$

definiert, die *Komposition* (Hintereinanderausführung) von f und g. Bezeichnung: $g \circ f$.

1.2.7 Gleichmächtige Mengen

Mengen A, B heissen *gleichmächtig* (Schreibweise $|A| = |B|$) falls es eine bijektive Abbildung $f\colon A \to B$ gibt.

Schreibweise:

- $|A| \le |B|$ gdw es eine injektive Abbildung $f\colon A \to B$ gibt.
- $|A| < |B|$ falls $|A| \le |B|$ und nicht gilt $|A| = |B|$.

Beispiel 1.2.8. Die Mengen $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{N}\times\mathbb{N}$, $\mathbb{Q}$ sind gleichmächtig (sie sind *abzählbar unendlich*). △

Satz 1.2.9 (Cantor). [4] *Für jede Menge A gilt $|A| < |\mathcal{P}(A)|$.*

Beweis. Ein Widerspruchsbeweis. Angenommen, es gäbe eine Bijektion $f\colon A \to \mathcal{P}(A)$. Sei

$$U := \{x \in A \mid x \notin f(x)\}$$

$U \subseteq A$, $U \in \mathcal{P}(A)$. Da f bijektiv ist, existiert $u \in A$ so dass $f(u) = U$. Entweder $u \in U$ oder $u \notin U$.

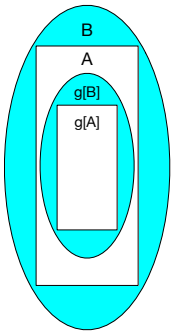
Wäre $u \in U$, so $u \in f(u)$, also $u \notin U$ nach Def. von U, Widerspruch.

Wäre $u \notin U$, so $u \notin f(u)$, also $u \in U$ nach Def. von U, Widerspruch. □

Lemma 1.2.10. *Injektive Abbildungen bilden Vereinigungen auf Vereinigungen ab.*

Satz 1.2.11 (Bernstein-Schröder). *Für alle Mengen A, B gilt: wenn $|A| \leq |B|$ und $|B| \leq |A|$, dann $|A| = |B|$.*

Beweis. Es genügt, den Fall zu betrachten, dass $A \subseteq B$ und dass f die identische Abbildung ist. Definiere $C := \{g^n(x) \mid n \in \mathbb{N}, x \in B \setminus A\}$. Es gilt $B \setminus C \subseteq A$ da $g^0(B \setminus A) = B \setminus A$. Siehe Abbildung.



Sei $h\colon B \to A$ gegeben durch

$$h(x) := \begin{cases} g(x) \text{ falls } x \in C \\ x \text{ falls } x \in B \setminus C. \end{cases}$$

Die Abbildung h ist injektiv:

- Falls $h(x) = h(y) \in C$, dann $g(x) = g(y)$, also $x = y \in C$ wegen der Injektivität von g.
- Falls $h(x) = h(y) \in B \setminus C$ dann gilt $x = h(x) = h(y) = y$.

Die Abbildung h ist auch surjektiv: für jedes $x \in A \cap C$ gibt es ein $y \in C$ mit $x = g(y)$ und für jedes $x \in A \setminus C$ gilt $x = h(x)$. □

1.2.8 Das Auswahlaxiom

Sei $g\colon A \to B$ eine Surjektion. Falls A und B endlich sind, so gibt es auch eine Injektion f von B nach A: denn für jedes $b \in B$ gibt es ein $a \in A$ so dass $g(a) = b$, und wir definieren $f(b) := a$. Wenn A und B unendlich sind, so stellt sich die Frage, ob eine solche Funktion f überhaupt existiert.

Das Auswahlaxiom (AC für englisch *Axiom of choice*) impliziert, dass solche Funktionen existieren (es entspricht aber der mathematischen Praxis, das Auswahlaxiom anzunehmen). Es gibt viele äquivalente Formulierungen des Auswahlaxioms; eine ist die folgende.

[4] Ein *Satz* in der Mathematik ist eine wahre Aussage, die von großer Bedeutung ist, und häufig nach ihrer Entdecker:in benannt wird. Das Wort ‚Theorem' bezeichnet besonders herausstehende Sätze, und wird im Deutschen sehr sparsam verwendet, deutlich seltener jedenfalls als das englische Wort *'theorem'*, was eher dem deutschen Wort 'Satz' entspricht.

(AC) Falls $g\colon A \to B$ eine Surjektion ist, so gibt es auch eine Injektion $f\colon B \to A$ so dass $g \circ f = \mathrm{id}_B$.

Tatsächlich ist bekannt, dass man in ZF die Existenz solcher Funktionen im allgemeinen nicht zeigen kann!

1.2.9 Die natürlichen Zahlen

Der Aufbau der natürlichen Zahlen als Mengen:

$$\begin{aligned} 0 &:= \varnothing \\ 1 &:= \{0\} = \{\varnothing\} \\ 2 &:= \{0,1\} = \{\varnothing, \{\varnothing\}\} \\ 3 &:= \{0,1,2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \\ &\dots \\ n+1 &:= \{0,1,\dots,n\} = \{n\} \cup n \end{aligned}$$

Für $n \in \mathbb{N}$ ist also $n+1$ die Menge, die n und alle Elemente von n enthält.
Vorteil dieser Definition: für alle natürlichen Zahlen m und n gilt:

$$\begin{aligned} m < n &\Leftrightarrow m \subset n \\ &\Leftrightarrow m \in n \end{aligned}$$

Bemerkung 1.2.12. '$<$' und '$\leq$' sind (zweistellige) Relationen auf $\mathbb{N}$:

$$\{(n,m) \in \mathbb{N}^2 \mid n \leq m\}$$

Die Relation $\leq$ auf $\mathbb{N}$ is eine *Wohlordnung*: jede Teilmenge T von $\mathbb{N}$ besitzt ein kleinstes Element. Das heißt, für jedes $T \subseteq \mathbb{N}$ gibt es ein $x \in T$ so dass für alle $y \in T$ gilt $x \leq y$.

Beispiel 1.2.13. Die folgenden Ordnungen sind keine Wohlordnungen:

- Die bekannte Ordnung $\leq$ der ganzen Zahlen $\mathbb{Z}$.
- Die bekannte Ordnung der nicht-negativen rationalen Zahlen

$$\mathbb{Q}_0^+ := \{x \in \mathbb{Q} \mid x \geq 0\}. \qquad \triangle$$

Addition und Multiplikation

Die *Addition* ist induktiv definiert: für alle $n, m \in \mathbb{N}$

$$\begin{aligned} n + 0 &:= n \\ n + (m+1) &:= (n+m) + 1 \end{aligned}$$

Die *Multiplikation* ist induktiv definiert mit Hilfe der Addition: $n, m \in \mathbb{N}$

$$n \cdot 0 := 0$$
$$n \cdot m^{+} := n \cdot m + n$$

Wir definieren auf $\mathbb{N}$ die *Teilbarkeitsrelation*: für $a, b \in \mathbb{N}$ gelte $a|b$ (sprich: *a teilt b*) genau dann wenn es ein $k \in \mathbb{N}$ gibt mit $a \cdot k = b$. Eine Zahl $p \in \mathbb{N}$ heißt *Primzahl* (oder *prim*), wenn sie größer als 1 ist und nur durch 1 und sich selbst teilbar ist.

1.2.10 Restklassen modulo n

Sei $n \in \mathbb{N}_+ = \{1, 2, 3, \dots\}$ und $x, y \in \mathbb{Z}$. Dann ist x ein *Teiler* von y falls ein $z \in \mathbb{Z}$ existiert so dass $y = xz$. Schreiben $x \equiv y \mod n$ falls n ein Teiler von $y - x$. Dadurch wird eine Äquivalenzrelation definiert, nämlich $\{(x, y) \mid x \equiv y \mod n\}$. Menge der Äquivalenzklassen: $\mathbb{Z}/n$ (die *Restklassen modulo n*; auch mit $\mathbb{Z}/(\mathrm{mod}\ n)$ oder $\mathbb{Z}/n\mathbb{Z}$ bezeichnet). Jedes Element $y \in [x]$ wird *Repräsentant* von $[x]$ genannt. Rechnen mit Restklassen ist repräsentantenweise möglich:

- Addition: $[x] + [y] := [x + y]$
- Multiplikation: $[x] \cdot [y] := [x \cdot y]$

Achtung: man muss beweisen, dass dies "wohldefiniert" ist, d.h., nicht von der Auswahl der Repräsentanten abhängt.

1.3 Beweisprinzipien

Was ist ein Beweis? Es gibt ein Gebiet der Mathematik, das sich damit beschäftigt: die *Beweistheorie*. 1930er Jahre: Axiomensysteme und Beweiskalküle, mit denen sich alle wahren mathematischen Aussagen herleiten lassen. Doch das sprengt den Rahmen der Vorlesung.

1.3.1 Logische Konnektoren

Für systematische und formale Definition verweisen wir auf eine Logikvorlesung, wie z.B. [1].

A, B, C, etc. stehen im folgenden für mathematische Aussagen, die entweder *wahr* (1) oder *falsch* (0) sind; man spricht hier auch von *aussagenlogischen Variablen*.

- Schreiben $A \wedge B$ für die Aussage A *und* B ('*Konjunktion*'). Die Aussage $A \wedge B$ ist genau dann wahr, wenn sowohl A als auch B wahr sind.
- Schreiben $A \vee B$ für die Aussage A *oder* B ('*Disjunktion*'). Die Aussage $A \vee B$ ist genau dann wahr, wenn A oder B wahr ist (was den Fall einschließt, dass sowohl A als auch B war sind).

- Schreiben $\neg A$ für die Aussage *nicht* A (‘*Negation*’). Die Aussage $\neg A$ ist genau dann wahr, wenn A nicht wahr ist.

Bemerkung 1.3.1. Die Aussage $\neg(A \wedge B)$ ist genau dann wahr, wenn $\neg A \vee \neg B$ wahr ist. Die Aussage $\neg(A \vee B)$ ist genau dann wahr, wenn $\neg A \wedge \neg B$ wahr ist.

1.3.2 Abkürzungen

Wir schreiben $A \Rightarrow B$ als Abkürzung für $\neg A \vee B$ (‘*Implikation*’).

Bemerkung 1.3.2. $A \Rightarrow B$ gilt genau dann, wenn $\neg B \Rightarrow \neg A$ gilt (‘*Kontraposition*’).

Bemerkung 1.3.3. $A \Rightarrow B$ gilt genau dann, wenn $A \wedge \neg B$ falsch ist (‘*Widerspruchsbeweis*’).

Bemerkung 1.3.4. Falls A gilt, und $B \Rightarrow A$ gilt, so gilt auch B.

Wir schreiben $A \Leftrightarrow B$ als Abkürzung für $(A \Rightarrow B) \wedge (B \Rightarrow A)$ (‘*Äquivalenz*’).

Um zu zeigen, dass die Aussagen $A_1, A_2, \ldots, A_n$ äquivalent sind (d.h., $A_i \Leftrightarrow A_j$ für alle $i, j \in \{1, \ldots, n\}$), genügt es, zu zeigen, dass

$$A_1 \Rightarrow A_2 \wedge A_2 \Rightarrow A_3 \wedge \cdots \wedge A_{n-1} \Rightarrow A_n \wedge A_n \Rightarrow A_1$$

Gute Wahl der Reihenfolge kann Arbeit sparen!

1.3.3 Aussagenlogik

Ein *aussagenlogischer Ausdruck* ist ein Ausdruck, der aus aussagenlogischen Variablen, $\wedge$, $\vee$, $\neg$, und Klammern aufgebaut ist, wie zum Beispiel $A \wedge (B \vee \neg C)$. Eine *Tautologie* ist ein aussagenlogischer Ausdruck, der wahr ist für *alle* Belegungen der aussagenlogischen Variablen mit wahr oder falsch.

Beispiel 1.3.5. Die folgenden aussagenlogischen Aussagen sind Tautologien:

- $A \vee \neg A$.
- $\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B)$ (siehe Bemerkung 1.3.1)
- $\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B)$ (siehe Bemerkung 1.3.1)
- $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ (siehe Bemerkung 1.3.2).
- $A \Rightarrow B \Leftrightarrow \neg(A \wedge \neg B)$ (siehe Bemerkung 1.3.3).
- $(A \wedge (A \Rightarrow B)) \Rightarrow B$ (siehe Bemerkung 1.3.4). △

1.3.4 Mengengleichheit

Um zu zeigen, dass zwei Mengen A und B gleich sind, genügt es zu zeigen, dass

$$A \subseteq B \text{ und } B \subseteq A\,.$$

Bei endlichen Mengen reicht zu zeigen:

$$A \subseteq B \text{ und } |A| = |B|\,.$$

1.3.5 Vollständige Induktion

Es seien $A_0, A_1, A_2, \ldots$ Aussagen. Wir wollen zeigen, dass A_i für alle $i \in \mathbb{N}$ gilt.
Dazu genügt es zu zeigen:

1. *Induktionsanfang*: es gilt A_0.
2. *Induktionsschritt*: für jedes $n \geq 0$ gilt: wenn A_n gilt *(Induktionsvoraussetzung)*, dann auch A_{n+1} *(Induktionsbehauptung)*.

Dann gilt A_i für jedes $i \in \mathbb{N}$ (*Induktionsschluss*).

Beispiel 1.3.6. Aussage A_n:

$$\sum_{i=1}^{n} i = 1 + \cdots + n = \frac{n(n+1)}{2}$$

<u>Induktionsanfang</u> $n = 1$.

$$\sum_{i=1}^{1} i = 1 = \frac{1(1+1)}{2}$$

<u>Induktionsschritt</u>: es gelte A_n, zu zeigen ist A_{n+1}.

$$\begin{aligned}
\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + (n+1) \\
&= \frac{n(n+1)}{2} + \frac{2}{2}(n+1) && \text{(Induktionsvoraussetzung)} \\
&= \frac{n^2 + n + 2n + 2}{2} \\
&= \frac{(n+1)(n+2)}{2}
\end{aligned}$$

△

Bemerkung 1.3.7. Es gibt einen Zusammenhang zwischen dem Prinzip der vollständigen Induktion mit der Aussage, dass $\mathbb{N}$ eine Wohlordnung ist. Dazu betrachten wir die Menge

$$S := \{i \in \mathbb{N} \mid A_i \text{ gilt } \textit{nicht}\}.$$

Angenommen, es stimmt *nicht*, dass $A_0, A_1, A_2, \ldots$ gelten. Dann ist $S \neq \varnothing$ und besitzt daher ein kleinstes Element. Das heisst, es gibt ein $i \in \mathbb{N}$, so dass $A_0, A_1, \ldots, A_{i-1}$ allesamt gelten, aber A_i gilt nicht. Die ist eine Situation, die wir im Induktionsschluss ausschliessen.

Kapitel 2

Gruppen, Körper, Vektorräume

Bekannteste Beispiele für Körper: $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ mit Addition und Multiplikation. Die Definition von Körpern besteht im Wesentlichen aus einer Axiomatisierung der Rechenregeln von Addition und Multiplikation. Motivationen für diese Abstraktion:

- Es gibt eine große Vielfalt an interessanten Körpern. Jede Definition und jeden Satz, den wir allgemein für Körper einführen beziehungsweise beweisen, können wir später auf alle möglichen Körper anwenden.
- Ein angemessener Grad an Abstraktion ist per se eine Errungenschaft, da die Abstraktion dann den Blick lenkt auf das Wesentliche, was wir für den Aufbau der linearen Algebras benötigen.

Jeder Körper ist insbesondere eine Gruppe; wir starten daher mit einem kurzen Abschnitt zu Gruppen.

2.1 Gruppen

Eine Menge G zusammen mit einer 2-stelligen Operation $m\colon G^2 \to G$ heißt *Gruppe*, wenn folgende Axiome erfüllt sind. Wir schreiben $m(x, y) = x \circ y$ der Einfachheit halber.

1. *Assoziativitätsgesetz*: für alle $x, y, z \in G$:
$$x \circ (y \circ z) = (x \circ y) \circ z$$
2. Existenz eines *neutralen* Elements: es gibt ein $e \in G$, so dass für alle $x \in G$ gilt: $x \circ e = x$ und $e \circ x = x$.
3. Existenz *inverser Elemente*: zu jedem $x \in G$ gibt es ein $y \in G$, so dass $x \circ y = e$ und $y \circ x = e$. Schreiben x^{-1} für y.

Eine Gruppe heißt *abelsch*, wenn die Operation $\circ$ zusätzlich das *Kommutativitätsgesetz* erfüllt:

$$\text{für alle } x, y \text{ gilt } x \circ y = y \circ x.$$

Bemerkung 2.1.1. Die genaue Bezeichnung für die Gruppenoperation, das neutrale Element, und das Inverse von x ist nicht von Bedeutung. Weitere Varianten sind: $\cdot$, 1, und x^{-1}, oder $+$, 0, $-x$.

Beispiel 2.1.2. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}; +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$. △

2.1.1 Erste Folgerungen

Lemma 2.1.3. *Das neutrale Element e einer Gruppe ist eindeutig bestimmt.*

Beweis. Seien e, e' zwei neutrale Elemente. Dann ist $e \circ e' = e$ (weil e neutrales Element) und $e \circ e' = e'$ (weil e' neutrales Element). Also $e = e'$. □

Lemma 2.1.4. *Das inverse Element x^{-1} von x ist in einer Gruppe eindeutig festgelegt.*

Beweis. Für Gruppenelemente y_1, y_2 mit

$$\begin{aligned} x \circ y_1 &= y_1 \circ x = e && \text{Voraussetzung 1} \\ x \circ y_2 &= y_2 \circ x = e && \text{Voraussetzung 2} \end{aligned}$$

folgt

$$\begin{aligned} y_1 &= y_1 \circ e && (e \text{ neutrales Element}) \\ &= y_1 \circ (x \circ y_2) && (\text{Voraussetzung 1}) \\ &= (y_1 \circ x) \circ y_2 && (\text{Assoziativität}) \\ &= e \circ y_2 && (\text{Voraussetzung 2}) \\ &= y_2 && (e \text{ neutrales Element}) \end{aligned}$$

□

Folgerung: $(x^{-1})^{-1} = x$.

Übung 2. Ein *links-inverses* Element zu x bezüglich einer 2-stelligen Operation $\circ$ mit neutralem Element e ist ein Element y, so dass $y \circ x = e$. Zeigen Sie, dass man das dritte Gruppenaxiom zur Existenz inverser Element abschwächen kann zur Existenz von Linksinversen. In anderen Worten: falls $(G, \circ)$ das Assoziativitätsgesetz erfüllt, es ein neutrales Element gibt, und zu jedem $x \in G$ ein linksinverses Element in G gibt, dann ist $(G, \circ)$ bereits eine Gruppe.

2.1.2 Beispiel: die symmetrische Gruppe

Sei X eine Menge. Schreiben $\mathrm{Sym}(X)$ für die Menge aller *Permutationen* von X, d.h., Bijektionen zwischen X und X. Dann ist $(\mathrm{Sym}(X), \circ)$ eine Gruppe, wobei $\circ$ die Komposition von Abbildungen ist. Das neutrale Element ist die Identität id_X, und zu $x \in \mathrm{Sym}(X)$ ist die Umkehrabbildung x^{-1} das inverse Element.

2.1.3 Untergruppen

Sei $(G, \circ)$ eine Gruppe, $U \subset G$ eine Teilmenge so dass

- $e \in U$;
- für alle $u \in U$ gilt $u^{-1} \in U$;
- für alle $u, v \in U$ gilt $u \circ v \in U$.

Dann heißt $(U, \circ)$ (genauer: $(U, \circ|_{U \times U})$) eine *Untergruppe* von $(G, \circ)$.

Beispiel 2.1.5. Beispiele zu Untergruppen.

- $\{e\}$ ist stets Untergruppe.
- $(\mathbb{Z}, +)$ ist Untergruppe von $(\mathbb{Q}, +)$.
- $(\mathbb{Q} \setminus \{0\}, \cdot)$ ist Untergruppe von $(\mathbb{R} \setminus \{0\}, \cdot)$. △

Anmerkung: Jede Gruppe G ist eine Untergruppe von $\mathrm{Sym}(G)$ (Beweis kommt später im Studium).

2.2 Körper

Ein *Körper* (englisch *field*; französisch *corps*) ist eine Menge K zusammen mit zwei binären Operationen

$$+ : K \times K \to K \qquad \text{Addition}$$
$$\cdot : K \times K \to K \qquad \text{Multiplikation}$$

die folgende Axiome erfüllen:

1. $(K, +)$ ist eine *abelsche Gruppe* mit neutralem Element 0 *(Nullelement)*, und inversem Element $-x$ zu jedem $x \in K$.
2. Für $(K, \cdot)$ gilt: Multiplikation ist assoziativ, kommutativ, es existiert ein neutrales Element 1 *(Eins-element)*, und für alle $x \in K \setminus \{0\}$ existiert ein inverses Element x^{-1}.
3. $0 \neq 1$.
4. *Distributivgesetz*: für alle $x, y, z \in K$ gilt
$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Häufig wird eine mathematische Struktur und die entsprechende Grundmenge mit dem gleichen Buchstaben bezeichnet, aber in einer anderen Schriftart. Etwa $\mathbb{K}$ für einen Körper und K für die Grundmenge. Häufig wird aber auch das gleiche Symbol für Grundmenge und Körper verwendet. Beispielsweise steht $\mathbb{R}$ sowohl für die Menge der reellen Zahlen als auch für den Körper der reellen Zahlen.

Beispiel 2.2.1. $\mathbb{R} = (\mathbb{R}, +, \cdot)$. △

Beispiel 2.2.2. $\mathbb{Q} = (\mathbb{Q}, +, \cdot)$. △

Kein Beispiel: $(\mathbb{Z}, +, \cdot)$. (Warum?)

Bemerkung 2.2.3. In einem Körper $\mathbb{K} := (K, +, \cdot)$ gilt für alle $x, y \in K$:

- $x \cdot 0 = 0$
- $(-1) \cdot x = -x$
- $(-x) \cdot (-x) = x \cdot x$
- $x \cdot y = 0 \Leftrightarrow x = 0$ oder $y = 0$

2.2.1 Der Körper mit zwei Elementen

Die Menge $\{0, 1\}$ mit folgender Addition und Multiplikation ist ein Körper:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

'Rechnen modulo 2'

- Nullelement ist 0.
- Eins-element ist 1.
- Inverse Elemente bzgl $+$ sind $-0 = 0$ und $-1 = 1$.
- Inverses Element von 1 bezüglich $\cdot$ ist $1^{-1} = 1$.

Bezeichnung für diesen Körper: $\mathrm{GF}(2) = (\{0, 1\}; +, \cdot)$ oder $\mathbb{F}_2$.

2.2.2 Weitere endliche Körper

Sei p eine Primzahl. Dann ist $(\mathbb{Z}/p, +, \cdot)$ ein Körper (mit Addition und Multiplikation wie in Abschnitt 1.2.10).

Bemerkung 2.2.4. $(\mathbb{Z}/n, +, \cdot)$ ist im allgemeinen kein Körper (aber ein *Ring*; Definition 4.2.1).

Bemerkung 2.2.5. Für jede Primzahlpotenz p^m gibt es einen Körper $\mathrm{GF}(p^m)$ mit p^m Elementen.

2.2.3 Der Körper der komplexen Zahlen

Die Gleichung $x^2 = -1$ hat keine reelle Lösung. *Imaginäre Zahlen*: Zahlen, deren Quadrat eine nicht-positive reelle Zahl ist. Mit i bezeichnen wir die imaginäre Zahl mit $i \cdot i = -1$. *Komplexe Zahlen* können in der Form $a + b \cdot i$ mit $a, b \in \mathbb{R}$ dargestellt werden. Hierbei heißt a der *Realteil* und b der *Imaginärteil*.

Formale Definition. Formal definieren wir die komplexen Zahlen mit Hilfe von $\mathbb{R}^2$:

- Addition:
$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} := \begin{pmatrix} a + c \\ b + d \end{pmatrix}$$
- Multiplikation:
$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} c \\ d \end{pmatrix} := \begin{pmatrix} ac - bd \\ ad + bc \end{pmatrix}$$

Schreibweise: schreiben a statt $\binom{a}{0}$ für alle $a \in \mathbb{R}$, und schreiben i statt $\binom{0}{1}$. Dann gilt

$$i \cdot i = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -1$$

Die Menge

$$\mathbb{C} := \{a + b \cdot i \mid a, b \in \mathbb{R}\}$$

bildet zusammen mit der obigen Addition und Multiplikation den *Körper der komplexen Zahlen*.

Nullelement ist 0, denn $\binom{0}{0} + \binom{a}{b} = \binom{a}{b}$.

Eins-element ist 1, denn $\binom{1}{0} \cdot \binom{a}{b} = \binom{a}{b}$.

Geometrische Interpretation. (Komplexe) Gaußsche Zahlenebene: $z = a + bi$ entspricht dem Punkt $\binom{a}{b} \in \mathbb{R}^2$ der Ebene.

Geometrische Interpretation der Multiplikation:

- Multiplikation mit -1:
$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -a \\ -b \end{pmatrix}$$
- Multiplikation mit i:
$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$$
- Multiplikation mit $-i$:
$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ -a \end{pmatrix}$$

2.2.4 Weitere Begriffsbildungen

Die *Charakteristik* $\operatorname{char}(\mathbb{K})$ eines Körpers $\mathbb{K}$ ist die kleinste Zahl $n \in \mathbb{N}^+$, so dass

$$\underbrace{1 + \cdots + 1}_{n \text{ mal}} = 0.$$

Falls eine solche Zahl nicht existiert, so ist $\operatorname{char}(\mathbb{K}) := 0$.

Bemerkung 2.2.6. Algebraische Strukturen, die alle Eigenschaften eines Körpers besitzen, außer dass die Multiplikation notwendigerweise kommutativ ist, heissen *Schiefkörper*. Der Begriff der Charakteristik ist auch für Schiefkörper definiert. Der Satz von Wedderburn besagt, dass jeder Schiefkörper mit endlich vielen Elementen bereits ein Körper ist. Ein Beispiel für einen Schiefkörper der Charakteristik 0, der kein Körper ist, sind die *Quaternionen*.

2.3 Vektorräume

Vektorräume sind das zentrale Thema der linearen Algebra. Sei $\mathbb{K}$ ein Körper mit Einselement 1. Die Elemente von K werden *Skalare* genannt. Ein *Vektorraum* über dem Körper $\mathbb{K}$ (kurz, ein $\mathbb{K}$-Vektorraum) ist eine Menge V zusammen mit zwei Abbildungen

$$V^2 \to V : (u, v) \mapsto u + v$$

(der *Addition*) und

$$K \times V \to V : (\lambda, u) \mapsto \lambda u$$

(der *skalaren Multiplikation*) so dass folgende Axiome erfüllt sind:

1. $(V, +)$ ist abelsche Gruppe mit Nullelement $\mathbf{0}$;
2. Für alle $\lambda, \mu \in K$ und $v \in V$ gilt $(\lambda\mu)v = \lambda(\mu v)$;
3. Für alle $v \in V$ gilt $1v = v$.
4. Für alle $u, v \in V$ und für alle $\lambda \in K$ gilt
$$\lambda(u + v) = \lambda u + \lambda v$$
5. Für alle $v \in V$ und für alle $\lambda, \mu \in K$ gilt
$$(\lambda + \mu)v = \lambda v + \mu v$$

Die Elemente von V heißen *Vektoren*. Wir definieren also zuerst Vektorräume, und dann Vektoren, nicht anders herum.

2.3.1 Beispiele

Der Vektorraum $\mathbb{K}^n$

Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}_+$. Die Menge $\mathbb{K}^n$ aller n-Tupel $(a_1, \ldots, a_n)$ von Elementen $a_1, \ldots, a_n \in K$ bildet einen Vektorraum über K wenn Addition und skalare Multiplikation wie folgt definiert werden:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

und für $\lambda \in K$

$$\lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} \lambda \cdot a_1 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}.$$

Der Nullvektor ist $\mathbf{0} := (0, \ldots, 0)$.

Wichtige Spezialfälle: $\mathbb{R} = \mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$.

Vektorräume durch Körpererweiterungen

Sei $\mathbb{K}$ ein Körper. Eine Teilmenge $U \subseteq K$ heißt *Teilkörper* (*Unterkörper*) wenn gilt

1. $0, 1 \in U$;
2. Für alle $a, b \in U$ ist $a + b \in U$;
3. Für alle $a \in U$ ist $-a \in U$;
4. Für alle $a, b \in U$ ist $a \cdot b \in U$;
5. Für alle $a \in U \setminus \{0\}$ ist $a^{-1} \in U$.

Dann ist U zusammen mit der Einschränkung der Addition und Multiplikation auf U^2 und dem gleichen Null- und Eins-element selbst ein Körper.

Schreibweise:

$$U \leq K$$

Beispiele:

$$\mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$$

Sei $K \leq K'$. Dann ist K' ein Vektorraum über K:

- Addition in K' schon vorhanden;
- Multiplikation von $u \in \mathbb{K}$ mit Skalar $\lambda \in K \subseteq K'$:

$$\lambda u := \lambda \cdot u.$$

Beispiele:

- $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum.
- $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum.
- $\mathbb{R}$ ist ein $\mathbb{R}$-Vektorraum.

Funktionenräume

Sei $\mathbb{K}$ ein Körper und X eine beliebige Menge. Dann bildet die Menge

$$\mathbb{K}^X := \{f \mid f\colon X \to K\}$$

aller Abbildungen von X in K einen $\mathbb{K}$-Vektorraum mit folgenden Operationen:

- Addition $f + g$:
$$(f + g)(x) := f(x) + g(x)$$
- Multiplikation mit Skalar $\lambda \in K$:
$$(\lambda f)(x) := \lambda \cdot f(x)$$

Nullvektor ist die *Nullfunktion*

$$\mathbf{0}\colon X \to K : x \mapsto 0$$

Potenzmenge als $\mathbb{F}_2$-Vektorraum

Die Potenzmenge $\mathcal{P}(A)$ einer beliebigen Menge A wird zu einem Vektorraum über $\mathbb{F}_2$ (siehe Abschnitt 1.1.2), mit folgenden Operationen, für $X, Y \in \mathcal{P}(A)$:

- Addition $X + Y := (X \cup Y) \setminus (X \cap Y)$ (*Symmetrische Differenz*)
- Skalare Multiplikation $0 \cdot X := \varnothing$ und $1 \cdot X := X$.

Nullvektor ist $\mathbf{0} := \varnothing$. Das additiv Inverse von $X \in \mathcal{P}(X)$ ist X selbst, denn

$$X + X = \varnothing = \mathbf{0}.$$

2.3.2 Erste Folgerungen

Lemma 2.3.1. *In einem $\mathbb{K}$-Vektorraum V gilt für alle $u \in V$:*

1. $0u = \mathbf{0}$
2. $(-1)\,u = -u$.

Beweis. Zu Teil 1.

$$\begin{aligned}
\mathbf{0} &= 0u + (-(0u)) && \text{(Vektorraumgesetz 1)}\\
&= (0+0)u + (-(0u)) && \text{(Körpergesetz)}\\
&= (0u + 0u) + (-(0u)) && \text{(Vektorraumgesetz 5)}\\
&= 0u && \text{(Vektorraumgesetz 1)}
\end{aligned}$$

Zu Teil 2.

$$\begin{aligned}
\mathbf{0} &= 0u && \text{(Teil 1)}\\
&= (1-1)u && \text{(Körpergesetz)}\\
&= 1u + (-1)u && \text{(Vektorraumgesetz 5)}\\
&= u + (-1)u && \text{(Vektorraumgesetz 3)}
\end{aligned}$$

Wegen der Eindeutigkeit des inversen Elements (Lemma 2.1.4) folgt $(-1)u = -u$. □

Lemma 2.3.2. *Für alle $\lambda \in K$ und $u \in V$ gilt $\lambda u = \mathbf{0}$ genau dann, wenn $\lambda = 0$ oder $u = \mathbf{0}$.*

Übung 3. Beweisen Sie Lemma 2.3.2.

2.3.3 Untervektorräume

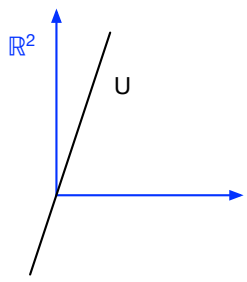
Sei V ein $\mathbb{K}$-Vektorraum.

Definition 2.3.3. Eine Teilmenge $U \subseteq V$ heißt *Untervektorraum* von V, wenn gilt

- $\mathbf{0} \in U$.
- Für alle $u, v \in U$ ist $u + v \in U$.
- Für alle $u \in U$ und $\lambda \in K$ ist $\lambda u \in U$.

Schreibweise:

$$U \leq V$$



Beispiel 2.3.4. Sei V der $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ aus Abschnitt 2.3.1.
Dann ist jede *Gerade durch den Ursprung* $(0,0)$,
also jede Teilmenge von V der Gestalt $\{(x,y) \in V \mid \lambda_1 x + \lambda_2 y = 0\}$,
für ein $\lambda_1, \lambda_2 \in \mathbb{R}$, ein Untervektorraum von V. △

Lemma 2.3.5. *Sei $U \subseteq V$. Dann gilt $U \leq V$ genau dann, wenn*

- *U nichtleer ist, und*
- *U zusammen mit der Addition (wie in V) und der skalaren Multiplikation (wie in V) selbst ein $\mathbb{K}$-Vektorraum ist.*

Beweis. Wenn $U \leq V$, dann gilt $\mathbf{0} \in U$, also $U \neq \varnothing$. Ausserdem ist für alle $u \in U$ auch $-u \in U$, da $-u = (-1) \cdot u \in U$ nach Lemma 2.3.1 und Voraussetzung. Die Einschränkung von $\circ$ auf U^2 und von $\cdot$ auf $K \times U$ liefert somit Funktionen von $U^2 \to U$ beziehungsweise $K \times U \to U$, und es gelten alle Vektorraumaxiome.

Umgekehrt sei $U \neq \varnothing$ so, dass $(U, +|_{U^2}, \cdot|_{K \times U})$ ein Vektorraum ist. Sei $u \in U$. Dann gilt $\mathbf{0} = 0 \cdot u \in U$. Weiterhin sind mit $u, v \in U$ und $\lambda \in K$ auch $u + v \in U$ und $\lambda u \in U$. □

Bemerkung 2.3.6. Der Schnitt von Untervektorräumen eines Vektorraumes ist wieder ein Vektorraum:

$$U_1, U_2 \leq V \implies U_1 \cap U_2 \leq V$$

Dies gilt *nicht* für Vereinigung! Betrachte dazu $V = \mathbb{R}^2$, $u = (1, 0)$, und $v = (0, 1)$. Dann sind $\mathbb{R}u := \{\lambda u \mid \lambda \in \mathbb{R}\}$ und $\mathbb{R}v$ Untervektorräume von V. Aber:

$$(1, 1) = u + v \notin \mathbb{R}u \cup \mathbb{R}v$$

2.4 Basen und Dimension

2.4.1 Linearkombinationen

Sei V ein $\mathbb{K}$-Vektorraum. Seien $v_1, \ldots, v_n \in V$ Vektoren und $\lambda_1, \ldots, \lambda_n \in K$ Skalare. Dann heißt

$$\sum_{i=1}^{n} \lambda_i v_i := \lambda_1 v_1 + \cdots + \lambda_n v_n$$

eine *Linearkombination* von $v_1, \ldots, v_n$. Die Menge aller Linearkombinationen von Vektoren aus $S \subseteq V$ wird mit $\langle S \rangle$ bezeichnet, und die *lineare Hülle* (oder auch: *(linearer) Abschluss*, *(linearer) Spann*, *(lineares) Erzeugnis*) von S genannt:

$$\langle S \rangle := \{\lambda_1 v_1 + \cdots + \lambda_n v_n \mid v_1, \ldots, v_n \in S, \lambda_1, \ldots, \lambda_n \in K, n \in \mathbb{N}\}$$

S darf auch unendlich sein! Legen fest $\langle \varnothing \rangle = \mathbf{0}$.

Vereinbaren außerdem: $\langle v_1, \ldots, v_n \rangle$ steht für $\langle \{v_1, \ldots, v_n\} \rangle$.

Bemerkung 2.4.1. Die Abbildung

$$\mathcal{P}(V) \to \mathcal{P}(V) : W \mapsto \langle W \rangle$$

ist ein *Hüllenoperator*, d.h., es gelten für alle $X, Y \subseteq V$:

- $X \subseteq \langle X \rangle$.
- $X \subseteq Y \implies \langle X \rangle \subseteq \langle Y \rangle$.
- $\langle \langle X \rangle \rangle = \langle X \rangle$.

Proposition 2.4.2. *$\langle v_1, \ldots, v_n \rangle$ ist ein Untervektorraum von V, und zwar der kleinste Untervektorraum von V, der $v_1, \ldots, v_n$ enthählt.*

Beweis. Sei $U := \langle v_1, \dots, v_n \rangle$.

1. $\mathbf{0} = 0 \cdot v_1 + \dots + 0 \cdot v_n \in U$.
2. Seien $u, v \in U$, d.h., $u = \lambda_1 v_1 + \dots + \lambda_n v_n$ und $v = \mu_1 v_1 + \dots + \mu_n v_n$. Dann ist $u + v = (\lambda_1 + \mu_1) v_1 + \dots (\lambda_n + \mu_n) v_n \in U$.
3. Seien $u \in U$, $\lambda \in K$, $u = \lambda_1 v_1 + \dots + \lambda_n v_n$. Dann ist $\lambda u = (\lambda\lambda_1) v_1 + \dots + (\lambda\lambda_n) v_n \in U$.

Also gilt $U \leq V$. Ist $v_1, \dots, v_n \in W$ für Untervektorraum $W \leq V$, so gehören auch alle Linearkombinationen von $v_1, \dots, v_n$ zu W, also $\langle v_1, \dots, v_n \rangle \leq W$. Daher ist $\langle v_1, \dots, v_n \rangle$ der *kleinste* Untervektorraum von V, der $v_1, \dots, v_n$ enthält. □

Man nennt $U = \langle v_1, \dots, v_n \rangle$ auch den *von $v_1, \dots, v_n$ erzeugten (aufgespannten) Vektorraum.* Die Menge $\{v_1, \dots, v_n\}$ heißt dann *Erzeugendensystem* von U.

2.4.2 Lineare Unabhängigkeit

Ein n-Tupel $(v_1, \dots, v_n) \in V^n$ heißt *linear unabhängig* falls gilt

$$\lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0} \Rightarrow \lambda_1 = \dots = \lambda_n = 0 \ .$$

Ansonsten: $(v_1, \dots, v_n)$ *linear abhängig.* Eine Menge $U = \{v_1, \dots, v_n\} \subseteq V$ heißt *linear unabhängig* wenn jedes n-Tupel $(v_1, \dots, v_n)$ mit paarweise verschiedenen Elementen aus U linear unabhängig ist. Ansonsten: U heißt *linear abhängig.*

Bemerkung 2.4.3. Ein einzelner Vektor $v \in V$ ist genau dann linear abhängig, wenn $v = \mathbf{0}$. Ein Tupel $(v_1, \dots, v_n)$ ist genau dann linear abhängig, wenn mindestens ein Vektor v_i Linearkombination der anderen ist:

$$v_i = \sum_{j \neq i} \lambda_j v_j .$$

Bemerkung 2.4.4. Jede Obermenge einer linear abhängigen Menge ist linear abhängig. Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

Beispiel 2.4.5. In $V = \mathbb{R}^2$ (als $\mathbb{R}$-Vektorraum):

- $v_1 = \binom{1}{0}$ und $v_2 = \binom{0}{1}$ sind linear unabhängig: denn $\lambda_1 v_1 + \lambda_2 v_2 = \binom{\lambda_1}{\lambda_2} = \mathbf{0}$ genau dann wenn $\lambda_1 = \lambda_2 = 0$.
- $v_1 = \binom{1}{2}$ und $v_2 = \binom{2}{4}$ sind linear abhängig, denn $v_2 = 2v_1$.
- $v_1 = \binom{1}{2}$ und $v_2 = \binom{\pi}{2\pi}$ sind linear abhängig (aber in $\mathbb{R}^2$ aufgefasst als $\mathbb{Q}$-Vektorraum sind v_1 und v_2 linear unabhängig, da $\pi \notin \mathbb{Q}$).
- Je drei Vektoren $v_1, v_2, v_3 \in \mathbb{R}^2$ sind linear abhängig.

$$\lambda_1 \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \lambda_2 \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} + \lambda_3 \begin{pmatrix} a_3 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0}.$$

Anders geschrieben,

$$\begin{aligned} a_1\lambda_1 + a_2\lambda_2 + a_3\lambda_3 &= 0 \\ \text{und } b_1\lambda_1 + b_2\lambda_2 + b_3\lambda_3 &= 0 \end{aligned}$$

hat für alle $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbb{R}$ *nichttriviale* (d.h., von $(0,0,0)$ verschiedene) Lösung für $\lambda_1, \lambda_2, \lambda_3$. △

2.4.3 Basen

Eine Teilmenge $B \subseteq V$ heißt *Basis* von V wenn

1. B linear unabhängig, und
2. $\langle B \rangle = V$.

Für Basis $B = \{v_1, \dots, v_n\}$, v_i paarweise verschieden, nennen wir $B = (v_1, \dots, v_n)$ *geordnete Basis* (oder auch kurz *Basis*).

Satz 2.4.6 (Eindeutigkeit der Koordinaten)**.** *Ist $B = (v_1, \dots, v_n)$ geordnete Basis von V, so gibt es für jeden Vektor $u \in V$* genau ein *n-Tupel $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$, so dass*

$$u = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

D.h., jedes Element u lässt sich *eindeutig* als Linearkombination von Basiselementen beschreiben. Das n-Tupel $(\lambda_1, \dots, \lambda_n)$ heißt *Koordinatenvektor* von u bezüglich B. Die Abbildung

$$\Phi_B \colon \mathbb{K}^n \to V : (\lambda_1, \dots, \lambda_n) \mapsto \lambda_1 v_1 + \cdots + \lambda_n v_n$$

ist bijektiv und heißt *kanonischer Basisisomorphismus.*[1]

Beweis von Satz 2.4.6. Da B eine Basis ist, gilt insbesondere $\langle B \rangle = V$ und jedes $u \in V$ lässt sich schreiben als $u = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Wenn nun gilt

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = \lambda_1' v_1 + \cdots + \lambda_n' v_n$$

so folgt

$$(\lambda_1 - \lambda_1')v_1 + \cdots + (\lambda_n - \lambda_n')v_n = \mathbf{0}\,.$$

Da $v_1, \dots, v_n$ linear unabhängig sind, so folgt $\lambda_1 - \lambda_1' = 0, \dots, \lambda_n - \lambda_n' = 0$, also $\lambda_1 = \lambda_1'$, $\dots$, $\lambda_n = \lambda_n'$. □

Beispiel 2.4.7. 1. Betrachten $B := \{e_1, e_2\}$ mit $e_1 := \binom{1}{0}$ und $e_2 := \binom{0}{1}$ ist Basis für $\mathbb{R}^2$. Zum Beispiel $u = \binom{8}{3} \in \mathbb{R}^2$ kann geschrieben werden als $u = 8 \cdot e_1 + 3 \cdot e_2$. Die Abbildung Φ_B ist die Identität auf $\mathbb{K}^2$.

[1] Der Begriff *'Isomorphismus'* wird für Vektorräume in Abschnitt 3.1 eingeführt werden).

2. Die beiden Vektoren $v_1 = \binom{1}{0}$, $v_2 = \binom{1}{1}$ bilden ebenfalls eine Basis für $\mathbb{R}^2$. Haben $\binom{8}{3} = 5 \cdot v_1 + 3 \cdot v_2$. △

Bemerkung 2.4.8. Für beliebigen Körper $\mathbb{K}$ ist

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

Basis des $\mathbb{K}$-Vektorraums $\mathbb{K}^n$. Das n-Tupel $(e_1, \ldots, e_n)$ heißt *Standardbasis von* $\mathbb{K}^n$.

Satz 2.4.9 (Charakterisierungssatz für Basen)**.** *Sei V ein $\mathbb{K}$-Vektorraum und $B \subseteq V$.*

- *(Basisergänzungssatz) B ist genau dann Basis von V, wenn B eine* maximale *linear unabhängige Menge ist (d.h., jede echte Obermenge von B ist linear abhängig).*

- *(Basisauswahlsatz) B ist genau dann Basis von V, wenn B ein* minimales *Erzeugendensystem von V ist (d.h., keine echte Teilmenge von B erzeugt V).*

Beweis. Basisergänzungssatz: '$\Rightarrow$' Sei B Basis. Angenommen es gibt ein $v \in V \setminus B$ mit $B \cup \{v\}$ linear unabhängig. Da $\langle B \rangle = V$ gibt es $\lambda_1, \ldots, \lambda_n \in K$ und $v_1, \ldots, v_n \in B$ mit $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, im Widerspruch zur linearen Unabhängigkeit von $B \cup \{v\}$.

'$\Leftarrow$' Sei B maximale linear unabhängige Menge. Wäre B keine Basis, so gäbe es ein $v \in V$ mit $v \notin \langle B \rangle$. Behauptung: Dann wäre $B' := B \cup \{v\}$ linear unabhängig (im Widerspruch zur Maximalität von B).

Wäre B' linear abhängig, so gäbe es $\lambda_0, \lambda_1, \ldots, \lambda_n \in K$ und $v_1, \ldots, v_n \in B$ so dass $\lambda_0 v + \lambda_1 v_1 + \cdots + \lambda_n b_n = \mathbf{0}$, aber $(\lambda_0, \lambda_1, \ldots, \lambda_n) \neq (0, 0, \ldots, 0)$. Falls $\lambda_0 = 0$ so sind $b_1, \ldots, b_n$ linear abhängig, Widerspruch. Falls $\lambda_0 \neq 0$ so ist $v = \lambda_0^{-1}(-\lambda_1 b_1 - \cdots - \lambda_n b_n) \in \langle B \rangle$ im Widerspruch zu $v \notin \langle B \rangle$. □

Übung 4. Beweisen Sie den Basisauswahlsatz (Satz 2.4.9).

Satz 2.4.10 (Satz über die Existenz von Basen)**.** *Jeder Vektorraum hat eine Basis.*

Beweis für endlich erzeugte Vektorräume. Angenommen $V = \langle v_1, \ldots, v_n \rangle$. Falls die Menge $M = \{v_1, \ldots, v_n\}$ linear unabhängig ist, dann ist $\{v_1, \ldots, v_n\}$ eine Basis. Ansonsten lässt sich ein Element $v \in M$ schreiben als Linearkombination der anderen. Sei $M' := M \setminus \{v\}$. Es gilt $\langle M' \rangle = V$. Starte das Verfahren mit der kleineren Menge M' anstatt von M von vorne. Nach endlich vielen Schritten muss das Verfahren abbrechen, und wir haben eine Basis gefunden.

Der Beweis für unendlich dimensionale Vektorräume erfordert das Auswahlaxiom (viele Beweise verwenden hier das *Zornsche Lemma*, welches äquivalent ist zum Auswahlaxiom) was wir uns für's nächste Semester aufheben. □

2.4.4 Austauschsatz

Bemerkung: Sei K *endlicher* Körper, $|K| = q \in \mathbb{N}$. Hat ein $\mathbb{K}$-Vektorraum V eine Basis mit n Elementen, so folgt aus Satz 2.4.6 (Eindeutigkeit der Koordinaten) dass $|V| = q^n$. Also hat *jede* Basis von V genau n Elemente.

Lemma 2.4.11 (Austauschlemma)**.** *Sei $B = (v_1, \dots, v_n)$ eine Basis eines Vektorraums V und $w = \lambda_1 v_1 + \dots + \lambda_n v_n \in V \setminus \{\mathbf{0}\}$ beliebig, und sei $\lambda_j \neq 0$ für ein $j \in \{1, \dots, n\}$. Dann ist $B' = (v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_n)$ ebenfalls eine Basis.*

'Austausch' von v_j gegen w.

Beweis. 1ter Teil: B' ist Erzeugendensystem von V.

$$v_j = \lambda_j^{-1} w - \lambda_j^{-1} \sum_{i \neq j} \lambda_i v_i$$

also $B \subseteq \left\langle B' \right\rangle$ und

$$V = \langle B \rangle \subseteq \left\langle \left\langle B' \right\rangle \right\rangle = \left\langle B' \right\rangle \subseteq V$$

also $V = \left\langle B' \right\rangle$.

2ter Teil: B' ist linear unabhängig. Ansonsten gäbe es nichttriviale Linearkombination

$$\mu_1 v_1 + \dots + \mu_j w + \dots + \mu_n v_n = \mathbf{0}$$

Falls $\mu_j = 0$ dann sind $v_1, \dots, v_n$ linear abhängig, Widerspruch zur Annahme dass B Basis. Also $\mu_j \neq 0$:

$$w = (-\mu_j^{-1}\mu_1)v_1 + \dots + (-\mu_j^{-1}\mu_{j-1})v_{j-1} + 0 \cdot v_j + (-\mu_j^{-1}\mu_{j+1})v_{j+1} + \dots + (-\mu_j^{-1}\mu_n)v_n$$

Andererseits

$$w = \lambda_1 v_1 + \dots + \lambda_j v_j + \dots + \lambda_n v_n$$

Für die geordnete Basis $(v_1, \dots, v_n)$ ergibt sich mit Satz 2.4.6 (Eindeutigkeit der Koordinaten) dass $\lambda_j = 0$, Widerspruch. □

Beispiel 2.4.12. $V = \mathbb{R}^3$ hat die folgende Basis:

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Sei $w := v_2 - v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. Nach Austauschlemma sind $(w, v_2, v_3), (v_1, w, v_3)$ Basen (nicht aber (v_1, v_2, w)). △

Übung 5. Zeigen Sie die Behauptungen in Beispiel 2.4.12.

Satz 2.4.13 (Austauschsatz von Steinitz)**.** *Es sei $B = \{v_1, \dots, v_n\}$ eine Basis eines $\mathbb{K}$-Vektorraums V, und $C = \{w_1, \dots, w_m\}$ sei beliebige Menge linear unabhängiger Vektoren. Dann gilt*

(*a*) $|C| \leq |B|$, *d.h.*, $m \leq n$:
'Jede linear unabhängige Menge besteht aus höchstens n Elementen'.

(*b*) *Durch Hinzunahme von $n - m$ geeignet gewählten Vektoren aus B kann man C zu einer Basis von V ergänzen.*

Beweis. Beweis von (a) und (b) durch Induktion über m.

Induktionsanfang $m = 1$: $w_1 \neq \mathbf{0}$, d.h., V hat mindestens ein Element ungleich $\mathbf{0}$. Also muss auch gelten $|B| \geq 1 = m$. Aussage (b) folgt direkt aus dem Austauschlemma, Lemma 2.4.11.

Induktionsschritt $m > 1$: Die Aussagen (a) und (b) seien richtig für $m - 1$ (Induktionsvoraussetzung). Zu zeigen: (a) und (b) gelten auch für m. Sei $C' := \{w_1, \dots, w_{m-1}\} \subset C$ (ist linear unabhängig). Nach Induktionsvoraussetzung gelten

(a') $m - 1 \leq n$, und

(b') es gibt $v_m, \dots, v_n \in B$ so dass $B' := \{w_1, \dots, w_{m-1}, v_m, \dots, v_n\}$ Basis von V.

(a) gilt für m:
1. Fall: $n > m - 1$. Dann ist $n \geq m$, fertig.
2. Fall: $n \leq m - 1$ (also $n = m - 1$). Nach (b') ist $\{w_1, \dots, w_{m-1}\}$ Basis von V, also ist $\{w_1, \dots, w_{m-1}, w_m\}$ linear abhängig, Widerspruch zur Voraussetzung.

(b) gilt für m: B' Basis. $w_m \in \langle B' \rangle$.

$$w_m = \lambda_1 w_1 + \cdots \lambda_{m-1} w_{m-1} + \lambda_m v_m + \cdots + \lambda_n v_n$$

Wäre $\lambda_i = 0$ für alle $i \geq m$, so wäre $w_m = \sum_{i=1}^{m-1} \lambda_i w_i$ im Widerspruch zur linearen Unabhängigkeit von $C = \{w_1, \dots, w_m\}$. Also gibt es $j \in \{m, \dots, n\}$ mit $\lambda_j \neq 0$. Nach Austauschlemma (Lemma 2.4.11) ist

$$\{w_1, \dots, w_{m-1}, v_m, \dots, v_{j-1}, w_m, v_{j+1}, \dots, v_n\}$$

Basis von V, also ist (b) erfüllt. Nach Induktion gelten (a) und (b) für alle $m \in \mathbb{N}$. □

Bemerkung 2.4.14. Sei V ein $\mathbb{K}$-Vektorraum mit Basis $\{v_1, \dots, v_n\}$.

1. Alle Basen von V haben gleiche Mächtigkeit (nämlich n).
   Beweis: B_1, B_2 Basen. Dann gilt nach Satz 2.4.13 (1) $|B_1| \leq |B_2|$ und analog $|B_2| \leq |B_1|$.

2. Jede linear unabhängige Menge C mit n Elementen ist eine Basis.
   Beweis: Nach Satz 2.4.13 (2), denn für $m = n$ gibt es nichts zu ergänzen.

3. Ist $U \leq V$ Untervektorraum, so hat jede Basis von U höchstens n Elemente und kann stets zu einer Basis von V ergänzt werden.
   Beweis: Basis von U ist linear unabhängige Menge $C \subseteq V$, kann nach Satz 2.4.13 (2) ergänzt werden.

2.4.5 Dimension

Dimension: "Die Anzahl der Freiheitsgrade in einem mathematischen Raum"

Definition 2.4.15. Sei V ein $\mathbb{K}$-Vektorraum und $B = \{v_1, \dots, v_n\}$ eine Basis. Dann nennt man n die *Dimension* von V.

$$\dim_{\mathbb{K}} V := n$$

Falls aus dem Kontext klar ist, welches der zu Grunde liegende Körper ist, oder der Körper keine Rolle spielt, so schreiben wir auch $\dim V$ anstatt $\dim_{\mathbb{K}} V$. Für $V = \{\mathbf{0}\}$ ist $\dim V = 0$ ($B = \varnothing$). Falls keine endliche Basis existiert, schreiben wir $\dim V = \infty$. Definition 2.4.15 hängt wegen Satz 2.4.13 nicht von der Auswahl der Basis ab.

Bemerkung 2.4.16. Seien $U_1 \leq U_2 \leq V$ Untervektorräume. Dann gilt

$$\dim U_1 \leq \dim U_2$$

sowie

$$\dim U_1 = \dim U_2 \Leftrightarrow U_1 = U_2$$

(siehe Bemerkung 2.4.14 (3)).

Definition 2.4.17. Sind $U_1, U_2 \leq V$ Untervektorräume von V, so heißt

$$U_1 + U_2 := \{u + v \mid u \in U_1, v \in U_2\}$$

die *Summe* von U_1 und U_2. Gilt zusätzlich $U_1 \cap U_2 = \{\mathbf{0}\}$, so schreibt man

$$U_1 \oplus U_2$$

und spricht von der *direkten Summe*. Falls $U_1 \oplus U_2 = V$, so heißt U_2 *Komplement* von U_1 (in V). Wir sagen auch, U_1 und U_2 sind *komplementär*.

Satz 2.4.18. *Seien $U_1, U_2 \leq V$. Dann gilt: $U_1 + U_2 \leq V$. Mehr noch: $U_1 + U_2$ ist der kleinste Untervektorraum von V, der U_1 und U_2 enthält, d.h.,*

$$U_1 + U_2 = \langle U_1 \cup U_2 \rangle$$

Beweis. $U_1 + U_2 \subseteq \langle U_1 \cup U_2 \rangle$: klar.
$U_1 \subseteq U_1 + U_2$, $U_2 \subseteq U_1 + U_2$: klar.
Nach Proposition 2.4.2 ist $\langle U_1 \cup U_2 \rangle$ der kleinste Untervektorraum von V, der $U_1 \cup U_2$ enthält. Also reicht es zu zeigen, dass $U_1 + U_2$ ein Untervektorraum ist:

$$\mathbf{0} + \mathbf{0} = \mathbf{0} \in U_1 + U_2$$
$$(u_1 + u_2) + (v_1 + v_2) = (u_1 + v_1) + (u_2 + v_2) \in U_1 + U_2$$
$$\lambda(u_1 + u_2) = (\lambda u_1) + (\lambda u_2) \in U_1 + U_2$$

Beispiele mit Zeichnung an der Tafel: Wie hängt die Dimension von Durchschnitt und Summe von den Dimensionen der einzelnen Teile ab? Betrachten Vereinigung und Schnitt von Gerade U_1 und Ebene U_2 im $\mathbb{R}^3$. Jeweils zwei Fälle:

- Gerade liegt in der Ebene, $U_1 \leq U_2$.
  $\dim(U_1 \cap U_2) = 1$.
  $\dim(U_1 + U_2) = 2$.
- Gerade liegt nicht in der Ebene.
  $\dim(U_1 \cap U_2) = 0$.
  $\dim(U_1 + U_2) = 3$. □

Satz 2.4.19 (Dimensionssatz)**.** *Sind $U_1, U_2 \leq V$ endlichdimensional, so gilt*

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

Speziell gilt also

$$\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2)\,.$$

Bemerkung:

$$\dim(U_1 \cap U_2) \leq \min\{\dim U_1, \dim U_2\}$$

da $U_1 \cap U_2 \leq U_1$ und $U_1 \cap U_2 \leq U_2$ (Abschnitt 2.3.3).

Beweis von Satz 2.4.19. Sei $U_0 := U_1 \cap U_2$ und $B = (v_1, \ldots, v_m)$ Basis von U_0. Dann kann B zu einer Basis

$$B_1 = (v_1, \ldots, v_m, w_1, \ldots, w_r)$$

von U_1 und zu Basis

$$B_2 = (v_1, \ldots, v_m, u_1, \ldots, u_s)$$

von U_2 ergänzt werden (Folgerung von Satz 2.4.13 da $U_0 \leq U_1$ und $U_0 \leq U_2$).

Behauptung:

$$(v_1, \ldots, v_m, w_1, \ldots, w_r, u_1, \ldots, u_s)$$

ist Basis von $U_1 + U_2$. Denn

$$\langle B_1 \cup B_2 \rangle = U_1 + U_2$$

und $v_1, \ldots, v_m, w_1, \ldots, w_r, u_1, \ldots, u_s$ sind linear unabhängig: sei

$$\lambda v_1 + \cdots + \lambda_m v_m + \mu_1 w_1 + \cdots \mu_r w_r + \gamma_1 u_1 + \cdots + \gamma_s u_s = \mathbf{0}$$

Es folgt

$$\underbrace{\gamma_1 u_1 + \cdots + \gamma_s u_s}_{\in U_1} = -(\underbrace{\lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 w_1 + \cdots + \mu_r w_r}_{\in U_2}) \in U_1 \cap U_2 \tag{2.1}$$

Also gibt es Darstellung durch Basis B:

$$\gamma_1 u_1 + \cdots + \gamma_s u_s = \alpha_1 v_1 + \cdots + \alpha_m v_m$$

Da $B_2 = (v_1, \ldots, v_m, u_1, \ldots, u_s)$ linear unabhängig erhalten wir

$$\gamma_1 = \cdots = \gamma_s = \alpha_1 = \cdots = \alpha_m = 0$$

Also folgt aus (2.1) dass

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 w_1 + \cdots + \mu_r w_r = \mathbf{0}$$

Da $B_1 = (v_1, \ldots, v_m, \mu_1, \ldots, \mu_r)$ linear unabhängig gilt

$$\lambda_1 = \cdots = \lambda_m = \mu_1 = \cdots = \mu_r = 0\,.$$

Damit ist lineare Unabhängigkeit von $(v_1, \ldots, v_m, w_1, \ldots, w_r, u_1, \ldots, u_s)$ bewiesen. □

Übung 6. Nach Satz 2.4.19 gilt $\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$. Was halten Sie von folgender Aussage:

$$\begin{aligned}\dim(U_1 + U_2 + U_3) \overset{?}{=}\ & \dim(U_1) + \dim(U_2) + \dim(U_3) \\ & - \dim(U_1 \cap U_2) - \dim(U_1 \cap U_3) - \dim(U_2 \cap U_3) \\ & + \dim(U_1 \cap U_2 \cap U_3)\end{aligned}$$

Satz 2.4.20 (Charakterisierungssatz für direkte Summen)**.** *Seien $U_1, U_2 \leq V$. Dann gilt genau dann $V = U_1 \oplus U_2$, wenn sich jeder Vektor $v \in V$* eindeutig *als Summe $v = v_1 + v_2$ mit $v_1 \in U_1$ und $v_2 \in U_2$ darstellen lässt.*

Beweis. Zuerst die Rückrichtung: Haben $V = U_1 + U_2$. Ist $u \in U_1 \cap U_2$, so ist $u + (-u) = \mathbf{0} = \mathbf{0} + \mathbf{0}$. Also folgt aus Eindeutigkeit $u = \mathbf{0}$, d.h., $U_1 \cap U_2 = \{\mathbf{0}\}$.

Hinrichtung: Da $V = U_1 \oplus U_2$ lässt sich jedes $v \in V$ als Summe $v_1 + v_2$ darstellen. Eindeutigkeit: ist $v = v_1 + v_2 = v_1' + v_2'$ mit $v_1, v_1' \in U_1$ und $v_2, v_2' \in U_2$, so folgt

$$u_1 - v_1' = v_2' - v_2 =: u$$

also

$$u \in U_1 \cap U_2 = \{\mathbf{0}\}$$

Also $v_1 - v_1' = \mathbf{0}$ und daher $v_1 = v_1'$, und $v_2 - v_2' = \mathbf{0}$ und daher $v_2 = v_2'$. □

Direkte Summen mit mehreren Summanden:

$$U_1 \oplus U_2 \oplus U_3 := (U_1 \oplus U_2) \oplus U_3$$

Man darf die Klammern weglassen:

$$(U_1 \oplus U_2) \oplus U_3 = U_1 \oplus (U_2 \oplus U_3)$$

Zunächst gilt $U_1 + (U_2 + U_3) = U_1 + (U_2 + U_3)$. Weiterhin gilt

$$\begin{aligned}(a) \quad & U_1 \cap U_2 = \{\mathbf{0}\} \\ \text{und } (b) \quad & (U_1 + U_2) \cap U_3 = \{\mathbf{0}\}\end{aligned}$$

genau dann wenn

$$\begin{aligned} (c) \quad & U_1 \cap (U_2 \oplus U_3) = \{\mathbf{0}\} \\ \text{und } (d) \quad & U_2 \cap U_3 = \{\mathbf{0}\}\,. \end{aligned}$$

Davon die Hinrichtung: $U_2 \cap U_3 \subseteq (U_1 + U_2) \cap U_3 = \{\mathbf{0}\}$ wegen (b), also gilt (d). Sei $v \in U_1 \cap (U_2 \oplus U_3)$. Dann gilt $v = u_2 + u_3$ für $u_2 \in U_2$ und $u_3 \in U_3$. Also $v - u_2 = u_3 \in (U_1 + U_2) \cap U_3 = \{\mathbf{0}\}$ wegen (b). Ausserdem $v = u_2 \in U_1 \cap U_2 = \{\mathbf{0}\}$, daher $v = \mathbf{0}$, also (c). Rückrichtung ähnlich.

Analoges gilt für beliebig viele Summanden.

Satz 2.4.21 (Zerlegungssatz)**.** *Sei V ein $\mathbb{K}$-Vektorraum, $n = \dim V$, und $\{v_1, \ldots, v_n\}$ eine Basis von V. Dann ist*

$$V = U_1 \oplus \cdots \oplus U_n$$

für $U_i := \langle u_i \rangle = Kv_i$.

Beispiel 2.4.22. $\mathbb{R}^n = \mathbb{R}e_1 \oplus \cdots \oplus \mathbb{R}e_n$. △

Folgerung: Jeder Untervektorraum eines endlichdimensionalen Vektorraums V hat ein Komplement.

Beweis von Satz 2.4.21.

$$V = U_1 + \cdots + U_n = \langle u_1, \ldots u_n \rangle$$

Es bleibt zu zeigen: $(U_1 + \cdots + U_i) \cap U_{i+1} = \{\mathbf{0}\}$ für $i \in \{1, \ldots, n-1\}$. □

Übung 7. Vervollständigen Sie den Beweis von Satz 2.4.21.

Kapitel 3

Lineare Abbildungen, Gleichungssysteme, Matrizen

3.1 Lineare Abbildungen I

Lineare Abbildungen sind die wesentlichen strukturerhaltende Abbildung für Vektorräume.

Definition 3.1.1 (Lineare Abbildung)**.** Es seien V und W Vektorräume über einem Körper $\mathbb{K}$. Eine Funktion $f\colon V \to W$ heißt *lineare Abbildung* oder *(Vektorraum-) Homomorphismus* wenn gilt

- für alle $v, v' \in V$:

$$f(v + v') = f(v) + f(v') \qquad \text{(Verträglichkeit mit der Addition)}$$

- für alle $v \in V$ und $\lambda \in K$:

$$f(\lambda v) = \lambda f(v) \qquad \text{(Verträglichkeit mit der skalaren Multiplikation)}$$

Insbesondere folgt für $\lambda = 0$, dass $f(\mathbf{0}) = \mathbf{0}$. Wir sprechen von einem *Isomorphismus* wenn f zusätzlich bijektiv ist. Weitere Bezeichnungen für den Spezialfall $V = W$:

- eine lineare Abbildung $f\colon V \to V$ heißt *Endomorphismus* (von V);
- ein Isomorphismus $f\colon V \to V$ heißt *Automorphismus* (von V).

Zwei Vektorräume V und W heißen *isomorph* wenn ein Isomorphismus $f\colon V \to W$ existiert. Es handelt sich also bei isomorphen V und W bis auf Umbenennung der Elemente um den gleichen Vektorraum.

Bemerkung 3.1.2. Wenn $f\colon V \to W$ ein Isomorphismus ist, dann auch $f^{-1}\colon W \to V$: seien $u, v \in W$ und $\lambda \in K$. Wegen der Surjektivität von f gibt es $u', v' \in V$ mit $f(u') = u$ und $f(v') = v$. Dann gilt

$$f^{-1}(u + v) = f^{-1}(f(u') + f(v')) = f^{-1}(f(u' + v')) = u' + v' = f^{-1}(u) + f^{-1}(v),$$
$$f^{-1}(\lambda f(u)) = f^{-1}(f(\lambda u)) = \lambda u.$$

Proposition 3.1.3. *Die Komposition linearer Abbildungen ist wieder linear: wenn $f\colon V_1 \to V_2$ und $g\colon V_2 \to V_3$ linear sind, dann ist auch $g \circ f\colon V_1 \to V_3$ linear.*

Beweis. Für alle $v, v' \in V_1$ und $\lambda \in K$ gilt

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

□

3.2 Matrizen

Beispiel einer 2×3-Matrix über $\mathbb{R}$:

$$\begin{pmatrix} 5 & -3 & 0 \\ 1 & 2.5 & \pi \end{pmatrix}$$

Formal:

Definition 3.2.1 (Matrizen)**.** Sei $\mathbb{K}$ ein Körper. Eine *$m \times n$-Matrix mit Einträgen aus* $\mathbb{K}$ ist ein Element des mn-dimensionalen Vektorraums $\mathbb{K}^{mn}$ über K. Schreiben $\mathbb{K}^{m \times n}$ für die Menge aller $(m \times n)$-Matrizen.

Insbesondere dürfen wir für $\lambda \in K$ und $m \times n$-Matrizen M, N schreiben: λM und $M + N$, und $\mathbf{0}$ steht für die Matrix in $\mathbb{K}^{m \times n}$, deren Einträge allesamt 0 sind.

Motivationen: Matrizen dienen der kompakten Beschreibung von z.B.

- linearen Abbildungen;
- linearen Gleichungssystemen.

Beschreibung linearer Abbildungen

Sei B eine $(m \times n)$-Matrix über $\mathbb{K}$. Dann beschreibt B die lineare Abbildung

$$f_B\colon \mathbb{K}^n \to \mathbb{K}^m : x \mapsto Bx$$

wobei

$$Bx := \sum_{i=1}^{n} x_i \begin{pmatrix} b_{1i} \\ \dots \\ b_{mi} \end{pmatrix}.$$

Beispiel 3.2.2. Für $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist $f_B = \mathrm{id}_{\mathbb{R}^2}$:

$$Bx = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x.$$

△

Beschreibung linearer Gleichungssysteme

$$\begin{aligned} b_{11}x_1 + \cdots + b_{1n}x_n &= z_1 \\ \cdots \quad &= \cdots \\ b_{m1}x_1 + \cdots + b_{mn}x_n &= z_m \end{aligned}$$

Übersichtlichere Schreibweise für das lineare Gleichungssystem:

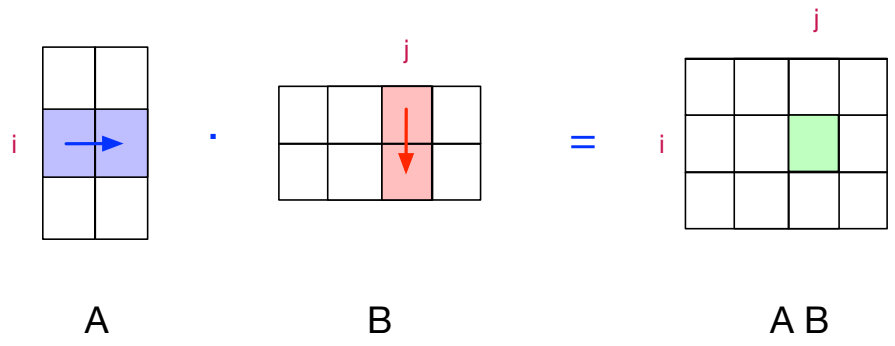
$$\underbrace{\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}}_{B} \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} z_1 \\ \cdots \\ z_m \end{pmatrix}$$

mit Hilfe der *Koeffizientenmatrix* B. Kurzschreibweise: $Bx = z$.

3.2.1 Matrizenmultiplikation

Seien $A \in \mathbb{K}^{r\times m}$ und $B \in \mathbb{K}^{m\times n}$.
Dann lassen sich die Funktionen $f_B\colon \mathbb{K}^n \to \mathbb{K}^m : z \mapsto Bz$ und $f_A\colon \mathbb{K}^m \to \mathbb{K}^r : z \mapsto Az$ komponieren: sei $f_{AB} := f_A \circ f_B$.

Beobachtung: es gibt ein $C \in \mathbb{K}^{n\times r}$ mit $f_C = f_{AB}$. Schreiben: "$C = AB$".



Definition 3.2.3. Das Produkt AB zweier Matrizen $A \in \mathbb{K}^{r\times m}$ und $B \in \mathbb{K}^{m'\times n}$ ist genau dann definiert, wenn $m = m'$ und zwar durch

$$AB := \left(\sum_{k=1}^{m} a_{ik} b_{kj} \right)_{i=1,\ldots,r,\, j=1,\ldots,n} \in \mathbb{K}^{r\times n}$$

für $A = (a_{ij})_{i=1,\ldots,r, j=1,\ldots,m}$ und $B = (b_{ij})_{i=1,\ldots,m', j=1,\ldots,n}$.

Proposition 3.2.4. *Für die Matrizenmultiplikation gelten:*

1. *Assoziativitätsgesetz*

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

2. *Die* Einheitsmatrizen

$$E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{K}^{n\times n}$$

sind Eins-elemente *für die Multiplikation: Für* $A \in \mathbb{K}^{m\times n}$ *gilt*

$$E_m A = A \text{ und } AE_n = A$$

3. *Distributivgesetze*

$$\begin{aligned} A(B+C) &= AB + AC \\ (B+C)A &= BA + CA \end{aligned}$$

4. *Verträglichkeit mit der Skalarmultiplikation: für* $\lambda \in K$ *gilt*

$$(\lambda A)B = \lambda(AB) = A(\lambda B) \tag{3.1}$$

Übung 8. Beweisen Sie Proposition 3.2.4.

Definition 3.2.5. Die i-te Zeile der Einheitsmatrix hat die Gestalt

$$e_i = (0, \ldots, 0, \underbrace{1}_{i\text{-te Stelle}}, 0, \ldots, 0) \in \mathbb{K}^n$$

und wird *Einheitsvektor* genannt.

Beispiel 3.2.6. Matrizenmultiplikation in $\mathbb{K}^{n\times n}$ ist nicht kommutativ.

$$\begin{aligned} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{3.2}$$

△

Die Menge $\mathbb{K}^{n\times n}$ mit Addition und der eben definierten Multiplikation ist für $n \geq 2$ nicht einmal ein Schiefkörper, da es Matrizen ungleich $\mathbf{0}$ gibt, die kein multiplikatives Inverses haben, wie Beispiel 3.2.6 (3.2) ebenfalls zeigt.

Definition 3.2.7. Eine Matrix $A \in \mathbb{K}^{m\times n}$ heißt *invertierbar* (oder *regulär* oder *nichtsingulär*) wenn $m = n$ (*quadratische* Matrix) und eine Matrix $B \in \mathbb{K}^{n\times n}$ existiert, so dass

$$AB = BA = E_n \,.$$

Die Matrix B ist durch A eindeutig bestimmt (Lemma 2.1.4!) und wird mit A^{-1} bezeichnet.

Bezeichnung:

$$\mathrm{GL}(n, K) := \{A \in \mathbb{K}^{n\times n} \mid A \text{ invertierbar}\}$$

(englisch: *"general linear group"*)

Eigenschaften.

1. Für $A, B \in \mathbb{K}^{n\times n}$ gilt
$$AB = E_n \Leftrightarrow BA = E_n \Leftrightarrow B = A^{-1} \Leftrightarrow A = B^{-1}$$
2. Ist A invertierbar, so auch A^{-1} und es gilt $(A^{-1})^{-1} = A$.
3. Sind $A, B \in \mathbb{K}^{n\times n}$ invertierbare Matrizen, so ist auch AB invertierbar und es gilt
$$(AB)^{-1} = B^{-1}A^{-1} \tag{3.3}$$
4. $\mathrm{GL}(n, K)$ ist bezüglich der Matrizenmultiplikation eine Gruppe.
5. Für $\lambda \in K \setminus \{0\}$ gilt $(\lambda A)^{-1} = \lambda^{-1}A^{-1}$. Denn: $\lambda^{-1}A^{-1}(\lambda A) = \lambda^{-1}\lambda A^{-1}A = E_n$.

Beispiel 3.2.8. Eine Matrix $A \in \mathbb{K}^{n\times n}$ der Gestalt

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

für $\lambda_1, \ldots, \lambda_n \in K$ heißt *Diagonalmatrix*. Eine Diagonalmatrix ist genau dann invertierbar, wenn $\lambda_1, \ldots, \lambda_n \neq 0$, und in diesem Fall gilt

$$A^{-1} = \begin{pmatrix} \lambda_1^{-1} & & & 0 \\ & \lambda_2^{-1} & & \\ & & \ddots & \\ 0 & & & \lambda_n^{-1} \end{pmatrix} \qquad \triangle$$

Übung 9. Wie berechnet sich das Produkt von Diagonalmatrizen?

Übung 10. Sei $D \in \mathbb{K}^{n\times n}$ eine Diagonalmatrix und $A \in \mathbb{K}^{n\times n}$ beliebig. Unter welchen Voraussetzungen an D gilt $DA = AD$?

3.2.2 Rang

Sei $A = (a_{ij}) \in \mathbb{K}^{m\times n}$.

$$\begin{array}{cc} \begin{array}{ccc} s_1 & \ldots & s_n \end{array} & \\ \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & & \ldots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} & \begin{array}{c} z_1 \\ \ldots \\ z_m \end{array} \end{array}$$

Zeilen von A sind Elemente aus $\mathbb{K}^n$ und die Spalten von A sind Elemente aus $\mathbb{K}^m$!

Definition 3.2.9 (Rang)**.** Die maximale Zahl linear unabhängiger Spalten von A (in $\mathbb{K}^m$) heißt *Spaltenrang* von A. Die maximale Zahl linear unabhängiger Zeilen von A (in $\mathbb{K}^n$) heißt *Zeilenrang* von A.

Bemerkung 3.2.10. Seien $s_1, \ldots, s_n$ die Spalten von A. Dann

$$\text{Spaltenrang von } A = \dim_{\mathbb{K}} \langle s_1, \ldots, s_n \rangle .$$

Seien $z_1, \ldots, z_m$ die Zeilen von A. Dann

$$\text{Zeilenrang von } A = \dim_{\mathbb{K}} \langle z_1, \ldots, z_m \rangle .$$

Satz 3.2.11. *Dann*

$$\text{Zeilenrang von } A = \text{Spaltenrang von } A$$

Definieren $\operatorname{rg}(A) :=$ Zeilenrang von A = Spaltenrang von A.

Beweis. Eine Spalte heiße *linear überflüssig* wenn sie Linearkombination der anderen Spalten ist. Analog für Zeilen. Weglassen einer linear überflüssigen Spalte ändert den Spaltenrang nicht.

Behauptung. Weglassen einer linear überflüssigen Spalte ändert auch den Zeilenrang nicht! Sei etwa letzte Spalte s_n linear überflüssig, d.h.,

$$v_n = \lambda_1 s_1 + \cdots + \lambda_{n-1} s_{n-1}$$

also $a_{in} = \lambda_1 a_{i1} + \cdots + \lambda_{n-1} a_{i,n-1}$ für $i \in \{1, \ldots, n\}$. Durch Weglassen der n-ten Spalte entstehe aus A die Matrix A' mit Zeilen $z_1', \ldots, z_m'$. Dann gilt

$$\alpha_1 z_1' + \cdots + \alpha_m z_m' = \mathbf{0} \Leftrightarrow \alpha_1 z_1 + \cdots + \alpha_m z_m = \mathbf{0}$$

Rückrichtung hierbei klar. Hinrichtung: für die letzte Komponente gilt

$$\begin{aligned} \alpha_1 a_{1n} + \cdots + \alpha_m a_{mn} &= \alpha_1 \left(\sum_{k=1}^{n-1} \lambda_k a_{1k} \right) + \cdots + \alpha_m \left(\sum_{k=1}^{n-1} \lambda_k a_{mk} \right) \\ &= \sum_{k=1}^{n-1} \lambda_k (\alpha_1 a_{1k} + \cdots + \alpha_m a_{mk}) = 0 \end{aligned}$$

Analog ändert das Weglassen einer linear überflüssigen Zeile nicht den Spaltenrang. Durch sukzessives Weglassen von linear überflüssigen Zeilen und Spalten gelangt man zu einer $m' \times n'$-Matrix $A' \in \mathbb{K}^{m' \times n'}$ ohne linear überflüssige Zeilen oder Spalten, mit m' Zeilen in $\mathbb{K}^{n'}$ und n' Spalten in $\mathbb{K}^{m'}$:

$$\begin{aligned} \text{Zeilenrang}(A) = \text{Zeilenrang}(A') = m' \leq \dim(\mathbb{K}^{n'}) = n' \\ \text{Spaltenrang}(A) = \text{Spaltenrang}(A') = n' \leq \dim(\mathbb{K}^{m'}) = m' \end{aligned}$$

Also $m' = n'$. □

Beispiel 3.2.12. Der Rang der Nullmatrix $\mathbf{0}$ (alle Einträge 0) ist Null: $\operatorname{rg}(\mathbf{0}) = 0$. △

Bemerkung 3.2.13. Der Beweis von Satz 3.2.11 zum Rang ein Matrix A ist leider nicht komplett algorithmisch: denn es ist bis zu dieser Stelle der Vorlesung (noch) nicht klar, wie man (effizient!) berechnet, ob eine Spalte beziehungsweise eine Zeile von A linear überflüssig ist. Die nächsten beiden Abschnitte werden dieses Problem lösen.

3.2.3 Zeilenumformungen

Sei $A \in \mathbb{K}^{m \times n}$. Die folgenden Umformungen von A heißen *elementare Zeilenumformungen* (manchmal auch *(elementare) Zeilentransformationen)*:

(1) Vertauschung zweier Zeilen;

(2) Multiplikation einer Zeile mit einem Skalar $\lambda \in K \setminus \{0\}$;

(3) Addition des λ-fachen ($\lambda \in K$) einer Zeile zu einer anderen Zeile.

Analog: *elementare Spaltenumformungen.*

Bemerkung 3.2.14. Mit (1) lassen sich die Zeilen beliebig permutieren (Satz 4.1.1).

Bemerkung 3.2.15. Jede elementare Zeilenumformung lässt sich wieder mit einer elementaren Zeilenumformung rückgängig machen.

Beispiel 3.2.16. In diesem Beispiel werden die einzelnen Zeilentransformationen durch Pfeile angedeutet, die unten beschriftet sind mit einem der genannten drei Typen (1), (2), oder (3) der Transformation, und oben beschriftet sind mit einer Beschreibung der Transformation; zum Beispiel steht $z_2 + z_1 \rightsquigarrow z_2$ für die Transformation, die die zweite Zeile durch die Summe der ersten beiden Zeilen ersetzt.

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow[(3)]{z_2+z_1 \rightsquigarrow z_2} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 4 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow[(3)]{z_2-4z_3 \rightsquigarrow z_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Vorteil der letzten Matrix: der Zeilenrang (nämlich 2) ist sofort ablesbar. △

Lemma 3.2.17. *Elementare Umformungen ändern den Rang einer Matrix nicht.*

Beweis. $\langle z_1, \ldots, z_n \rangle$ bleibt bei elementaren Zeilenumformungen erhalten:

- $\langle z_1, z_2 \rangle = \langle \lambda z_2, z_1 \rangle$
- für $\lambda \in K \setminus \{0\}$ gilt: $\langle z \rangle = \langle \lambda z \rangle$
- $\langle z_1, z_2 + \lambda z_1 \rangle = \langle z_1, z_2 \rangle$ □

Bemerkung 3.2.18. Jede Zeilenumformung einer Matrix A lässt sich beschreiben als Matrizenmultiplikation TA von A mit einer geeigneten Matrix T:

1. $\lambda z_i \rightsquigarrow z_i$ (Multiplikation der Zeile z_i mit λ): Wähle

$$T := \begin{pmatrix} 1 & 0 & \cdots & & & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & \cdots & 0 & 1 \end{pmatrix}$$

2. $z_i \leftrightsquigarrow z_j$ (Vertauschung von Zeile z_i und z_j): Wähle

$$T := \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & \cdots & 1 & & & \\ & & & \vdots & & \vdots & & & \\ & & & 1 & \cdots & 0 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}$$

3. $z_i + \lambda z_j \rightsquigarrow z_i$ (Addition der Zeile z_i mit dem λ-fachen der Zeile z_j):

$$T := \begin{array}{cc} & \begin{array}{cccc} & & j & \\ \end{array} \\ & \begin{pmatrix} 1 & & & \\ & \ddots & \lambda & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{array}{c} \\ i \\ \\ \\ \end{array} \end{array}$$

Diese Matrizen T werden auch *Elementarmatrizen* genannt.

Bemerkung 3.2.19. Offensichtlich ist jede Elementarmatrix invertierbar, und die inverse Matrix ist ebenfalls eine Elementarmatrix:

- Falls T die Elementarmatrix ist von $\lambda z_i \rightsquigarrow z_i$, dann ist T^{-1} die Elementarmatrix von $\frac{1}{\lambda} z_i \rightsquigarrow z_i$;
- Falls T die Elementarmatrix ist von $z_i \leftrightsquigarrow z_j$, dann ist $T^{-1} = T$;
- Falls T die Elementarmatrix ist von $z_i + \lambda z_j \rightsquigarrow z_i$, dann ist T^{-1} die Elementarmatrix von $z_i - \lambda z_j \rightsquigarrow z_i$.

Bemerkung 3.2.20. Analog lassen sich elementare Spaltenumformungen durch Multiplikation mit Elementarmatrizen *von rechts* beschreiben.

3.2.4 Algorithmus zur Umwandlung einer Matrix in Stufenform

In diesem Abschnitt stellen wir eine Prozedur zur Rangbestimmung vor. Es handelt sich um das Kernstück des Gaußschen Algorithmus zur Lösung von linearen Gleichungssystemen.

Idee: Erzeugen mit Hilfe von elementaren Zeilenumformungen eine Matrix, deren Rang direkt sichtbar ist.

Definition 3.2.21. $A \in \mathbb{K}^{m\times n}$ ist in *(oberer) Stufenform*, falls A von der Gestalt

$$A = \begin{pmatrix} 0 & \cdots & 0 & a_{1j_1} & \cdots & a_{1j_2} & \cdots & a_{1j_3} & \cdots & a_{1j_r} & \cdots & a_{1n} \\ 0 & \cdots & 0 & \cdots & 0 & a_{2j_2} & \dots & a_{2j_3} & \cdots & & & \vdots \\ \vdots & & & & \cdots & 0 & \cdots & 0 & \ddots & & & \vdots \\ \vdots & & & & & & & & \ddots & a_{rj_r} & \cdots & a_{rn} \\ 0 & \cdots & & & & & & & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & \cdots & & & & & & & \cdots & 0 & \cdots & 0 \end{pmatrix} \tag{3.4}$$

mit $0 < j_1 < j_2 < \cdots < j_r \leq n$ und $a_{1j_1}, \dots, a_{rj_r} \in K \setminus \{0\}$.

Bemerkung 3.2.22. Für $A \in \mathbb{K}^{m\times n}$ von der Form (3.4), so gilt $\operatorname{rg}(A) = r$ (gleich der Anzahl der *Stufen*). Grund: die ersten r Zeilen sind linear unabhängig, denn $\lambda_1 z_1 + \cdots + \lambda_r z_r = \mathbf{0}$ impliziert

$$\begin{aligned} \lambda_1 \cdot a_{1j_1} + \lambda_2 \cdot 0 + \cdots + \lambda_r \cdot 0 = 0 \quad &(\Rightarrow \lambda_1 = 0) \qquad && (j_1\text{-te Spalte}) \\ \lambda_2 \cdot a_{2j_2} + \cdots + \cdot\lambda_r \cdot 0 = 0 \quad &(\Rightarrow \lambda_2 = 0) \qquad && (j_2\text{-te Spalte}) \\ &\text{usw.} \end{aligned}$$

Also $j_i = 0$ für alle $i \in \{1, \dots, r\}$, und $\operatorname{rg}(A) = r$.

Beispiel 3.2.23. Betrachten $A \in \mathbb{Q}^{3\times 4}$ wie folgt.

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 8 \\ 1 & 3 & 5 & 1 \end{pmatrix} \xrightarrow[(1)]{z_1 \leftrightarrow z_3} \begin{pmatrix} 1 & 3 & 5 & 1 \\ 0 & 2 & 4 & 8 \\ 0 & 1 & 2 & 3 \end{pmatrix} \xrightarrow[(3)]{z_3 - \frac{1}{2} z_2 \rightsquigarrow z_3} \begin{pmatrix} 1 & 3 & 5 & 1 \\ 0 & 2 & 4 & 8 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Stufenform, $\operatorname{rg}(A) = 3$. △

Allgemein: induktiver Algorithmus zur Überführung einer Matrix in Stufenform mit Hilfe von elementaren Zeilenumformungen. Hat die Matrix $A \in \mathbb{K}^{m\times n}$ die Gestalt[1]

$$\left(\begin{array}{cccccc|ccc} 0 & \cdots & 0 & a_{1j_1} & \cdots & * & * & \cdots & * \\ \vdots & & & \ddots & \ddots & \vdots & \vdots & \text{beliebig} & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & a_{k-1j_{k-1}} & * & \cdots & * \\ \hline & & & & & & a_{k,j_k} & & \\ & & & \mathbf{0} & & & \vdots & B & \\ & & & & & & a_{m,j_k} & & \end{array}\right) \tag{3.5}$$

[1] Sterne in Matrizen stehen für beliebige Einträge.

mit $a_{1j_1}, \ldots, a_{k-1,j_{k-1}} \neq 0$ und k größtmöglich. Falls Stufenform noch nicht erreicht ist, so lässt sich weiter wie folgt verfahren:

<u>1ter Fall:</u> $a_{k,j_k} = 0$. Vertauschen der k-ten Zeile mit einer Zeile, für die $a_{i,j_k} \neq 0$ $(i > k)$ (die gibt es, da Stufenform noch nicht erreicht und k größtmöglich gewählt). Damit o.B.d.A. 2ter Fall.

<u>2ter Fall:</u> $a_{k,j_k} \neq 0$. Von jeder Zeile z_l $(l > k)$ subtrahiere man $(a_{l,j_k} a_{k,j_k}^{-1}) z_k$.

Dies ergibt Matrix der Gestalt

$$\left(\begin{array}{ccccccc|c} 0 & \cdots & 0 & a_{1j_1} & * & \cdots & * & \\ \vdots & & & \ddots & \ddots & \ddots & \vdots & \\ \vdots & & & & \ddots & a_{k-1j_{k-1}} & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & a_{k,j_k} & \\ \hline & & & & \mathbf{0} & & & B' \end{array}\right)$$

Der Algorithmus endet, wenn $B' = \mathbf{0}$ oder wenn keine Zeilen mehr vorhanden sind.

Bemerkung 3.2.24. Im Verfahren wurden keine elementaren Zeilenumformungen vom Typ (2) verwendet (Multiplikation einer Zeile mit einem Skalar).

3.2.5 Bestimmung von Dimension und Basen

Seien $u_1, \ldots, u_m \in \mathbb{K}^n$. Bestimmung von $d := \dim \langle u_1, \ldots, u_m \rangle$: Sei

$$A := \begin{pmatrix} -u_1- \\ \vdots \\ -u_m- \end{pmatrix} \in \mathbb{K}^{m \times n}$$

Dann gilt $d = \operatorname{rg}(A)$ und d ist durch Umformen von A in Zeilen-Stufenform bestimmbar. Siehe Lemma 3.2.17 (elementare Zeilenumformungen ändern den Rang nicht) und Beobachtung am Ende von Abschnitt 3.2.3 (Ablesen des Rangs in der Stufenform).

Bestimmung einer Basis von $V := \langle u_1, \ldots, u_m \rangle$: Umformen von

$$A = \begin{pmatrix} -u_1- \\ \vdots \\ -u_m- \end{pmatrix} \in \mathbb{K}^{m \times n}$$

in Stufenform $B \in \mathbb{K}^{m \times n}$. Die vom Nullvektor verschiedenen Zeilen von B bilden eine Basis von V. Denn: falls

$$\begin{pmatrix} -z_1- \\ \vdots \\ -z_m- \end{pmatrix} \rightsquigarrow \begin{pmatrix} -u_1- \\ \vdots \\ -u_m- \end{pmatrix}$$

durch elementare Zeilenumformung, so ist $\langle z_1, \ldots, z_m \rangle = \langle u_1, \ldots, u_m \rangle$ (siehe Beweis von Satz 3.2.17, "Elementare Umformungen ändern den Rang einer Matrix nicht").

Rechnungen in beliebigen endlichdimensionalen Vektorräumen:
Sei nun V ein beliebiger n-dimensionaler $\mathbb{K}$-Vektorraum. Das Rechnen in V lässt sich auf das Rechnen mit Koordinatenvektoren bzgl. einer Basis zurückführen (Grundlage: Satz 2.4.6).

Wiederholung: Sei V ein $\mathbb{K}$-Vektorraum mit geordneter Basis $B = (v_1, \dots, v_n)$. Dann ist

$$\Phi_B \colon \mathbb{K}^n \to V : (\lambda_1, \dots, \lambda_n) \mapsto \lambda_1 v_1 + \dots + \lambda_n v_n$$

ein Isomorphismus, d.h., eine bijektive Abbildung mit den Eigenschaften

- $\Phi(z_1 + z_2) = \Phi(z_1) + \Phi(z_2)$;
- $\Phi(\lambda z) = \lambda \Phi(z)$.

D.h., $\mathbb{K}^n$ und V sind "im Prinzip" der gleiche Vektorraum (Vergleiche: Satz 2.4.6).

Gegeben: $w_1, \dots, w_m \in V$.
Gesucht: Basis von $\langle w_1, \dots, w_m \rangle$.
Ansatz: Für $i \in \{1, \dots, m\}$ gibt es $u_i = (\lambda_{i1}, \dots, \lambda_{in}) \in \mathbb{K}^n$ so dass $w_i = \Phi_B(u_i)$. Also:

$$\begin{aligned} w_1 &= \lambda_{11} v_1 + \dots + \lambda_{1n} v_n \\ &\vdots \\ w_m &= \lambda_{m1} v_1 + \dots + \lambda_{mn} v_n \end{aligned}$$

Setzen

$$A := \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix}$$

Dann gelten

$$\dim_V \langle w_1, \dots, w_m \rangle = \dim_{\mathbb{K}^n} \langle u_1, \dots, u_m \rangle = \operatorname{rg}(A) \tag{3.6}$$

und

$$\langle w_1, \dots, w_m \rangle = \Phi_B(\langle u_1, \dots, u_m \rangle)\,. \tag{3.7}$$

Insbesondere: $(b_1, \dots, b_s)$ ist genau dann Basis von $\langle u_1, \dots, u_m \rangle$, wenn $(\Phi(b_1), \dots, \Phi(b_s))$ Basis von $\langle w_1, \dots, w_m \rangle$ ist.

3.2.6 Invertierbarkeitskriterium

Wir lernen nun ein wichtiges notwendiges und hinreichendes Kriterium für die Invertierbarkeit von Matrizen kennen.

Satz 3.2.25. *Eine Matrix $A \in \mathbb{K}^{n \times n}$ ist genau dann invertierbar, wenn $\operatorname{rg}(A) = n$.*

Beweis. "$\Rightarrow$": Sei A invertierbar. Wir müssen zeigen, dass die Spalten $v_1, \dots, v_n$ von A linear unabhängig sind. Wir nehmen an, dass $\lambda_1 v_1 + \cdots + \lambda_n v_n = \mathbf{0}$, also dass

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0}.$$

Dann gilt

$$A^{-1} A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \mathbf{0}.$$

Da $A^{-1}A = E_n$ erhalten wir $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \mathbf{0}$ and daher $\lambda_1 = \cdots = \lambda_n = 0$. Also ist der Spaltenrang von A gleich n.

"$\Leftarrow$": Sei $\operatorname{rg}(A) = n$. Dann sind Spalten $v_1, \dots, v_n$ von A linear unabhängig, also eine Basis von $\mathbb{K}^n$, und $\langle v_1, \dots, v_n \rangle = \mathbb{K}^n$. Also gibt es Linearkombinationen

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_1 = b_{11} v_1 + \cdots + b_{n1} v_n$$

$$\vdots$$

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = e_n = b_{1n} v_1 + \cdots + b_{nn} v_n$$

Also ist

$$A \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} | & & | \\ b_{11} v_1 + \cdots + b_{n1} v_n & \cdots & b_{1n} v_1 + \cdots + b_{nn} v_n \\ | & & | \end{pmatrix} = E_n$$

und haben damit das Inverse

$$A^{-1} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}$$

zu A gefunden. $\square$

3.2.7 Konstruktion der inversen Matrix

Es gelte $AB = C$ für drei Matrizen $A, B, C \in \mathbb{K}^{n \times n}$. Überführt man A und C durch die *gleichen* elementaren Zeilenumformungen in Matrizen A' und C', dann gilt auch

$A'B = C'$. Denn: eine elementare Zeilenumformung entspricht Multiplikation von links mit einer Elementarmatrix T, und damit:

$$\begin{aligned} AB = C &\Rightarrow T(AB) = TC \\ &\Rightarrow (TA)B = TC \qquad \text{(nach Proposition 3.2.4).} \end{aligned}$$

Folgerung:

$$\begin{array}{ccc} AA^{-1} & = & E_n \\ \wr & & \wr \\ E_nA^{-1} & = & A^{-1} \end{array}$$

Das bedeutet: erhält man E_n durch elementare *Zeilen*umformungen aus einer Matrix A, so verwandeln die gleichen Zeilenumformungen die Matrix E_n in die Matrix A^{-1}.

Beispiel 3.2.26.

$$\begin{array}{rcl} A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_3 \\ z_3 - z_1 \rightsquigarrow z_3: & \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\ 2^{-1}z_3 \rightsquigarrow z_3: & \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2^{-1} & 0 & 2^{-1} \end{pmatrix} \\ z_1 - (-2)z_3 \rightsquigarrow z_1: & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -2^{-1} & 0 & 2^{-1} \end{pmatrix} = A^{-1} \end{array} \qquad \triangle$$

Zur Konstruktion von A^{-1} benötigen wir also einen Algorithmus, der A durch Zeilenumformungen in E_n überführt.

1. Teil: Umformung von A mit Algorithmus in Zeilen-Stufenform (Abschnitt 3.2.4). Gibt es weniger als n Stufen, so ist $\operatorname{rg}(A) \leq n-1$, und A ist nicht invertierbar (siehe Abschnitt 3.2.5). Ansonsten hat A die Form

$$\begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

wobei alle $a_{ii} \in K \setminus \{0\}$.

2. Teil, 1. Schritt: Alle Diagonalelemente zu 1:
Multiplikation von z_i mit a_{ii}^{-1} für $i \in \{1, \dots, n\}$:

$$\begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

2. Schritt: Alle Elemente $*$ zu Null machen: Für $j = 2, 3, \dots, n$ (der Reihe nach) Bearbeitung der Spalte j: Von Zeile z_i mit $i \in \{1, \dots, j-1\}$ wird $a_{ij} z_j$ abgezogen (ersetze z_i durch $z_i - a_{ij} z_j$). Dies ergibt

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & * & \cdots & * \\ 0 & 1 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ \vdots & & 0 & 1 & * & & \vdots \\ \vdots & & & 0 & 1 & \ddots & \vdots \\ \vdots & & & \vdots & \ddots & \ddots & * \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & 1 \end{pmatrix}$$

und führt schließlich zu E_n.

Satz 3.2.27. *Sei $A \in \mathbb{K}^{n \times n}$. Die folgenden Aussagen sind äquivalent:*

1. *A ist invertierbar (A^{-1} existiert)*
2. $\operatorname{rg}(A) = n$
3. *Die Spalten von A sind linear unabhängig*
4. *Die Zeilen von A sind linear unabhängig*
5. *A kann durch elementare Zeilenumformungen in E_n umgewandelt werden*
6. $\det A \neq 0$ *(kommt später in Abschnitt 4.1, Satz 4.1.6)*

Beweis. (1) $\Leftrightarrow$ (2): Satz 3.2.25.
(2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4): Satz 3.2.11.
(1) $\Leftrightarrow$ (5): Abschnitt 3.2.7. □

3.3 Lineare Gleichungssysteme

3.3.1 Definitionen

Ist $A \in \mathbb{K}^{m \times n}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$ so heißt

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

kurz

$$Ax = b \qquad \text{für } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

lineares Gleichungssystem (LGS) (mit m Gleichungen und n Unbekannten $x_1, \dots, x_n$ und Koeffizienten aus K).

- $b \neq \mathbf{0}$: *inhomogenes* LGS
- $b = \mathbf{0}$: *homogenes* LGS

$Ax = \mathbf{0}$ ist das *zu $Ax = b$ gehörige homogene Gleichungssystem.*

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

heißt *Lösungsmenge* des LGS $Ax = b$.

3.3.2 Lösbarkeitskriterium

Sei $A \in \mathbb{K}^{m \times n}$. Dann heißt $Ax = b$ *lösbar* falls $\text{Lös}(A, b) \neq \varnothing$.

Bemerkung 3.3.1. Seien $s_1, \dots, s_n \in \mathbb{K}^m$ die Spalten von A. Dann gilt

$$f_A(x) = Ax = x_1 s_1 + \cdots + x_n s_n.$$

Also ist $Ax = b$ genau dann lösbar, wenn $b \in \langle s_1, \dots, s_n \rangle$.

Es gibt drei Möglichkeiten:

1. $Ax = b$ ist nicht lösbar: $\text{Lös}(A, b) = \varnothing$.
   Zum Beispiel $x_1 + x_2 = 1$, $x_1 + x_2 = 2$.
2. $Ax = b$ ist eindeutig lösbar: $|\text{Lös}(A, b)| = 1$.
   Zum Beispiel $x_1 + x_2 = 1$, $x_2 = 2$.
3. $Ax = b$ hat mehrere Lösungen: $|\text{Lös}(A, b)| > 1$.
   Zum Beispiel $x_1 + x_2 = 1$, $2x_1 + 2x_2 = 2$.

Proposition 3.3.2. *Ein LGS $Ax = b$ ist genau dann lösbar, wenn*

$$\text{rg}(A) = \text{rg}(A|b).$$

Dabei bezeichne $A|b$ die Matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

Beweis. Schreiben $s_1, \ldots, s_n$ für die Spalten von A.

$$\begin{aligned}
& Ax = b \text{ lösbar} \\
\Leftrightarrow\ & b \in \langle s_1, \ldots, s_n \rangle && \text{(Bemerkung 3.3.1)} \\
\Leftrightarrow\ & \langle s_1, \ldots, s_n \rangle = \langle s_1, \ldots, s_n, b \rangle && \text{(siehe Abschnitt 2.4.1)} \\
\Leftrightarrow\ & \dim\langle s_1, \ldots, s_n \rangle = \dim\langle s_1, \ldots, s_n, b \rangle && \text{(siehe Abschnitt 2.4.5)} \\
\Leftrightarrow\ & \operatorname{rg}(A) = \operatorname{rg}(A|b) && \text{(siehe Abschnitt 3.2.2).}
\end{aligned}$$

□

Bemerkung 3.3.3. $\operatorname{rg}(A) = m$ ist hinreichende (aber nicht notwendige) Bedingung für Lösbarkeit (denn $\operatorname{rg}(A) \leq \operatorname{rg}(A|b) \leq m$).

Satz 3.3.4. *Sei v_0 eine Lösung des LGS $Ax = b$ (d.h., $Av_0 = b$). Dann gilt*

$$\text{Lös}(A, b) = v_0 + \text{Lös}(A, \mathbf{0}) := \{v_0 + v \mid v \in \text{Lös}(A, \mathbf{0})\}$$

"Allgemeine Lösung des inhomogenen LGS
= spezielle Lösung des inhomogenen LGS
+ allgemeine Lösung des zugehörigen homogenen LGS."

Beweis. Sei $v \in \text{Lös}(A, \mathbf{0})$. Nach Distributivitätsgesetz gilt

$$A(v_0 + v) = Av_0 + Av = b + \mathbf{0} = b$$

D.h., $v_0 + v \in \text{Lös}(A, b)$.
Umgekehrt: Sei $w \in \text{Lös}(A, b)$. Dann ist $v := w - v_0 \in \text{Lös}(A, \mathbf{0})$, denn

$$Av = A(w - v_0) = Aw - Av_0 = b - b = \mathbf{0}\,.$$

Also $w = v_0 + v \in v_0 + \text{Lös}(A, \mathbf{0})$. □

Beispiel 3.3.5. $K = \mathbb{R}$, $m = 1$, $n = 2$.

$$2x_1 + 4x_2 = 12$$

'$Ax = b$':

$$A = (2\ 4) \in \mathbb{R}^2, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad b = (1\ 2) \in \mathbb{R}^2$$

$v_0 = \binom{4}{1}$ ist spezielle Lösung.

$$\begin{aligned}
\text{Lös}(A, \mathbf{0}) &= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid 2x_1 + 4x_2 = 0 \right\} \\
&= \left\{ \begin{pmatrix} -2\lambda \\ \lambda \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} =: \mathbb{R}\begin{pmatrix} -2\lambda \\ \lambda \end{pmatrix}
\end{aligned}$$

ist Gerade $x_2 = -\frac{1}{2}x_1$. Zeichnung!

Menge aller Lösungen:

$$\text{Lös}(A, b) = v_0 + \text{Lös}(A, \mathbf{0}) = \left\{ \begin{pmatrix} 4 - 2\lambda \\ 1 + \lambda \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$$

ist Gerade $x_2 = -\frac{1}{2}x_1 + 3$. Zeichnung! △

3.3.3 Bild und Kern

Sei $A \in \mathbb{K}^{m \times n}$. Definieren *Kern* und *Bild* von A. Entspricht Kern und Bild der Abbildung

$$f_A \colon \mathbb{K}^n \to \mathbb{K}^m : x \mapsto Ax$$

Bild von A (ist das Bild von f_A; vergleiche Abschnitt 1.2.3):

$$\text{Bild}\, A := \{Ax \mid x \in \mathbb{K}^n\} = \langle s_1, \ldots, s_n \rangle \leq \mathbb{K}^m$$

wobei $s_1, \ldots, s_n$ die Spalten von A. Also:

$$\text{rg}(A) = \dim(\text{Bild}\, A).$$

Kern von A (ist der Kern von f_A, allerdings etwas anders definiert als in Abschnitt 1.2.3):

$$\text{Kern}\, A := \{x \in \mathbb{K}^n \mid Ax = \mathbf{0}\} = \text{Lös}(A, \mathbf{0}).$$

Einfach nachzurechnen: Kern A ist Untervektorraum von $\mathbb{K}^n$.

Satz 3.3.6. *Es gilt*

$$\dim(\text{Kern}\, A) + \dim(\text{Bild}\, A) = n$$
$$\text{rg}(A) = n - d \qquad \text{'Dimensionsformel'}$$

wobei $d := \dim(\text{Kern}\, A)$ *der* Defekt *von* A.

Beweis: später für lineare Abbildungen, Satz 3.4.15.

Seien $A \in \mathbb{K}^{m \times n}$ und $b \in \mathbb{K}^m$. Ist $v_0 \in \mathbb{K}^n$ spezielle Lösung von $Ax = b$ und ist $(v_1, \ldots, v_d)$ Basis von $\text{Lös}(A, \mathbf{0})$, so ist

$$\text{Lös}(A, b) = \{v_0 + \lambda_1 v_1 + \cdots + \lambda_d v_d \mid \lambda_1, \ldots, \lambda_d \in K\}.$$

Das bedeutet, dass $d = n - \text{rg}(A)$ nach Satz 3.3.6 die Anzahl der frei wählbaren Parameter $(\lambda_1, \ldots, \lambda_d)$ ist für allgemeine Lösung $x = v_0 + \lambda_1 v_1 + \cdots + \lambda_d v_d \in \mathbb{K}^n$.

Korollar 3.3.7. [2] *Seien $A \in \mathbb{K}^{m \times n}$ und $b \in \mathbb{K}^m$ so, dass das LGS $Ax = b$ lösbar. Dann ist $Ax = b$ genau dann* eindeutig *lösbar, wenn* $\text{rg}(A) = n$.

[2] Ein *Korollar* bezeichnet in der Mathematik eine wahre Aussage von Interesse, die sich unmittelbar, oder mit vergleichsweise geringem Aufwand, aus einer (meist direkt davor) bewiesenen Aussage ergibt.

Beweis.

$$\begin{aligned}
Ax = b \text{ ist eindeutig lösbar} &\iff |\operatorname{Lös}(A,b)| = 1 \\
&\iff \operatorname{Lös}(A,\mathbf{0}) = \{\mathbf{0}\} && \text{(Satz 3.3.4)} \\
&\iff \dim(\operatorname{Kern} A) = 0 \\
&\iff n = n - \dim(\operatorname{Kern} A) = \operatorname{rg}(A) && \text{(Satz 3.3.6)}
\end{aligned}$$

□

Bemerkung 3.3.8. Eindeutigkeit der Lösung hängt nur von A ab (nicht von b).

Bemerkung 3.3.9. Für $n = m$ und $\operatorname{rg}(A) = n$ ist $A^{-1}b$ die (eindeutig bestimmte) Lösung von $Ax = b$.

Bemerkung 3.3.10. Für $n > m$ (mehr Unbekannte als Gleichungen) gilt: wenn lösbar, dann nie eindeutig (weil $\operatorname{rg}(A) \leq m < n$).

Bemerkung 3.3.11. Für $n < m$: alle 3 Fälle sind möglich (siehe Abschnitt 3.3.2). Wegen $\operatorname{rg}(A) \leq n < m$ (insbesondere: Zeilenrang$(A) \leq n$) sind einige der Zeilen Linearkombinationen der anderen (also linear überflüssig, falls LGS lösbar).

3.3.4 Der Gaußsche Algorithmus

Wie löst man ein lineares Gleichungssystem $Ax = b$? Prinzip: $Ax = b$ wird durch Zeilenumformungen in ein gleichwertiges LGS $A'x = b'$ umgewandelt, d.h., $\operatorname{Lös}(A,b) = \operatorname{Lös}(A',b')$, so dass Lösbarkeit von $A'x = b'$ leicht entscheidbar ist.

Wir verwenden den Algorithmus aus Abschnitt 3.2.4 mit der erweiterten Koeffizientenmatrix $(A|b)$. Umformung von $(A|b)$ diesem Algorithmus ergibt Matrix $(A'|b')$ in Stufenform.

$$(A'|b') = \left(\begin{array}{ccccccccccc|c}
0 & \cdots & 0 & a'_{1j_1} & \cdots & a'_{1j_2} & \cdots & a'_{1j_3} & \cdots & a'_{1j_r} & \cdots & a'_{1n} & b'_1 \\
0 & \cdots & 0 & \cdots & 0 & a'_{2j_2} & \dots & a'_{2j_3} & \cdots & & & \vdots & b'_2 \\
\vdots & & & & \cdots & 0 & \cdots & 0 & \ddots & & & \vdots & \vdots \\
\vdots & & & & & & & & \ddots & a'_{rj_r} & \cdots & a'_{rn} & b'_r \\
0 & \cdots & & & & & & & \cdots & 0 & \cdots & 0 & b'_{r+1} \\
\vdots & & & & & & & & & & & \vdots & \vdots \\
0 & \cdots & & & & & & & \cdots & 0 & \cdots & 0 & b'_m
\end{array}\right)$$

für $1 \leq j_1 < \cdots < j_r \leq n$ und mit $a'_{1j_1} \neq 0, \dots, a'_{rj_r} \neq 0$.
Merke: $r = \operatorname{rg}(A)$.

1. Fall: $b'_{r+1}, \dots, b'_n$ nicht alle 0. Dann ist $\operatorname{rg}(A|b) > r = \operatorname{rg}(A)$, und nach Proposition 3.3.2 ist $Ax = b$ nicht lösbar.

2. Fall: $b'_{r+1}, \dots, b'_n = 0$. Dann ist $Ax = b$ lösbar nach Proposition 3.3.2 (und eindeutig lösbar falls $r = n$, Korollar 3.3.7).

Lemma 3.3.12. $\operatorname{Lös}(A,b) = \operatorname{Lös}(A',b')$.

Beweis. Elementare Zeilenumformungen ändern nichts (vergleiche Abschnitt 3.2.7):

$$\begin{array}{ccc} Ax & = & b \\ \wr & & \wr \\ A'x & = & b' \end{array} \qquad \text{(elementare Zeilenumformung)}$$

Lösung von $A'x = b'$ einfach berechenbar:

$$a'_{rj_r}x_{j_r} + a'_{rj_r+1}x_{j_r+1} + \cdots + a'_{rn}x_n = b'_r$$

Auflösen nach x_{j_r} ('an Stufe'):

$$x_{j_r} = (a'_{rj_r})^{-1}(b'_r - a'_{rj_r+1}x_{j_r+1} - \cdots - a'_{rn}x_n)$$

Zeile Nr. i:

$$a'_{ij_i}x_{j_i} + a'_{ij_i+1}x_{j_i+1} + \cdots a'_{in}x_n = b'_i$$

auflösen nach x_{j_i}, bereits berechnete x_k für $k > j_i$ einsetzen. Schließlich: die restlichen $x_{j_i}, \ldots, x_{j_{i_1}-1}$ als freie Parameter wählen. □

Beispiel 3.3.13. LGS $Ax = b$:

$$\begin{aligned} x_3 + 3x_4 + 3x_5 &= 2 \\ x_1 + 2x_2 + x_3 + 4x_4 + 3x_5 &= 3 \\ x_1 + 2x_2 + 2x_3 + 7x_4 + 6x_5 &= 5 \\ 2x_1 + 4x_2 + x_3 + 5x_4 + 3x_5 &= 4 \end{aligned}$$

Erweiterte Koeffizientenmatrix $(A|b)$:

$$\left(\begin{array}{ccccc|c} 0 & 0 & 1 & 3 & 3 & 2 \\ 1 & 2 & 1 & 4 & 3 & 3 \\ 1 & 2 & 2 & 7 & 6 & 5 \\ 2 & 4 & 1 & 5 & 3 & 4 \end{array}\right)$$

Umformung zu Stufenform: $z_1 \leftrightsquigarrow z_2$ (Stufenelement muss ungleich Null sein): *"Pivotelement"* (für numerische Berechnungen in $\mathbb{Q}$ ist Wahl wichtig).

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 3 \\ 0 & 0 & 1 & 3 & 3 & 2 \\ 1 & 2 & 2 & 7 & 6 & 5 \\ 2 & 4 & 1 & 5 & 3 & 4 \end{array}\right)$$

$z_3 - z_1 \leftrightsquigarrow z_3$ und $z_4 - 2z_1 \rightsquigarrow z_4$ (1. Spalte unterhalb von Stufe alles zu 0):

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

$z_3 - z_2 \rightsquigarrow z_3$ und $z_4 - (-z_2) \rightsquigarrow z_4$ (3. Spalte unterhalb von Stufe alles zu 0).

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 3 \\ 0 & 0 & 1 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Stufenform $(A'|b')$ erreicht. Wegen $\operatorname{rg}(A) = 2 = \operatorname{rg}(A|b)$ ist LGS lösbar (Proposition 3.3.2).

Lösungen berechnen. Haben freie Parameter $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{K}$.
$x_5 = \lambda_3$
$x_4 = \lambda_2$
$x_3 = 2 - 3\lambda_2 - 3\lambda_3$ (aus dritter Zeile $x_3 + 3x_4 + 3x_5 = 2$)
$x_2 = \lambda_1$
$x_1 = 3 - 2\lambda_1 - (2 - 3\lambda_2 - 3\lambda_3) - 4\lambda_2 - 3\lambda_3) = 1 - 2\lambda_1 - \lambda_2$.

Es gilt also:

$$\text{Lös}(A, b) = \left\{ \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & -3 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} \mid \lambda_1, \lambda_2, \lambda_3 \in \mathbb{K}^3 \right\}. \qquad \triangle$$

Bemerkung. Mit dem gaußschen Algorithmus[3] lassen sich folgende Probleme lösen:

1. Enscheiden, ob ein LGS (Abschnitt 3.3.4) eine Lösung besitzt;
2. Bestimmung von Kern A (Abschnitt 3.3.5);
3. Bestimmung von Bild A (Abschnitt 3.3.6);
4. Den Rang einer Matrix berechnen (Abschnitt 3.2.3);
5. $\dim \langle z_1, \ldots, z_m \rangle$ berechnen (Abschnitt 3.2.5);
6. Basis von $\langle z_1, \ldots, z_m \rangle$ ausrechnen (Abschnitt 3.2.5);
7. Bestimmung der Determinante von A (später in Abschnitt 4.1).

3.3.5 Bestimmung des Kerns

Sei A' die Zeilen-Stufenform von einer Matrix $A \in \mathbb{K}^{m \times n}$. Dann gilt

$$\operatorname{Kern} A = \text{Lös}(A, \mathbf{0}) = \text{Lös}(A', \mathbf{0}).$$

[3] Das Verfahren war schon vor Gauß in Europa und unabhängig in China bekannt.

Im Beispiel (aus Abschnitt 3.3.4):

$$A' = \begin{array}{c} \begin{array}{ccccc} x_1 & \lambda_1 & x_3 & \lambda_2 & \lambda_3 \end{array} \\ \begin{pmatrix} 1 & 2 & 1 & 4 & 3 \\ 0 & 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}$$

Lösungen des homogenen Systems $A'x = \mathbf{0}$ haben die Form

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2\lambda_1 - \lambda_2 \\ \lambda_1 \\ -3\lambda_2 - 3\lambda_3 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \text{ für } \lambda_1, \lambda_2, \lambda_3 \in K.$$

Wegen $\dim \operatorname{Kern} A = n - \operatorname{rg}(A) = 5 - 2 = 3$ (Satz 3.3.6) müssen für Basis von $\operatorname{Kern} A$ drei linear unabhängige Lösungen gefunden werden. Diese erhält man, wenn man für $(\lambda_1, \lambda_2, \lambda_3)$ eine Basis von $\mathbb{K}^3$ einsetzt, zum Beispiel die Einheitsvektoren $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Also:

$$w_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

ist Basis von $\operatorname{Kern} A$.

3.3.6 Bestimmung des Bilds

Für gegebene Matrix $A \in \mathbb{K}^{m \times n}$ und $b \in \mathbb{K}^m$ wollen wir wissen, ob $b \in \operatorname{Bild} A$. Es seien $s_1, \ldots, s_n$ Spalten von $A \in \mathbb{K}^{m \times n}$. Dann: $\operatorname{Bild} A = \langle s_1, \ldots, s_n \rangle$ (siehe Abschnitt 3.3.3). Ein Vektor $b \in \mathbb{K}^m$ ist also genau dann im Bild von A, wenn $Ax = b$ eine Lösung besitzt.

Ziel: Berechnung einer Basis von $\operatorname{Bild} A \subseteq \mathbb{K}^m$. Idee: Auch hierfür kann das Verfahren aus Abschnitt 3.2.5 (Umformung in Stufenform) verwendet werden.

Definition 3.3.14. Sei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{m \times n}.$$

Die Matrix

$$A^\top := \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{K}^{n \times m}$$

heißt *transponierte Matrix* von A (Zeilen und Spalten vertauscht).
Entspricht Spiegelung an der Diagonalen: ($\diagdown$).

Beispiel 3.3.15.

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

(Zeilenvektor)$^\top$ = (Spaltenvektor):

$$\begin{pmatrix} 1 & 2 \end{pmatrix}^\top = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

△

Bemerkung 3.3.16. Zu Rechnungen mit der Transposition.

- Offenbar: $(A^\top)^\top = A$.
- $(AB)^\top = B^\top A^\top$. Denn

$$(AB)^\top = \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right)_{ki} = \left(\sum_{j=1} b_{jk} a_{ij} \right)_{ki} = B^\top A^\top$$

- Ist $A \in \mathbb{K}^{n \times n}$ invertierbar, so gilt $(A^{-1})^\top = (A^\top)^{-1}$. Denn:

$$A^\top \cdot (A^{-1})^\top = (A^{-1} A)^\top = E^\top = E$$

Beispiel 3.3.17. Beispiel 3.3.13 aus Abschnitt 3.3.4:

$$A := \begin{array}{c} \begin{array}{ccccc} s_1 & s_2 & s_3 & s_4 & s_5 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 3 & 3 \\ 1 & 2 & 1 & 4 & 3 \\ 1 & 2 & 2 & 7 & 6 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix} \end{array}$$

Spalten von A als Zeilen der Matrix

$$A^\top = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 0 & 2 & 2 & 4 \\ 1 & 1 & 2 & 1 \\ 3 & 4 & 7 & 5 \\ 3 & 3 & 6 & 3 \end{pmatrix} \begin{array}{c} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array}$$

Auf Zeilen-Stufenform bringen:

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 3 & 4 & 7 & 5 \\ 3 & 3 & 6 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Basis von Bild A:

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

Wegen Bild $A = \operatorname{rg}(A) = 2$ (Abschnitt 3.3.4) braucht man nur zwei linear unabhängige Spalten von A zu finden, z.B.

$$s_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} \text{ und } s_5 = \begin{pmatrix} 3 \\ 3 \\ 6 \\ 3 \end{pmatrix}. \qquad \triangle$$

3.3.7 Unlösbarkeitskriterium

Wenn $Ax = b$ für $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$ eine Lösung hat, kann man das einfach durch Angabe einer Lösung in $\mathbb{K}^n$ beweisen. Gibt es auch einfache Beweise dafür, dass $Ax = b$ *keine* Lösung hat? Etwas einfacheres, als den Gaußschen Algorithmus durchzuführen?

Satz 3.3.18 (Dualität)**.** *Sei $A \in \mathbb{K}^{m\times n}$ und $b \in \mathbb{K}^m$. Dann ist das LGS $Ax = b$ genau dann unlösbar, wenn das System*

$$(A|b)^\top y = (\mathbf{0}|1)^\top \tag{3.8}$$

lösbar ist.

Beweis. Seien $z_1, \dots, z_m$ die Zeilen von $(A|b)$. Angenommen, $(A|b)^\top y = (\mathbf{0}|1)^\top$ hat eine Lösung $y \in K^m$. Das bedeutet, $y_1 z_1 + \cdots + y_m z_m = (\mathbf{0}|1)$. Also $(\mathbf{0}|1) \in \langle z_1, \dots, z_m \rangle$. Das bedeutet, man kann die Zeile $(\mathbf{0}|1) = (0 \cdots 0\ 1)$ mit elementaren Zeilenumformungen aus $Ax = b$ herleiten. Diese Zeile entspricht der unerfüllbaren Gleichung $0x_1 + \cdots + 0x_n = 1$. Dann ist auch $Ax = b$ unerfüllbar.

Die andere Richtung in Satz 3.3.18 ist etwas schwieriger zu zeigen, allerdings nicht viel, wenn man die Stufenform kennt. Wir überführen mit Hilfe von elementaren Zeilenumformungen die Matrix $(A|b)$ in eine Matrix $(C|d)$ in Stufenform. Wenn nun $(A|b)$ unerfüllbar ist, dann auch $(C|d)$, und $r := \operatorname{rg}(C) < \operatorname{rg}(C|d)$ nach Satz 3.2.11. Es gilt dann insbesondere $z_{r+1} = (\mathbf{0}|d_{r+1})$ mit $d_{r+1} \in \mathbb{K} \setminus \{0\}$. Ersetze z_{r+1} durch $d_{r+1}^{-1} z_{r+1} = (\mathbf{0}|1)$. Da diese Zeile durch elementare Zeilenumformungen aus $(A|b)$ hervorgegangen ist, ist also $(\mathbf{0}|1) \in \langle z_1, \dots, z_m \rangle$. Daher ist das System $(A|b)^\top y = (\mathbf{0}|1)^\top$ lösbar. $\square$

3.4 Lineare Abbildungen II

Lineare Abbildungen waren bereits Gegenstand von Abschnitt 3.1. In diesem frühen Abschnitt jedoch haben uns die Werkzeuge gefehlt, um wirklich interessante Aussagen über lineare Abbildungen formulieren zu können. Dies ist jetzt anders, weshalb wir dieses Thema hier nochmal aufgreifen.

3.4.1 Beispiele

1. Die identische Abbildung $\mathrm{id}_V : V \to V$ ist linear (ein Endomorphismus).
2. Die Nullabbildung $V \to V : v \mapsto \mathbf{0}$ stets linear.
3. Allgemeiner: für $\lambda \in K$ ist $x \mapsto \lambda x$ linear.
4. Die Abbildung $f : \mathbb{C} \to \mathbb{C} : z \mapsto i \cdot z$ ist linear (betrachten $\mathbb{C}$ als $\mathbb{R}$-Vektorraum).

   $$f : \mathbb{R}^2 \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -y \\ x \end{pmatrix}$$

   Geometrische Interpretation: Drehung um 90 Grad.
   Es gilt z.B. $f(v + v') = f(v) + f(v')$ und $f(2u) = 2f(u)$.

Wichtiges Beispiel:

Proposition 3.4.1. *Sei $A \in \mathbb{K}^{m \times n}$. Dann ist*

$$f_A : \mathbb{K}^n \to \mathbb{K}^m : x \mapsto Ax$$

eine lineare Abbildung.

Beispiel 3.4.2. $n = m = 2$, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$:

$$f_A : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$ △

Beispiel 3.4.3. $n = 3$, $m = 2$, $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$,

$$f_A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + 2x_3 \end{pmatrix}$$

$$f_A(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad f_A(e_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad f_A(e_3) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$ △

Übung 11. Es sei $\mathbb{K}$ ein Körper. Beweisen Sie, dass eine Abbildung $f : \mathbb{K} \to \mathbb{K}$ genau dann linear ist, wenn es ein $\lambda \in K$ gibt, so dass für alle $v \in K$ gilt, dass $f(v) = \lambda v$.

3.4.2 Beschreibung linearer Abbildungen

Im gesamten Abschnitt stehen V und W für zwei $\mathbb{K}$-Vektorräume. Weiterhin seien $v_1, \dots, v_n \in V$ und $f\colon V \to W$ eine lineare Abbildung. Dann gilt:

$$f[\langle\{v_1, \dots, v_n\}\rangle] = \langle\{f(v_1), \dots, f(v_n)\}\rangle \tag{3.9}$$

Denn:

$$f(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n)$$

Insbesondere:

$$f[V] \leq W \qquad \text{(Proposition 2.4.2)} \tag{3.10}$$

Satz 3.4.4. *Sei $(v_1, \dots, v_n)$ eine Basis von V. Dann gibt es zu jedem n-Tupel $(w_1, \dots, w_n) \in W^n$ genau eine lineare Abbildung $f\colon V \to W$ mit $v_i \mapsto w_i$ für alle $i \in \{1, \dots, n\}$.*

Damit ist jede lineare Abbildung f eindeutig festgelegt, wenn man die Bilder $f(v_i)$ einer Basis kennt. Die Bilder der Basiselemente sind beliebig wählbar.

Beweis von Satz 3.4.4. Jedes $v \in V$ lässt sich eindeutig schreiben als $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in V$ (Satz 2.4.6). Wir definieren dann (wohldefiniert!)

$$f(v) := \lambda_1 w_1 + \dots + \lambda_n w_n$$

Diese Abbildung ist linear, und $f(v_i) = w_i$ für alle $i \in \{1, \dots, n\}$. Um die Eindeutigkeit dieser Abbildung nachzuweisen, sei f' eine beliebige lineare Abbildung mit $f(v_i) = w_i$ für alle $i \in \{1, \dots, n\}$. Dann gilt

$$\begin{aligned} f'(v) &= f'(\lambda_1 v_1 + \dots + \lambda_n v_n) \\ &= \lambda_1(f(v_1) + \dots + \lambda_n(f(v_n)) \\ &= \lambda_1(w_1) + \dots + \lambda_n(w_n) = f(v) \end{aligned}$$

und damit $f' = f$. □

3.4.3 Kern, Bild, Rang, Defekt

Sei $f\colon V \to W$ lineare Abbildung. Dann definieren wir

- $\operatorname{Kern} f := \{v \in V \mid f(v) = \mathbf{0}\}$.
- $\operatorname{Bild} f = f[V] := \{f(v) \mid v \in V\}$ (siehe Abschnitt 1.2.2).
- $\operatorname{rg} f := \dim(\operatorname{Bild} f)$ (Definition sinnvoll wegen $\operatorname{Bild} f \leq W$, siehe (3.10))
- $\operatorname{dfkt} f := \dim(\operatorname{Kern} f)$ der *Defekt* von f.

Bemerkung 3.4.5. Begriffe für Matrizen A stimmen mit denen für die zugehörige lineare Abbildung $f_A\colon x \mapsto Ax$ überein (siehe Abschnitt 3.3.3):

- $\operatorname{Kern} A = \operatorname{Kern} f_A$;
- $\operatorname{Bild} A = \operatorname{Bild} f_A$;
- $\operatorname{rg} A = \operatorname{rg} f_A$.

Satz 3.4.6. *Sei $f\colon V \to W$ eine lineare Abbildung. Dann:*

1. *f injektiv $\Leftrightarrow$ $\operatorname{Kern} f = \{\mathbf{0}\}$;*
2. *Falls W endlich dimensional, so gilt*
$$f \text{ surjektiv} \Leftrightarrow \operatorname{Bild} f = W \Leftrightarrow \dim \operatorname{Bild} f = \dim W;$$
3. *Falls $\dim V = \dim W = n < \infty$, so gilt f injektiv $\Leftrightarrow$ f surjektiv $\Leftrightarrow$ f bijektiv (also Isomorphismus)*

Beweis. Zu 1.

$$\begin{aligned} f(v) = f(v') &\Leftrightarrow f(v) - f(v') = \mathbf{0} \\ &\Leftrightarrow f(v - v') = \mathbf{0} \\ &\Leftrightarrow v - v' \in \operatorname{Kern} f \end{aligned}$$

Zu 2. Siehe Abschnitt 2.4.5.
Zu 3. (Gilt nur in endlich dimensionalen Vektorräumen.)

$$\begin{aligned} f \text{ injektiv} &\Leftrightarrow \operatorname{Kern} f = \{\mathbf{0}\} \\ &\Leftrightarrow \dim \operatorname{Kern} f = 0 \\ &\Leftrightarrow \dim \operatorname{Bild} f = n - 0 = n && \text{(nach Dimensionsformel, Satz 3.3.6)} \\ &\Leftrightarrow \operatorname{Bild} f = W && \text{(siehe 2.)} \\ &\Leftrightarrow f \text{ surjektiv} && \text{(nach Definition).} \end{aligned}$$

□

Satz 3.4.7. *Sei $f\colon V \to W$ eine lineare Abbildung, und $v_1, \ldots, v_n$ eine Basis von V. Dann:*

1. *f ist genau dann injektiv wenn $f(v_1), \ldots, f(v_n)$ linear unabhängig;*
2. *f ist genau dann surjektiv wenn $\langle f(v_1), \ldots, f(v_n)\rangle = W$;*
3. *f ist genau dann Isomorphismus wenn $\{f(v_1), \ldots, f(v_n)\}$ eine Basis ist von W.*

Beweis. 1, '$\Rightarrow$': Sei f injektiv und $\lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = \mathbf{0}$. (Z.z.: $\lambda_1 = \cdots = \lambda_n = 0$). Dann:

$$\begin{aligned} & \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = \mathbf{0} \\ \Rightarrow\ & f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \mathbf{0} && \text{(Linearität von } f\text{)} \\ \Rightarrow\ & \lambda_1 v_1 + \cdots + \lambda_n v_n \in \operatorname{Kern} f \\ \Rightarrow\ & \lambda_1 v_1 + \cdots + \lambda_n v_n = \mathbf{0} && \text{(Satz 3.4.6 (1.), da } f \text{ injektiv)} \\ \Rightarrow\ & \lambda_1 = \cdots = \lambda_n = 0 && \text{(da } v_1, \ldots, v_n \text{ linear unabhängig)} \end{aligned}$$

'$\Leftarrow$': Seien $f(v_1), \dots, f(v_n)$ linear unabhängig und $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in V$. Dann:

$$\begin{aligned}
& f(v) = \mathbf{0} \\
\Rightarrow\ & v \in \operatorname{Kern} f \\
\Rightarrow\ & f(v) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = \mathbf{0} \\
\Rightarrow\ & \lambda_1 = \dots = \lambda_n = 0 \qquad\qquad (\text{da } v_1, \dots, v_n \text{ linear unabhängig}).
\end{aligned}$$

Also ist $v = \mathbf{0}$ und f ist injektiv nach Satz 3.4.6 (1.).

2. f ist nach Definition genau dann surjektiv wenn $\operatorname{Bild} f = W$. Man rechnet leicht nach dass $\operatorname{Bild} f \leq W$. Da $\langle f(v_1), \dots, f(v_n) \rangle$ der kleinste Untervektorraum von W der $f(v_1), \dots, f(v_n)$ enthält, ist f genau dann surjektiv wenn $W = \langle f(v_1), \dots, f(v_n) \rangle$.

3. Direkt aus 1. und 2. □

Satz 3.4.8 (Fundamentalsatz für endlich dimensionale Vektorräume)**.** *Je zwei $\mathbb{K}$-Vektorräume, die von n Elementen erzeugt werden, sind isomorph. Insbesondere ist jeder n-dimensionale $\mathbb{K}$-Vektorraum V isomorph zu $\mathbb{K}^n$, geschrieben $V \simeq \mathbb{K}^n$.*

Im Prinzip könnte man sich also beim Studium von endlichdimensionalen Vektorräumen auf $\mathbb{K}^n$ beschränken; das wäre aber unpraktisch, da viele Vektorräume ganz anders angegeben sind. Trotzdem ist es eine wichtige Einsicht, dass Rechnen mit Koordinaten (nach Wahl einer Basis!) möglich ist.

Beweis. Sei $(v_1, \dots, v_n)$ eine Basis von V, und $(w_1, \dots, w_n)$ eine Basis von W (Satz 2.4.10: Basen existieren). Nach Satz 3.4.4 gibt es eine lineare Abbildung f mit $f\colon v_i \mapsto w_i$, und f is Isomorphismus gemäß Satz 3.4.6. □

Sei $B = (v_1, \dots, v_n)$ Basis von $\mathbb{K}$-Vektorraum V, und $(e_1, \dots, e_n)$ die *kanonische* Basis von $\mathbb{K}^n$. Nach Satz 3.4.4 gibt es einen eindeutig bestimmten Isomorphismus

$$\Phi_B\colon \mathbb{K}^n \to V \text{ mit } \Phi_B(e_i) = v_i$$

Dieser heißt der *kanonische Basisisomorphismus*. Beschreibt den Zusammenhang zwischen Vektoren und ihren Koordinatenvektoren

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \lambda_1 v_1 + \dots + \lambda_n v_n$$

Aussagen über V in gleichwertige Aussagen über Elemente von $\mathbb{K}^n$ umwandeln.

Beispiel: Bereits in Abschnitt 3.2.5.

$$\begin{aligned}
(u_1, \dots, u_n) \text{ Basis von } \mathbb{K}^n &\Leftrightarrow (\Phi_B(u_1), \dots, \Phi_B(u_n)) \text{ Basis von } V \\
(w_1, \dots, w_n) \text{ Basis von } V &\Leftrightarrow (\Phi_B^{-1}(w_1), \dots, \Phi_B^{-1}(w_n)) \text{ Basis von } \mathbb{K}^n
\end{aligned}$$

3.4.4 Faktorräume

Sei V ein $\mathbb{K}$-Vektorraum, und $U \leq V$ Untervektorraum. Für $v \in V$ heißt

$$v + U := \{v + u \mid u \in U\}$$

Nebenklasse von v bezüglich U. Das Element v heißt *Repräsentant* von $v+U$. Die Menge der Nebenklassen

$$V/U := \{v + U \mid v \in V\}$$

heißt *Faktorraum* von V nach U. Wie wir sehen werden, ist der Faktorraum auch wieder ein $\mathbb{K}$-Vektorraum.

Bemerkung 3.4.9. Es gilt

$$\begin{aligned} w \in v + U &\Leftrightarrow (w = v + u \text{ für ein } u \in U) \\ &\Leftrightarrow (v - w = u' \text{ für ein } u' \in U) \\ &\Leftrightarrow v - w \in U \\ &\Leftrightarrow v \in w + U \\ &\Leftrightarrow v + U = w + U \end{aligned} \tag{3.11}$$

Das heißt, jedes Element einer Nebenklasse ist Repräsentant dieser Nebenklasse. Durch

$$v \sim w :\Leftrightarrow u + U = w + U$$

ist eine Äquivalenzrelation definiert (siehe Abschnitt 1.2.1), die Äquivalenzklassen sind gerade die Nebenklassen:

$$[v]_\sim = v + U$$

Abschnitt 1.2.1 besagt, dass $V/U = V/\sim$ eine Zerlegung von V in disjunkte Nebenklassen ist. Bild malen!

Übung 12. Beweisen Sie die Behauptungen in Bemerkung 3.4.9.

Beispiel 3.4.10. $V := \mathbb{R}^2$, $U \leq V$ sei die Gerade durch $\mathbf{0}$ mit Richtungsvektor v. Durch jeden Punkt $p \in \mathbb{R}^2$ gibt es genau eine Gerade g' parallel zu g:

$$[p] := g' = p + \mathbb{R}v = p + U$$

Die Gerade g' hängt nicht von der Auswahl des Repräsentanten ab: $g' = [p] = [q]$ genau dann, wenn p und p' auf der gleichen Geraden g' parallel zu g liegen. Der Faktorraum V/U ist die Schar der zu g parallelen Gerade $p + U$. △

Auf der Menge der Geraden lässt sich folgende Vektorraumstruktur definieren:

$$\begin{aligned} [p] + [q] &:= [p + q] \\ \lambda[p] &:= [\lambda p] \end{aligned}$$

Das funktioniert, weil die Gerade $[p+q]$ nur von den Geraden $[p]$ und $[q]$ abhängt, aber nicht von der konkreten Wahl der Repräsentanten $p, q \in \mathbb{R}^2$.

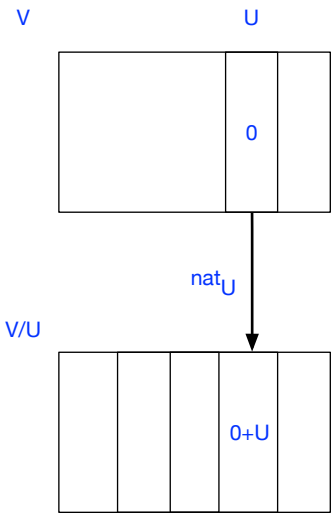
Satz 3.4.11. *Sei V ein $\mathbb{K}$-Vektorraum und $U \leq V$. Dann ist V/U zusammen mit den folgenden Operationen ein $\mathbb{K}$-Vektorraum (der* Faktorraum*):*

- *Addition:*
$$(v + U) + (w + U) := (v + w) + U$$
- *Multiplikation mit Skalar $\lambda \in K$:*
$$\lambda \cdot (v + U) := (\lambda v) + U$$

Dabei ist

- Nullvektor $\mathbf{0}_{V/U} = \mathbf{0}_V + U = U$
- additives inverses Element $-(v + U) = (-v) + U$.

Siehe Abbildung rechts.



Beweisskizze: + und Multiplikation mit Skalar sind wohldefiniert: die Definition hängt nicht von der Wahl der Repräsentanten ab.

Zu zeigen: wenn $v + U = v' + U$, $w + U = w' + U$, dann $(v' + w') + U = (v + w) + U$.
Nach (3.11) gilt $v' \in v + U$ und $w' \in w + U$. Also

$$\begin{aligned} v' + w' &\in (v + U) + (w + U) \\ &= v + w + U + U \\ &= (v + w) + U. \end{aligned}$$

Und mit (3.11) gilt $(v' + w') + U = (v + w) + U$.

Analog für skalare Multiplikation.

Die Vektorrauaxiome übertragen sich damit von den Axiomen für die Repräsentanten $v \in V$ auf die Nebenklassen $v + U \in V/U$. □

Proposition 3.4.12. *Die Abbildung* $\mathrm{nat}_U \colon V \to V/U$ *gegeben durch $v \mapsto v + U$ ist eine lineare Abbildung, mit* $\mathrm{Kern}(\mathrm{nat}_U) = U$*, und heißt* natürliche Homomorphismus*.*

Beweis. Linearität folgt aus Definition von V/U.

$$\begin{aligned} \mathrm{Kern}(\mathrm{nat}_U) &= \{v \in V \mid \mathrm{nat}_U(v) = \mathbf{0}_{V/U}\} \\ &= \{v \in V \mid v + U = U\} \\ &= \{v \in V \mid v \in U\} \\ &= U \end{aligned}$$

□

Satz 3.4.13 (Homomorphiesatz)**.** *Es sei $f\colon V \to W$ eine lineare Abbildung (Homomorphismus) und $U := \operatorname{Kern} f$. Dann gilt*

$$\operatorname{Bild} f \simeq V/U$$

Ein Isomorphismus ist gegeben durch

$$h\colon V/U \to \operatorname{Bild} f : (v + U) \mapsto f(v)$$

Beweis. Achtung! Definition darf nicht vom Repräsentanten abhängen.

- h ist wohldefiniert: Seien $u, v' \in V$ beliebig so dass $v + U = v' + U$. Zu zeigen ist, dass $f(v) = f(v')$. Es gilt $v' \in v + U$ und daher gibt es ein $u \in U$ mit $v' = v + u$. Da $u \in U = \operatorname{Kern} f$ gilt dann

$$f(v') = f(v + u) = f(v) + f(u) = f(v) + \mathbf{0} = f(v)$$

- h ist linear:

$$\begin{aligned} h((v + U) + (w + U)) &= h((v + w) + U) \\ &\overset{\text{Def.}}{=} f(v + w) = f(v) + f(w) \\ &= h(v + U) + h(w + U) \end{aligned}$$

- Multiplikation mit Skalar: analog.

- h ist surjektiv: Nach Definition gilt für beliebiges $f(v) \in \operatorname{Bild} f$

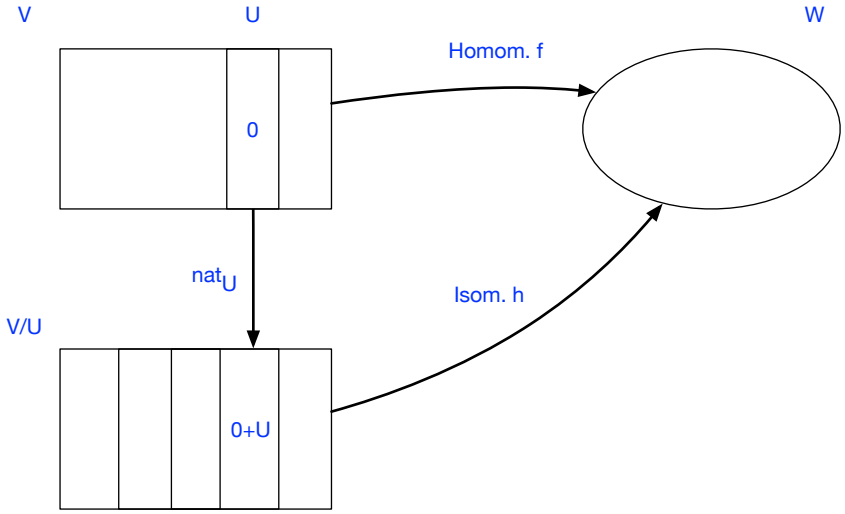
$$h(v + UV) = f(v) \in \operatorname{Bild} h\,.$$

- h ist injektiv: Nach Satz 3.4.6 genau dann, wenn $\operatorname{Kern} h = \{\mathbf{0}\}$.

$$\begin{aligned} h(v + U) = \mathbf{0} &\Leftrightarrow f(v) = \mathbf{0} \\ &\Leftrightarrow v \in \operatorname{Kern} f = U \\ &\Leftrightarrow v + U = U = \mathbf{0}_{V/U} \end{aligned}$$

□

Also: homomorphe Bilder entsprechen Faktorräumen V/U, gegeben durch Unterräume U von V. Siehe Abbildung.



Wie lässt sich eine Basis eines Faktorraums finden?

Lemma 3.4.14. *Sei $U \leq V$ mit Basis $(v_1, \ldots, v_d)$ und $(v_1, \ldots, v_d, v_{d+1}, \ldots, v_{d+r})$ (ergänzte) Basis von V. Dann ist $(v_{d+1} + U, \ldots, v_{d+r} + U)$ Basis von V/U.*

(Siehe Abschnitt 2.4.4 zum Austauschsatz von Steinitz.)

Folgerungen:

$$V/U \simeq \langle v_{d+1}, \ldots, v_{d+r} \rangle \qquad \text{(gleiche Dimension und Satz 3.4.8)}$$
$$V = U \oplus \langle v_{d+1}, \ldots, v_{d+r} \rangle$$

Beweis. Basis: 1) Erzeugendensystem: Sei $v \in V$ beliebig. Da $(v_1, \ldots, v_d)$ Basis lässt sich v schreiben als $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Dann:

$$\begin{aligned} v + U &= \operatorname{nat}_U(v) = \sum_{i=1}^{n} \lambda_i \operatorname{nat}_U(v_i) \\ &= \mathbf{0}_{V/U} + \sum_{i=d+1}^{n} \lambda_i \operatorname{nat}_U(v_i) \qquad \text{(weil } v_i \in U = \operatorname{Kern}(\operatorname{nat}_U) \text{ für } i \leq d) \\ &= \lambda_{d+1}(v_{d+1} + U) + \cdots + \lambda_{d+r}(v_{d+r} + U) \end{aligned}$$

2) Lineare Unabhängigkeit: Sei $\sum_{i=d+1}^{n} \lambda_i (v_i + U) = \mathbf{0}_{V/U} = U$. Da

$$\sum_{i=d+1}^{n} \lambda_i (v_i + U) = \sum_{i=d+1}^{n} \lambda_i \operatorname{nat}_U(v_i) = \operatorname{nat}_U\left(\sum_{i=d+1}^{n} \lambda_i v_i\right)$$

gilt also:

$$\begin{aligned}
&\mathrm{nat}_U\Big(\sum_{i=d+1}^{n} \lambda_i v_i\Big) = \mathbf{0}_{V/U} \\
\Rightarrow\ &\sum_{i=d+1}^{n} \lambda_i v_i \in \mathrm{Kern}(\mathrm{nat}_U) = U \\
\Rightarrow\ &\sum_{i=d+1}^{n} \lambda_i v_i = \lambda_1 v_1 + \cdots + \lambda_d v_d && \text{für geeignete } \lambda_1, \ldots, \lambda_d \in \mathbb{K} \\
\Rightarrow\ &\left(\sum_{i=d+1}^{n} \lambda_i v_i\right) - \lambda_1 v_1 - \cdots - \lambda_d v_d = \mathbf{0} \\
\Rightarrow\ &\lambda_1 = \cdots = \lambda_d = \lambda_{d+1} = \cdots = \lambda_n = 0 && (v_1, \ldots, v_n \text{ linear unabhängig}).
\end{aligned}$$

□

Satz 3.4.15 (Dimensionssatz)**.** *Sei $U \leq V$ endlichdimensional. Dann gilt*

$$\dim V/U = \dim V - \dim U$$

(Dimension verhält sich wie Logarithmus $\log(a/b) = \log(a) - \log(b)$.)
Sei $f\colon V \to W$ lineare Abbildung. Dann gilt

$$\dim(\mathrm{Bild}\, f) = \dim V - \dim(\mathrm{Kern}\, f) \tag{3.12}$$

Beweis. Der erste Teil folgt mit $r = \dim(V/U)$, $n = \dim V$, und $d = \dim U$ direkt aus Lemma 3.4.14. Der zweite Teil folgt aus dem ersten Teil und dem Homomorphiesatz: $\dim(\mathrm{Bild}\, f) = \dim(V/\mathrm{Kern}\, f)$. Setze $U := \mathrm{Kern}\, f$. □

Korollar 3.4.16. *Seien $f\colon U \to V$ und $g\colon V \to W$ lineare Abbildungen zwischen endlichdimensionalen Vektorräumen. Dann gilt*

$$\mathrm{rg}(g \circ f) \leq \min(\mathrm{rg}(g), \mathrm{rg}(f)).$$

Beweis. Offensichtlich gilt: $\mathrm{Bild}(g \circ f) \subseteq \mathrm{Bild}(g)$. Also

$$\mathrm{rg}(g \circ f) = \dim(\mathrm{Bild}(g \circ f)) \leq \dim(\mathrm{Bild}(g)) = \mathrm{rg}(g).$$

Auch gilt

$$\begin{aligned}
&\mathrm{rg}(g \circ f) = \dim(\mathrm{Bild}(g \circ f)) \\
&= \dim V - \dim(\mathrm{Kern}(g \circ f)) && \text{(Dimensionsformel)} \\
&= \dim(\mathrm{Bild}(f)) + \dim(\mathrm{Kern}(f)) - \dim(\mathrm{Kern}(g \circ f)) && \text{(Dimensionsformel)} \\
&\leq \dim(\mathrm{Bild}(f)) = \mathrm{rg}(f) && (\mathrm{Kern}(f) \subseteq \mathrm{Kern}(g \circ f)).
\end{aligned}$$

□

Übung 13. Es sei $f\colon V \to W$ eine lineare Abbildung. Sei $U \leq V$ so, dass V/U isomorph ist zu $\mathrm{Bild}(f)$. Zeigen Sie, dass dann $U = \mathrm{Kern}(f)$.[4]

[4] Übungsaufgabe inspiriert von Student:innenfrage WS'23/24.

3.4.5 Lineare Abbildungen und Matrizen

Jede Matrix $A \in \mathbb{K}^{m\times n}$ beschreibt eine lineare Abbildung, nämlich

$$f_A\colon \mathbb{K}^n \to \mathbb{K}^m : x \mapsto Ax$$

Umgekehrt gilt: *jede* lineare Abbildung lässt sich durch eine Matrix beschreiben.
Sei $f\colon V \to W$ eine lineare Abbildung, wobei:

- V ein n-dimensionaler Vektorraum mit Basis $B = (v_1, \dots, v_n)$, und
- W ein m-dimensionaler Vektorraum mit Basis $C = (w_1, \dots, w_m)$.

Nach Satz 3.4.4 ist f eindeutig durch die Bilder der Basisvektoren festgelegt:

$$f(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m = \sum_{i=1}^{m} a_{ij}w_i$$

Das heißt, f wird eindeutig festgelegt durch die Matrix

$$A := M_C^B(f) := \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \tag{3.13}$$

die sogenannte *Darstellungsmatrix*.

> **Merkregel:** Die Spalten der zu f gehörigen Darstellungsmatrix $M_C^B(f)$ sind die Koordinatenvektoren (bezüglich der Basis C von W) der Bilder von B (der Basisvektoren von V) unter $f\colon V \to W$.

Dabei gilt: Hat $v \in V$ den Koordinatenvektor $u = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ (bezüglich B), so hat $f(v)$ den Koordinatenvektor Au (bezüglich C) für $A = M_C^B(f)$, d.h., die lineare Abbildung

$$f_A\colon \mathbb{K}^n \to \mathbb{K}^m : u \mapsto Au$$

beschreibt die lineare Abbildung f in ihren Koordinatenvektoren:

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{K}^n \arrow[u, "\Phi_B"'] \arrow[r, "A:=M_C^B(f)"] & \mathbb{K}^m \arrow[u, "\Phi_C"'] \end{tikzcd}$$

Ein sogenanntes *kommutatives Diagram*. Konkret:

$$\begin{tikzcd}
v = \lambda_1 v_1 + \cdots + \lambda_n v_n \arrow[r, "f"] & f(v) = \Phi_C(Au) \\
u = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \arrow[u, "\Phi_B"'] \arrow[r, "f_A"] & Au \arrow[u, "\Phi_C"']
\end{tikzcd}$$

Denn:

$$\begin{aligned}
f(\Phi_B(u)) = f(v) &= \sum_{j=1}^{n} \lambda_j f(v_j) \\
&= \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} w_i \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \lambda_j \right) w_i = \Phi_C(Au)
\end{aligned}$$

Beispiel 3.4.17. $V := \mathbb{K}^n$, $W := \mathbb{K}^m$. Standardbasen $B_n := (e_1, \ldots, e_n)$ von V und B_m von W. Der kanonische Basisisomorphismus $\Phi_{B_n} \colon \mathbb{K}^n \to \mathbb{K}^n : v \mapsto v$ ist die identische Abbildung, also ist jede lineare Abbildung

$$f \colon \mathbb{K}^n \to \mathbb{K}^m$$

darstellbar als

$$f_A \colon \mathbb{K}^n \to \mathbb{K}^m : u \mapsto Au$$

mit $A = M_{B_m}^{B_n}(f)$. Hier: Spalten von A sind die Bilder der Einheitsvektoren.

Beispiel zum Beispiel: $f \colon \mathbb{R}^3 \to \mathbb{R}^2$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

(Projektion auf xy-Ebene) ist linear, zugehörige Matrix

$$A = M_{B_2}^{B_3}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Spalten von A sind Bilder der Koordinatenvektoren, $A = (f(e_1) f(e_2) f(e_3))$. △

Beispiel 3.4.18. Für die identitsche Abbildung $\mathrm{id}_V \colon V \to V$ gilt (nur eine Basis, $B = C$)

$$M_B^B(\mathrm{id}_V) = E_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Aber für $B \neq C$ entsteht *nicht* E_n! △

Beispiel 3.4.19. Die Polynome

$$a_0 + a_1X + a_2X^2 + a_3X^3$$

in einer Unbekannten X vom Grad höchstens drei mit reellen Koeffizienten bilden einen $\mathbb{R}$-Vektorraum. Was ist ein Polynom? Aus der Schule bekannt. Ein *Ausdruck* geformt mit Hilfe von Variable, Skalaren, Addition, und Multiplikation.

$$V := \mathbb{R}^{(\leq 4)}[X]$$

Addition und Multiplikation mit Skalar $\lambda \in \mathbb{R}$ wie üblich.
Basis ist z.B. $B = (v_0, v_1, v_2, v_3)$ mit $v_0 = 1$, $v_1 = X$, $v_2 = X^2$, $v_3 = X^3$.
Kanonischer Basisisomorphismus:

$$\Phi_B\colon \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto a_0v_0 + a_1v_1 + a_2v_2 + a_3v_3 = a_0 + a_1X + x_2X^2 + a_3X^3$$

Das heißt $\Phi_B(e_i) = v_{i-1}$ für $i \in \{1, \dots, 4\}$. Das Differenzieren

$$\mathrm{diff}\colon \mathbb{R}^{(\leq 4)}[X] \to \mathbb{R}^{(\leq 4)}[X]$$

ist lineare Abbildung mit Matrix

$$A = M_B^B(\mathrm{diff}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{tikzcd}
a_0 + a_1X + a_2X^2 + a_3X^3 \arrow[r, "\mathrm{diff}"] & a_1 + 2a_2X + 3a_3X^3 = \Phi_B(Au) \arrow[d, "\Phi_B^{-1}"] \\
u = \begin{pmatrix} a_0 \\ \vdots \\ a_3 \end{pmatrix} \arrow[u, "\Phi_B"] \arrow[r, "f_A"] & Au = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{pmatrix}
\end{tikzcd}$$

△

Proposition 3.4.20. *Die Komposition von linearen Abbildungen entspricht dem Produkt der zugehörigen Matrizen: V_1, V_2, V_3 seien $\mathbb{K}$-Vektorräume mit Basen B_1, B_2, B_3 und $f_1\colon V_1 \to V_2$, $f_2\colon V_2 \to V_3$ lineare Abbildungen.*

$$\begin{tikzcd}
V_1 \arrow[r, "f_1"] & V_2 \arrow[r, "f_2"] & V_3 \\
\mathbb{K}^n \arrow[u, "\Phi_{B_1}"] \arrow[r, "f_{A_1}"] & \mathbb{K}^m \arrow[u, "\Phi_{B_2}"] \arrow[r, "f_{A_2}"] & \mathbb{K}^r \arrow[u, "\Phi_{B_3}"]
\end{tikzcd}$$

Dann gilt

$$M_{B_3}^{B_1}(f_2 \circ f_1) = M_{B_3}^{B_2}(f_2) M_{B_2}^{B_1}(f_1)$$

(Funktionenkomposition) *(Matrizenmultiplikation)*

Folgerung für Inverse:

$$M_B^C(f^{-1}) = (M_C^B(f))^{-1}$$

Bemerkung 3.4.21. Die linearen Abbildungen $f\colon V \to W$ bilden selbst einen Vektorraum, Bezeichnung:

$$\mathrm{Hom}_{\mathbb{K}}(V, W)$$

Operationen komponentenweise:

$$(f + g)(v) := f(v) + g(v)$$
$$(\lambda f)(v) := \lambda \cdot f(v)$$

Falls $\dim V = n$ und $\dim W = m$ so gilt:

$$\mathrm{Hom}(V, W) \cong \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m) \cong \mathbb{K}^{m \times n}$$

3.4.6 Basiswechsel und Koordinatentransformation

Sei V ein $\mathbb{K}$-Vektorraum, $B = (v_1, \ldots, v_n)$ eine Basis von V, und $B' = (v_1', \ldots, v_n')$ eine andere Basis von V. Bei *Basiswechsel* $B \rightsquigarrow B'$ ändern sich die Koordinatenvektoren eines Vektors $v \in V$: die *Koordinatentransformation* bei einem Basiswechsel wird beschrieben durch die *Transformationsmatrix*

$$T := M_{B'}^{B}(\mathrm{id}_V)$$

oder

$$S := M_B^{B'}(\mathrm{id}_V) = T^{-1}.$$

Ist

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \Phi_B^{-1}(v)$$

der Koordinatenvektor von $v \in V$ bzgl. B, d.h., $v = x_1 v_1 + \cdots + x_n v_n$ und

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

der Koordinatenvektor bzgl. B', d.h., $v' = x_1 v_1' + \cdots + x_n v_n'$, dann gilt

$$x' = Tx \text{ und } x = Sx'$$

$$\begin{tikzcd}
\mathbb{K}^n \arrow[dr, "\Phi_B"] \arrow[dd, bend right, "S"'] \arrow[dd, bend left, "T"] & \\
& V \\
\mathbb{K}^n \arrow[ur, "\Phi_{B'}"'] &
\end{tikzcd}$$

Beispiel 3.4.22. $V = \mathbb{R}^2$.

$$B := (e_1, e_2), B' := (w_1, w_2), w_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w_2 := \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Dann: (siehe Merkregel, (3.13))

$$S = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = M_B^{B'}(\mathrm{id}_V)$$
$$T = S^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = M_{B'}^{B}(\mathrm{id}_V)$$

Koordinatentransformation: Koordinatenvektor von $v \in V$

$$\text{bzgl. Basis } B = (e_1, e_2) \qquad \text{bzgl. Basis } B = (w_1, w_2)$$
$$x \mapsto Tx = S^{-1}x$$
$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$
$$e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

△

3.4.7 Transformationsformel für Matrizen einer linearen Abbildung

B_1, B_2: Basen eines $\mathbb{K}$-Vektorraumes V.
C_1, C_2: Basen eines $\mathbb{K}$-Vektorraumes W.
$f\colon V \to W$ lineare Abbildung.

Wie hängen $A_1 := M_{C_1}^{B_1}(f)$ und $A_2 := M_{C_2}^{B_2}(f)$ zusammen?

Wegen Proposition 3.4.20 und da

$$f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$$

gilt dass

$$M_{C_2}^{B_2}(f) = M_{C_2}^{C_1}(\mathrm{id}_w) \circ M_{C_1}^{B_1}(f) \circ M_{B_1}^{B_2}(\mathrm{id}_w) \tag{3.14}$$
$$A_2 = \tilde{S}^{-1} A_1 S \tag{3.15}$$

wobei $\tilde{S} := M_{C_1}^{C_2}(\mathrm{id}_W)$ und $S := M_{B_1}^{B_2}(\mathrm{id}_V)$ Transformationsmatrizen (siehe Abschnitt 3.4.6).

$$\begin{tikzcd}
\mathbb{K}^n \arrow[dr, "\Phi_{B_1}"] \arrow[dd, bend right, "S"'] \arrow[dd, bend left, "T"] & & \mathbb{K}^m \arrow[dl, "\Phi_{C_1}"'] \arrow[dd, bend right, "\tilde{S}"'] \arrow[dd, bend left, "\tilde{T}"] \\
& V \arrow[r, "f"] & W \\
\mathbb{K}^n \arrow[ur, "\Phi_{B_2}"'] & & \mathbb{K}^m \arrow[ul, "\Phi_{C_2}"]
\end{tikzcd}$$

Spezialfall: $V = W$, $B_1 = C_1$, $B_2 = C_2$. Hier gilt $S = \tilde{S}$ und damit

$$A_2 = S^{-1} A_1 S.$$

3.4.8 Äquivalenz von Matrizen

Es gibt verschiedene wichtige Äquivalenzrelationen auf der Menge der Matrizen.

Definition 3.4.23 (Äquivalenz)**.** Eine Matrix $A \in \mathbb{K}^{m\times n}$ heißt *äquivalent* zu einer Matrix $B \in \mathbb{K}^{m\times n}$ wenn es invertierbare Matrizen $S \in \mathbb{K}^{n\times n}$ und $T \in \mathbb{K}^{m\times m}$ gibt so dass

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

Als kommutatives Diagram:

$$\begin{tikzcd}
\mathbb{K}^n \arrow[r, "f_A"] \arrow[d, "f_S"] & \mathbb{K}^m \arrow[d, "f_T"] \\
\mathbb{K}^n \arrow[r, "f_B"] & \mathbb{K}^m
\end{tikzcd}$$

Definition 3.4.24 (Zeilenäquivalenz)**.** Eine Matrix $A \in \mathbb{K}^{m\times n}$ heißt *zeilenäquivalent* (auch: *links-äquivalent*) zu einer Matrix $B \in \mathbb{K}^{m\times n}$ wenn es *eine* invertierbare Matrix $S \in \mathbb{K}^{m\times m}$ gibt so dass

$$B = SA.$$

Spaltenäquivalenz ist analog definiert, wird aber hier nicht weiter betrachtet.

A und B sind genau dann zeilenäquivalent, wenn man B aus A durch elementare Zeilenoperationen gewinnen kann (dies folgt aus Satz 3.2.27); also ist jede Matrix äquivalent zu einer in Stufenform (Abschnitt 3.2.4).

Definition 3.4.25 (Ähnlichkeit)**.** Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *ähnlich* zu einer Matrix $B \in \mathbb{K}^{n\times n}$ wenn es *eine* invertierbare Matrix $S \in \mathbb{K}^{n\times n}$ gibt so dass

$$B = S^{-1} A S.$$

Bemerkung 3.4.26. Äquivalenz, Zeilenäquivalenz, und Ähnlichkeit sind Äquivalenzrelationen (auf $\mathbb{K}^{m\times n}$ bzw. auf $\mathbb{K}^{n\times n}$). (Vergleiche Abschnitt 1.2.1.) Ähnlichkeit impliziert Äquivalenz, und Zeilenäquivalenz impliziert Äquivalenz. $A \sim B$.

Übung 14. Beweisen Sie die Behauptungen in Bemerkung 3.4.26.

Satz 3.4.27 (Charakterisierung von Äquivalenz)**.** *Für $A_1, A_2 \in \mathbb{K}^{m\times n}$ sind die folgenden Aussagen gleichbedeutend:*

1. *A_1 und A_2 sind äquivalent.*
2. *Es gibt eine lineare Abbildung $f\colon \mathbb{K}^n \to \mathbb{K}^m$, Basen B_1, B_2 von $\mathbb{K}^n$, und Basen C_1, C_2 von $\mathbb{K}^m$ so dass $A_1 = M_{C_1}^{B_1}(f)$ und $A_2 = M_{C_2}^{B_2}(f)$.*
3. $\operatorname{rg}(A_1) = \operatorname{rg}(A_2)$.
4. *Die Matrix A_1 lässt sich durch elementare Zeilen und(!) Spaltenumformungen in die Matrix A_2 umwandeln.*

Beweis. 1. $\Rightarrow$ 2.: Sei $A_2 = \tilde{S}^{-1}A_1S$. Wähle $f := f_{A_1}$. Wähle $B_1 := (e_1, \dots, e_n)$ und $C_1 := (e_1, \dots, e_m)$ Standardbasen. Dann ist $A_1 = M_{C_1}^{B_1}(f_{A_1})$. Für $B_2 := B_1S$ (neue Basis von $V = \mathbb{K}^n$) und $C_2 := C_1\tilde{S}$ (neue Basis von $W = \mathbb{K}^m$) ist

$$M_{B_1}^{B_2}(\mathrm{id}_V) = S \text{ und } M_{C_1}^{C_2}(\mathrm{id}_W) = \tilde{S}$$

also

$$\begin{aligned} A_2 = \tilde{S}^{-1}A_1S &= M_{C_2}^{C_1}(\mathrm{id}_V)M_{C_1}^{B_1}(f_{A_1})M_{B_1}^{B_2}(\mathrm{id}_V) \\ &= M_{C_2}^{B_2}(f_{A_1}) && \text{(siehe (3.14))} \end{aligned}$$

2. $\Rightarrow$ 3.:

$$\operatorname{rg}(A_1) = \operatorname{rg}(f) = \operatorname{rg}(A_2) \qquad \text{(Abschnitt 3.4.3)}$$

3. $\Rightarrow$ 4.: Aus Zeilenstufenform A lässt sich durch elementare Spaltenumformungen die Matrix

$$\begin{pmatrix} E_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

konstruieren, mit $r = \operatorname{rg}(A)$. Gilt sowohl für A_1 als auch für A_2 da $\operatorname{rg}(A_1) = \operatorname{rg}(A_2)$.

4. $\Rightarrow$ 1.: Jede elementare Zeilenumformung einer Matrix A ist als Multiplikation TA mit *invertierbarer* Matrix T beschreibbar (Abschnitt 3.2.3). Analog ist jede Spaltenumformung als Multiplikation AS mit invertierbarer Matrix S beschreibbar. Produkte invertierbarer Matrizen sind invertierbar, also

$$A_2 = TAS$$

für geeignete invertierbare Matrizen T und S. □

Satz 3.4.28 (Charakterisierung von Ähnlichkeit)**.** *Die folgenden Aussagen sind äquivalent für $A_1, A_2 \in \mathbb{K}^{n\times n}$:*

1. *A_1 und A_2 sind ähnlich: es existiert invertierbare Matrix $S \in \mathbb{K}^{n\times n}$ mit $A_2 = S^{-1}A_1S$;*

2. *Es gibt Basis B von $\mathbb{K}^n$ so dass $A_2 = M_B^B(f_{A_1})$.*

Beweis. (1) $\Rightarrow$ (2): Wähle für $B \subset \mathbb{K}^n$ die Spalten von S.
(2) $\Rightarrow$ (1): Falls $B = (b_1, \dots, b_n)$, so wähle $S := (b_1, \dots, b_n) \in \mathbb{K}^{n\times n}$. □

3.4.9 Homogene Gleichungssysteme und Untervektorräume

Proposition 3.4.29. *Sei V ein $\mathbb{K}$-Vektorraum, und $U \subseteq V$. Dann sind äquivalent.*

1. *U ist ein Untervektorraum von V;*

2. *U ist die Lösungsmenge eines homogenen linearen Gleichungssystems.*

Beispiel 3.4.30. Es sei $V = \mathbb{R}^4$. Betrachten

$$U = \langle \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \rangle.$$

Dann kann U beschrieben werden als

$$\begin{aligned} U &= \{\lambda_1 \begin{pmatrix}1\\1\\0\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \mid \lambda_1, \lambda_2 \in \mathbb{K}\} \\ &= \{\begin{pmatrix}v_1\\v_2\\v_3\\v_4\end{pmatrix} \mid v_1 = v_2, v_4 = 0\} \\ &= \text{Lös}(\begin{pmatrix}1 & -1 & 0 & 0\\0 & 0 & 0 & 1\end{pmatrix}, \mathbf{0}) \,. \end{aligned}$$

△

Beweis von Proposition 3.4.29. Wir nehmen zunächst an, dass $U = \text{Lös}(A, \mathbf{0})$ für $A \in \mathbb{K}^{m\times n}$. Seien $u, v \in U$. Dann gilt $Au = Av = \mathbf{0}$, und damit gilt $A(u + v) = Au + Av = \mathbf{0}$. Analog: nachrechnen, dass $\alpha u \in U$ für alle $\alpha \in \mathbb{K}$ und $u \in U$.

Umgekehrt sei $U \leq V$ und $(u_1, \dots, u_m)$ eine Basis von U. Nach dem Satz von Steinitz findet sich eine Basis B von V der Gestalt $(u_1, \dots, u_m, u_{m+1}, \dots, u_n)$. Definiere $T := M_B^{E_n}(\text{id})$, eine invertierbare Matrix, und seien $X \in \mathbb{K}^{(m,n)}$ und $R \in \mathbb{K}^{(n-m,n)}$ so dass

$T = \begin{pmatrix} X \\ R \end{pmatrix}$. Dann gilt

$$\begin{aligned} v \in U &\Leftrightarrow v \in \langle u_1, \dots, u_m \rangle \\ &\Leftrightarrow Tv \in \langle Tu_1, \dots, Tu_m \rangle = \langle e_1, \dots, e_m \rangle \\ &\Leftrightarrow (Tv)_{m+1} = \cdots = (Tv)_n = 0 \\ &\Leftrightarrow Rv = \mathbf{0} \\ &\Leftrightarrow v \in \text{Lös}(R, \mathbf{0}). \end{aligned}$$

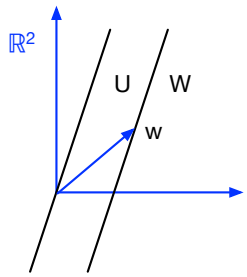
□

3.4.10 Gleichungssysteme und affine Unterräume

Idee: Lösungsmengen von allgemeinen linearen Gleichungssystemen: "affiner Unterraum". Offiziell definieren wir affine Unterräume mit Hilfe von Untervektorräumen.

Definition 3.4.31. Sei V ein $\mathbb{K}$-Vektorraum. Dann ist $W \subseteq V$ ein *affiner Unterraum* von V falls es einen Untervektorraum $U \leq V$ und ein $w \in W$ gibt, so dass $W = \{w + u \mid u \in U\}$.

Siehe Abbildung rechts. In anderen Worten: die affinen Unterräume von V sind genau die *Nebenklassen* von Untervektorräumen von V.



Die *Dimension* eines affinen Raumes $W = \{w + u \mid u \in U\}$ ist definiert als die Dimension des Untervektorraumes U.

Proposition 3.4.32. *Sei V ein $\mathbb{K}$-Vektorraum, und $W \subseteq V$. Dann sind äquivalent:*

(1) *W ist affiner Unterraum von V, oder $W = \varnothing$;*

(2) *W ist die Lösungsmenge eines linearen Gleichungssystems über V;*

(3) *für alle $w_1, \dots, w_l \in W$ and $\alpha_1, \dots, \alpha_l \in K$ mit $\alpha_1 + \cdots + \alpha_l = 1$ ist*

$$\alpha_1 w_1 + \cdots + \alpha_l w_l \in W\,.$$

Abbildung 3.1 veranschaulicht den dritten Punkt in Proposition 3.4.32 anhand von $w, v \in W \subseteq V = \mathbb{R}^2$ und $\alpha_1 w + \alpha_2 v \in W$ für $\alpha_1 = \alpha_2 = \frac{1}{2}$.

Beweis. (2) $\Rightarrow$ (1): Sei $Ax = b$ ein LGS. Falls $Ax = b$ eine Lösung besitzt, so gilt nach Abschnitt 3.3.2

$$\text{Lös}(A, b) = v_0 + \text{Lös}(A, \mathbf{0}).$$

Da $\text{Lös}(A, \mathbf{0}) \leq V$, ist die Lösungsmenge eines LGS also ein affiner Unterraum von V.

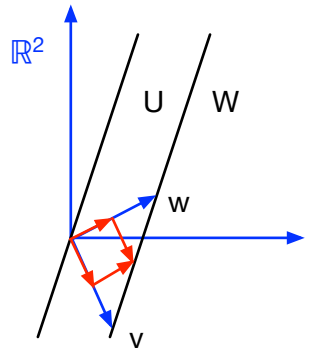


Abbildung 3.1: Illustration einer Affinkombination von w und v in $\mathbb{R}^2$.

(1) $\Rightarrow$ (2): Falls $W = \varnothing$, so ist W Lösungsmenge des LGS $0 = 1$. Ansonsten ist $W = \{w + u \mid u \in U\}$ für $U \leq V$. Nach Proposition 3.4.29 gibt es $A \in \mathbb{K}^{m\times n}$ mit $U = \text{Lös}(A, \mathbf{0})$. Dann ist

$$\begin{aligned} W &= \{w + u \mid u \in \text{Lös}(A, \mathbf{0})\} \\ &= \{w + u \mid Au = \mathbf{0}\} \\ &= \{w' \mid A(w' - w) = \mathbf{0}\} = \text{Lös}(A, Aw). \end{aligned}$$

(1) $\Rightarrow$ (3): Falls $W = \varnothing$, so gilt (3) trivialerweise. Ansonsten ist $W = \{w + u \mid u \in U\}$ für ein $U \leq V$. Es seien $w_1, \dots, w_l \in W$ und $\alpha, \dots, \alpha_l \in K$ so dass $\alpha_1 + \cdots + \alpha_l = 1$. Wir schreiben u_i für $w_i - w_1$. Dann gilt

$$\sum_i \alpha_i w_i = \sum_i \alpha_i (w_1 + u_i) = \underbrace{\sum_i \alpha_i}_{=1} w_1 + \sum_i \alpha_i u_i \in W.$$

(3) $\Rightarrow$ (1): Falls $W = \varnothing$, so ist nichts zu zeigen. Falls $W \neq \varnothing$, dann zeigen wir zunächst, dass $U := \{v - v' \mid v, v' \in W\} \leq V$. Siehe Abbildung 3.2.

Seien $u_1, u_2 \in U$. Dann ist $u_1 = v_1 - v_1'$ und $u_2 = v_2 - v_2'$ für $v_1, v_1', v_2, v_2' \in W$. Also ist nach Annahme $w := v_1 - v_1' + v_2 \in W$, und damit gilt $u_1 + u_2 = w - v_2' \in U$. Sei $u \in U$ und $\alpha \in \mathbb{K}$. Dann gibt es $v, v' \in W$ mit $v - v' = u$. Es gilt

$$\alpha u = \alpha v - \alpha v' = \underbrace{\alpha v - \alpha v' + v'}_{\in W} - v' \in U.$$

Wir zeigen nun, dass $W = w + U = \{w + u \mid u \in U\}$ für ein beliebiges $w \in W$. Sei $u \in U$, also $u = v - v'$ für $v, v' \in W$. Nach Annahme ist auch $w + u = w + v - v' \in W$, da $1 + 1 - 1 = 1$. Umgekehrt lässt sich jedes $w' \in W$ schreiben als $w' = w + w' - w$ und ist damit in $w + U$. $\square$

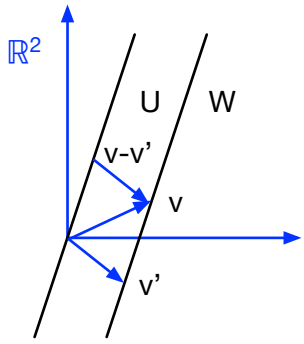


Abbildung 3.2: Illustration zum Beweis der Implikation (3) ⇒ (1).

Definition 3.4.33. Sei V ein $\mathbb{K}$-Vektorraum und $M \subseteq V$. Die *affine Hülle* $\langle M \rangle_{\text{Aff}}$ einer Teilmenge $M \subseteq V$ ist der kleinste affine Unterraum von V, der M enthält.

Ist ein Hüllenoperator (vgl. Abschnitt 2.4.1).

Proposition 3.4.34. *Es gilt*

$$\langle M \rangle_{\text{Aff}} = \{\alpha_1 w_1 + \cdots + \alpha_n w_n \mid w_1, \ldots, w_n \in M, \alpha_1, \ldots, \alpha_n \in K, \sum_i \alpha_i = 1\}$$

Beweis. Sicherlich ist

$$M \subseteq M' := \{\alpha_1 w_1 + \cdots + \alpha_n w_n \mid w_1, \ldots, w_n \in M, \sum_i \alpha_i = 1\}.$$

Behauptung: M' ist affiner Unterraum von V. Seien $\alpha_1, \ldots, \alpha_l \in K$ mit $\alpha_1, \ldots, \alpha_l = 1$ und $w_1, \ldots, w_l \in M'$ mit $w_i = \sum_{j=1}^{l_i} \alpha_{i,j} w_{i,j}$ für $w_{i,j} \in M$ und $\sum_{j=1}^{l_i} \alpha_{i,j} = 1$ für alle $i \leq l$. Dann gilt

$$\alpha_1 w_1 + \cdots + \alpha_l w_l = \sum_{i=1}^{l} \alpha_i \sum_{j=1}^{l_i} \alpha_{i,j} w_{i,j} \in W'$$

da $\sum_i \alpha_i \sum_j \alpha_{i,j} = \sum_i \alpha_i = 1$. Also ist M' nach Proposition 3.4.32 ((3) ⇒ (1)) ein affiner Unterraum von V. Da $\langle M \rangle_{\text{Aff}}$ der kleinste affine Unterraum von V ist, der M enthält, folgt, dass $\langle M \rangle_{\text{Aff}} \subseteq M'$.

Umgekehrt sei $w \in M'$. Dann gibt es $w_1, \ldots, w_n \in M$ und $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ mit $\sum_i \alpha_i = 1$ so dass $w = \alpha_1 w_1 + \cdots + \alpha_n w_n$. Wegen Proposition 3.4.32 ((1) ⇒ (3)) angewandt auf $\langle M \rangle_{\text{Aff}}$ ist $w \in \langle M \rangle_{\text{Aff}}$. □

Kapitel 4

Determinanten, Polynome, Diagonalisierbarkeit

4.1 Determinanten

Determinanten spielen eine Rolle bei

- Lösbarkeit von linearen Gleichungssystemen,
- der Frage, wie lineare Abbildungen das *Volumen* von Körpern verändern,
- und vielem mehr.

4.1.1 Permutationen

Eine bijektive Abbildung $f\colon A \to A$ heißt auch *Permutation* von A.
Lateinisch 'permutere': vertauschen.
Oft ist $A = \{1, 2, \dots, n\}$. Definiere

$$S_n := \{\sigma \mid \sigma \text{ eine Permutation auf } \{1, 2, \dots, n\}\}$$

Es gilt

$$|S_n| = n! := n \cdot (n-1) \cdots 2 \cdot 1$$

Schreibweise für Permutationen:

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

(Zwei-Zeilen-Schreibweise)

Alternativ: Zyklenschreibweise (als Produkt disjunkter Zyklen):

$$\sigma = (a_1 a_2 \dots a_r)(b_1 b_2 \dots b_s)(\dots) \cdots (\dots)$$

falls σ die Elemente wie folgt abbildet (Bild malen!):

$$\sigma(a_i) = a_{i+1(\bmod r)}, \sigma(b_i) = b_{i+1(\bmod s)}, \dots$$

Zyklen der Länge 1, das heißt, "Fixpunkte" c mit $\sigma(c) = c$, werden häufig nicht mitgeschrieben falls Grundmenge aus Kontext klar.

Beispiel.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 8 & 4 & 7 & 6 & 5 & 1 \end{pmatrix}$$

Zyklenschreibweise:

$$(1238)(57)$$

Schreibweise nicht eindeutig:

$$(57) = (75)$$
$$(1238)(57) = (57)(1238)$$

Bemerkungen.

- S_n ist bezüglich der Hintereinanderausführung $\circ$ (Kompositionsoperation) von Abbildungen eine Gruppe, die *(volle) symmetrische Gruppe* auf $A = \{1, 2, \dots, n\}$.
- Eins-element: $\mathrm{id}_A = (1)(2)(3)\cdots(n)$.
- Inverses Element zu σ: die Umkehrabbildung von σ.
  Beispiel: $(a_1 \dots a_s)^{-1} = (a_s \dots a_1)$.
- Die Gruppe S_n ist *nicht* abelsch: Beispiel für $n = 4$:

$$(123) \circ (124) = (13)(24)$$
$$(124) \circ (123) = (14)(23)$$

Permutationen der Form $\tau = (ij)$ (zwei Elemente vertauschen) heißen *Transpositionen.*

Satz 4.1.1. *Jede Permutation lässt sich als Komposition von Transpositionen darstellen.*

Beweis. Jeder Zyklus $(a_1 a_2 \dots a_r)$ ist darstellbar als

$$(a_1 a_2) \circ (a_2 a_3) \circ \cdots \circ (a_{r-1} a_r)$$ □

Sei $\sigma \in S_n$ und sind

$$\sigma = \tau_1 \tau_2 \cdots \tau_k$$
$$\sigma = \tau_1' \tau_2' \cdots \tau_{k'}'$$

zwei Darstellungen als Produkt von Transpositionen, so gilt

$$k = k' \mod 2$$

Definieren das *Signum* von σ:

$$\operatorname{sign}(\sigma) := (-1)^k$$

Also: $\operatorname{sign}(\sigma) = 1$ falls k gerade (σ ist *gerade Permutation*) und $\operatorname{sign}(\sigma) = -1$ falls k ungerade (σ ist *ungerade Permutation*).

Bemerkungen.

- $\operatorname{sign}(\sigma_1\sigma_2) = \operatorname{sign}(\sigma_1)\operatorname{sign}(\sigma_2)$
- $\operatorname{sign}(\tau) = -1$ für alle Transpositionen τ.
- Die geraden Permutationen ($\operatorname{sign}(\sigma)$ positiv) bilden eine Untergruppe von S_n, die sogenannte *alternierende Gruppe*, geschrieben A_n.
- Ist ρ eine ungerade Permutation ($\operatorname{sign}(\rho)$ negativ) so gilt

$$S_n = A_n \cup \rho A_n \qquad \text{wobei } \rho A_n := \{\rho \circ \sigma \mid \sigma \in A_n\}.$$

4.1.2 Determinantenfunktionen

Es gibt (mindestens) zwei grundverschiedene Möglichkeiten, Determinanten einzuführen: Mit einer expliziten Formel, oder über ihre Eigenschaften. Wir wählen letzteren Zugang.

Definition 4.1.2. Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}_+$. Eine Funktion

$$\det\colon \mathbb{K}^{n\times n} \to \mathbb{K} : A \mapsto \det A$$

heißt *Determinantenfunktion* wenn sie folgende Eigenschaften hat:

(D1) *Linearität in jeder Zeile*, d.h., für alle Zeilenvektoren $z_1, \dots, z_n, z_i' \in \mathbb{K}^n$ mit $i \in \{1, \dots, n\}$ und $\lambda \in \mathbb{K}$ gilt

$$\det\begin{pmatrix} z_1 \\ \vdots \\ z_i + z_i' \\ \vdots \\ z_n \end{pmatrix} = \det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} + \det\begin{pmatrix} z_1 \\ \vdots \\ z_i' \\ \vdots \\ z_n \end{pmatrix}$$

$$\det\begin{pmatrix} z_1 \\ \vdots \\ \lambda \cdot z_i \\ \vdots \\ z_n \end{pmatrix} = \lambda \cdot \det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix};$$

(D2) det ist *alternierend*, d.h., hat A zwei gleiche Zeilen, so ist $\det A = 0$;

(D3) $\det E_n = 1$.

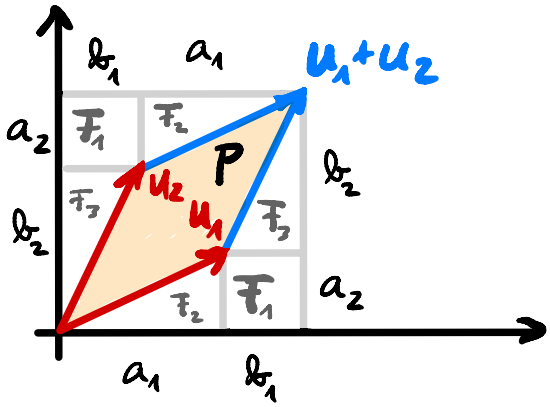


Abbildung 4.1: Die Determinante misst den vorzeichenbehafteten Flächeninhalt eines Parallelograms.

Gibt es solche Funktionen?

Beispiel 4.1.3. $n = 2$.

$$\det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} := a_1b_2 - a_2b_1 \tag{4.1}$$

ist Determinantenfunktion (nachprüfen!). Geometrische Interpretation für $\mathbb{K} = \mathbb{R}$: Ausdruck in (4.1) misst den 'vorzeichenbehafteten' Flächeninhalt des Parallelogramms P, das von den folgenden beiden Vektoren aufgespannt wird.

$$u_1 = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{und} \quad u_2 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Siehe Abbildung 4.1: falls $0 < \measuredangle(u_1, u_2) < 180^\circ$ so ist Flächeninhalt von P gleich

$$\begin{aligned} &(a_1 + b_1)(a_2 + b_2) - 2F_1 - 2F_2 - 2F_3 \\ &= a_1a_2 + a_1b_2 + b_1a_2 + b_1b_2 - 2b_1a_2 - a_1a_2 - b_1b_2 \\ &= a_1b_2 - a_2b_1. \end{aligned}$$

Vorzeichen ist negativ falls $180^\circ < \alpha = \measuredangle(u_1, u_2) < 360^\circ$. △

Bemerkung 4.1.4. Für $n = 3$ sind Determinanten als Volumina interpretierbar.

4.1.3 Eigeschaften von Determinantenfunktionen

Die folgende Proposition klärt das Verhalten einer Determinantenfunktion $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ bei elementaren Zeilenumformungen.

Proposition 4.1.5. *Sei* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ *eine Determinantenfunktion.*

1. $z_i \leftrightsquigarrow z_j$ *für* $i \neq j$*: Entsteht* A' *aus* $A \in \mathbb{K}^{n\times n}$ *durch Vertauschen von zwei Zeilen, so gilt*
$$\det A' = -\det A$$
2. $\lambda z_i \rightsquigarrow z_i$*: Entsteht* A' *aus* A *durch Multiplikation einer Zeile mit einem Skalar* $\lambda \in \mathbb{K}$*, so gilt:*
$$\det A' = \lambda \cdot \det A$$
3. $\lambda z_i + z_j \rightsquigarrow z_j$ *für* $i \neq j$*: Entsteht* A' *aus* A *durch Addition eines Vielfachen einer Zeile zu einer anderen Zeile, so gilt:*
$$\det A' = \det A$$

Das gleiche gilt auch für elementare Spaltenumformungen: folgt später.

Beweis. 2.: folgt direkt aus der Linearität (D1).
3.:
$$\det A' = \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j + \lambda z_i \\ \vdots \\ z_n \end{pmatrix} \overset{(D1)}{=} \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ z_n \end{pmatrix} + \lambda \cdot \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} \overset{(D2)}{=} \det A + 0$$
1.: Seien
$$A = \begin{pmatrix} \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \end{pmatrix} \text{ und } A' = \begin{pmatrix} \vdots \\ z_j \\ \vdots \\ z_i \\ \vdots \end{pmatrix}$$
Betrachten
$$B := \begin{pmatrix} \vdots \\ z_i + z_j \\ \vdots \\ z_j \\ \vdots \end{pmatrix}, \quad B' := \begin{pmatrix} \vdots \\ z_i + z_j \\ \vdots \\ z_i \\ \vdots \end{pmatrix}, \text{ und } B'' := \begin{pmatrix} \vdots \\ z_i + z_j \\ \vdots \\ z_i + z_j \\ \vdots \end{pmatrix}.$$
Dann ist
$$0 \overset{(D2)}{=} \det B'' \overset{(D1)}{=} \det B + \det B' \qquad (*)$$
Also
$$\det A \overset{(3)}{=} \det B \overset{(*)}{=} -\det B' \overset{(3)}{=} -\det A'. \qquad \square$$

(D2) lässt sich verschärfen:

Satz 4.1.6. *Für eine Determinantenfunktion* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ *gilt:*

$$\text{(D2')} \quad \operatorname{rg}(A) < n \iff \det A = 0$$

Beweis. "$\Rightarrow$": Wenn $\operatorname{rg}(A) < n$ dann ist eine Zeile Linearkombination der anderen. O.B.d.A: $z_n = \sum_{i=1}^{n-1} \lambda_i z_i$. Wegen der Linearitätsbedingung (D1) gilt:

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_{n-1} \\ z_n \end{pmatrix} \overset{\text{(D1)}}{=} \sum_{i=1}^{n-1} \lambda_i \det \begin{pmatrix} z_1 \\ \vdots \\ z_{n-1} \\ z_i \end{pmatrix} \overset{\text{(D2)}}{=} 0$$

"$\Leftarrow$": Wenn $\operatorname{rg}(A) = n$ dann kann A durch elementare Zeilenumformungen in E_n umgeformt werden. Wäre $\det A = 0$, so wäre auch $\det E_n = 0$ gemäß Proposition 4.1.5, im Widerspruch zu (D3). Also $\det A \neq 0$. □

Rückführung auf E_n stets möglich bei $\operatorname{rg}(A) = n$, also ist Determinante eindeutig berechenbar. Wenn es sie überhaupt gibt!

Lemma 4.1.7. *Es gibt höchstens eine Determinantenfunktion* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$.

Beweis. Seien $\det, \det'$ Determinantenfunktionen. Wegen Satz 4.1.6 gilt $\det A = \det' A = 0$ falls $\operatorname{rg} A < n$. Sei also $\operatorname{rg} A = n$. Nach Satz 3.2.27 erhalten wir E_n aus A durch elementare Zeilenumformungen. Dabei ändert sich die Determinante gemäß Proposition 4.1.5 für det und für $\det'$ auf die gleiche Weise. Weil $\det E_n = \det' E_n \overset{\text{(D3)}}{=} 1$, folgt $\det A = \det' A$. □

4.1.4 Die Leibnizsche Formel

Satz 4.1.8 (Die Leibnizsche Formel)**.** *Es gibt (genau) eine Determinantenfunktion*

$$\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}.$$

Für $A = (a_{ij}) \in \mathbb{K}^{n\times n}$ *gilt*

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \tag{4.2}$$

Verallgemeinerung vom Fall $n = 2$ aus Abschnitt 4.1.2.

Beweis. Wegen Lemma 4.1.7 muss nur gezeigt werden, dass die in (4.2) definierte Funktion det die Eigenschaften (D1), (D2) und (D3) hat.

(D1) Seien

$$B := \begin{pmatrix} z_1 \\ \vdots \\ z_i + \lambda z_i' \\ \vdots \\ z_n \end{pmatrix}, \quad A := \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} \text{ und } A' := \begin{pmatrix} z_1 \\ \vdots \\ z_i' \\ \vdots \\ z_n \end{pmatrix}$$

Dann gilt

$$\begin{aligned} \det B &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) b_{1\sigma(1)} \cdots b_{i,\sigma(i)} \cdots b_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots (a_{i,\sigma(i)} + a'_{i,\sigma(i)}) \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i,\sigma(i)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a'_{i,\sigma(i)} \cdots a_{n\sigma(n)} \\ &= \det A + \det A' \end{aligned}$$

Analog:

$$\det \begin{pmatrix} \vdots \\ \lambda z_i \\ \vdots \end{pmatrix} = \lambda \det \begin{pmatrix} \vdots \\ z_i \\ \vdots \end{pmatrix}$$

(D2) Die Matrix $A = (a_{ij})$ habe zwei gleiche Zeilen, o.B.d.A. $z_1 = z_2$. Das heißt:

$$a_{1j} = a_{2j} \qquad \text{für } j \in \{1, \ldots, n\}$$

Sei $\tau = (12) \in S_n$ (Transposition). Dann

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in A_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma' \in A_n \tau} \operatorname{sign}(\sigma') a_{1\sigma'(1)} \cdots a_{n\sigma'(n)} \\ &= \sum_{\sigma \in A_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in A_n} -a_{1\sigma(2)} a_{2\sigma(1)} a_{3\sigma(3)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in A_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} - \sum_{\sigma \in A_n} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= 0 \end{aligned}$$

Die zweite Gleichung gilt, da $S_n = A_n \cup A_n\tau$, die dritte, da $\sigma\tau(1) = \sigma(2)$, $\sigma\tau(2) = \sigma(1)$, und $\sigma\tau(n) = \sigma(n)$ für alle $n \in \{3, \ldots, n\}$, und die vierte, da $a_{1j} = a_{2j}$ für alle $j \in \{1, \ldots, n\}$, und damit $a_{1\sigma(2)} a_{2\sigma(1)} = a_{1\sigma(1)} a_{2\sigma(2)}$.

(D3)

$$\det E_n = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)} = \operatorname{sign}(\mathrm{id}) \cdot 1 \cdots 1 = 1$$

wobei $a_{ij} = \delta_{ij} = 1$ für $i = j$ und $a_{ij} = \delta_{ij} = 0$ sonst. □

Neue Schreibweise: $|A| := \det A$,

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Übung 15. Beweisen Sie, dass für quadratische Matrizen A und B gilt

$$\begin{vmatrix} A & 0 \\ * & B \end{vmatrix} = \det(A) \cdot \det(B).$$

Beispiel 4.1.9. Betrachten $n = 3$ und $A \in \mathbb{K}^{3\times 3}$. Es gibt $3! = 6$ Permutationen $\sigma \in S_n$. Die Leibnizformel ergibt:

$$\left|\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right.$$

Schräglinien einzeichnen! Vorzeichen!

Gerade Permutationen (Diagonalen):

$+a_{11}a_{22}a_{33}$	$+a_{12}a_{23}a_{31}$	$+a_{13}a_{21}a_{32}$
$\sigma = \mathrm{id}$	$\sigma = (123)$	$\sigma = (132)$

Ungerade Permutationen (Nebendiagonalen):

$-a_{31}a_{22}a_{13}$	$-a_{32}a_{23}a_{11}$	$-a_{33}a_{21}a_{12}$
$\sigma = (13)$	$\sigma = (23)$	$\sigma = (12)$

Regel von Sarrus (gesprochen Sarrüs). △

4.1.5 Berechnung der Determinante

Wie berechnet man Determinanten? Die Leibnizsche Regel ist im allgemeinen zu aufwändig. Idee: die Determinante einer *Dreiecksmatrix* ist das Produkt der Hauptdiagonalelemente.

$$\det \begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = a_{11}a_{22}\cdots a_{nn} \tag{4.3}$$

Folgt sofort aus Leibnizformel. Für $\sigma \neq \mathrm{id}$ ist ein Faktor in $a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}$ gleich Null (denn $a_{ij} = 0$ für $i > j$).

Diese Idee liefert ein Berechnungsverfahren!

- Umwandlung der Matrix in Stufenform. Dabei ändert sich Determinante gemäß Proposition 4.1.5 in Abschnitt 4.1.3.

- Determinante der Stufenform ist Produkt der Hauptdiagonalelemente (wie oben erklärt).

Beispiel 4.1.10.

$$\begin{aligned}
\begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 2 \\ 0 & 1 & 0 \end{vmatrix} &= \begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 1 & 0 \end{vmatrix} && (z_2 - 2z_1 \rightsquigarrow z_2) \\
&= \begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{vmatrix} && (z_3 - z_2 \rightsquigarrow z_3) \\
&= 1 \cdot 1 \cdot 2 = 2
\end{aligned}$$

Alternative Rechnung:

$$\begin{aligned}
\begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 2 \\ 0 & 1 & 0 \end{vmatrix} &= \begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 1 & 0 \end{vmatrix} && (z_2 - 2z_1 \rightsquigarrow z_2) \\
&= - \begin{vmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{vmatrix} && (s_2 \leftrightsquigarrow s_3) \\
&= -(1 \cdot -2 \cdot 1) = 2
\end{aligned}$$

△

Übung 16. Zeigen Sie, dass das skizzierte Verfahren zur Berechnung der Determinante einer Matrix aus $\mathbb{K}^{n \times n}$ über die Stufenform mit einer Anzahl an Rechenoperationen in $\mathbb{K}$ auskommt, die beschränkt ist durch ein Polynom in n.

Es gibt viele Möglichkeiten, das Berechnen von Determinanten zu erleichtern.

Proposition 4.1.11. *Sei $A \in \mathbb{K}^{n \times n}$. Dann gilt:*

$$\det A^\top = \det A$$

Beweis. $A = (a_{ij})$, $A^\top = (b_{ij})$, $b_{ij} = a_{ji}$.

$$\begin{aligned}
\det A^\top &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) b_{1\sigma(1)} \dots b_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{k_1\sigma^{-1}(k_1)} \dots a_{k_n\sigma^{-1}(k_n)} && \text{wobei } k_i := \sigma(i) \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma^{-1}(1)} \dots a_{n\sigma^{-1}(n)} && \text{Umordnen: } \{1, \dots, n\} = \{k_1, \dots, k_n\} \\
&= \sum_{\sigma' \in S_n} \operatorname{sign}(\sigma') a_{1\sigma'(1)} \dots a_{n\sigma'(n)} && \text{denn } \operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1}) \\
&= \det A
\end{aligned}$$

□

Folgerung: Das Verhalten der Determinante bei elementaren Spaltenumformungen ist das gleiche wie bei elementaren Zeilenumformungen (ersetze 'Zeile' durch 'Spalte' in Proposition 4.1.5).

Satz 4.1.12. *Seien $A, B \in \mathbb{K}^{n\times n}$. Dann gilt:*

1. $\det(AB) = \det A \cdot \det B$
2. *Für $A \in \mathrm{GL}(n, \mathbb{K})$ (invertierbare Matrizen, siehe Abschnitt 3.2.1) gilt*

$$\det(A^{-1}) = (\det A)^{-1}.$$

Beweis. 1.: Falls $\operatorname{rg}(B) < n$ dann $\operatorname{rg}(AB) < n$ (siehe Korollar 3.4.16), und mit Satz 4.1.6

$$\det(AB) = 0 = \det A \cdot \det B.$$

Also muss (1) nur für invertierbares B bewiesen werden ($\operatorname{rg}(B) = n$, $\det B \neq 0$). Sei $B \in \mathrm{GL}(n, \mathbb{K})$ fest. Die Abbildung

$$f\colon \mathbb{K}^{n\times n} \to \mathbb{K} : A \mapsto \det(AB)$$

hat folgende Eigenschaften:

(D1) f ist linear in den Zeilen von A

(D2) Falls A zwei gleiche Zeilen hat, dann ist $f(A) = 0$:

$$\begin{aligned}
\operatorname{rg}(A) < 0 &\Rightarrow \operatorname{rg}(AB) < n && \text{(Korollar 3.4.16)} \\
&\Rightarrow \det AB = 0 && \text{wie oben für } B \text{ statt } A.
\end{aligned}$$

Weiterhin gilt

$$f(E_n) = \det(E_n B) = \det B\,.$$

Damit erfüllt die Abbildung

$$\tilde{f}\colon \mathbb{K}^{n\times n} : A \mapsto \underbrace{(\det B)}_{\neq 0}{}^{-1} f(A)$$

alle drei Bedingungen (D1), (D2) und (D3) einer Determinantenfunktion. Wegen Eindeutigkeit der Determinantenfunktion (Lemma 4.1.7) folgt $\tilde{f}(A) = \det A$, also

$$\begin{aligned} & (\det B)^{-1} \det(AB) = \det A \\ \Rightarrow \quad & \det(AB) = \det A \cdot \det B\,. \end{aligned}$$

2.:

$$\begin{aligned} 1 = \det E_n &= \det(AA^{-1}) \\ &= \det A \cdot \det(A^{-1}) && \text{nach Teil 1.} \end{aligned}$$

Also $\det(A^{-1}) = (\det A)^{-1}$. □

Direktes Nachrechnen mit Leibnizformel ebenfalls möglich.

Weitere Rechenregeln. Wir definieren nun die *Entwicklung* einer Determinante nach einer Zeile oder Spalte. Eine Matrix, die aus $A \in \mathbb{K}^{m\times n}$ durch wiederholtes Streichen von beliebig vielen Zeilen und Spalten entsteht, heißt *Untermatrix* (oder *Teilmatrix*) von A.

Definition 4.1.13. Sei $A \in \mathbb{K}^{n\times n}$. Dann steht $A_{ij} \in \mathbb{K}^{(n-1)\times(n-1)}$ für die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

$$A = \begin{pmatrix} & & a_{1j} & & \\ & B_1 & \vdots & B_2 & \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{i,n} \\ & B_3 & \vdots & B_4 & \\ & & a_{nj} & & \end{pmatrix}, \quad A_{ij} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$$

Satz 4.1.14 (Entwicklungssatz)**.** *Es gilt*

$(*)_i$ *Entwicklung nach der i-ten Zeile:*

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

$(**)_j$ *Entwicklung nach der j-ten Spalte:*

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

Zum Beweis. $(*)_i$:

$$\det\begin{pmatrix} & & & 0 & & & \\ & B_1 & & \vdots & & B_2 & \\ & & & 0 & & & \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ & & & 0 & & & \\ & B_3 & & \vdots & & B_4 & \\ & & & 0 & & & \end{pmatrix} \overset{(D1)}{=} (-1)^{i+j} \det\begin{pmatrix} a_{ij} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & B_1 & B_2 \\ \mathbf{0} & B_3 & B_4 \end{pmatrix}$$

$$= (-1)^{i+j} a_{ij} \det A_{ij}$$

folgt aus der Leibnizformel.
$(**)_j$: folgt aus $(*)_j$ und Proposition 4.1.11: $\det A^\top = \det A$. □

Beispiel 4.1.15. $n = 3$:

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & 4 & 0 \\ 3 & 0 & 0 \end{vmatrix} = \quad ? \qquad \text{Vorzeichen } (-1)^{i+j}: \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

Entwicklung nach 1. Zeile $(*)_1$:

$$\begin{aligned} &+ 1 \cdot \begin{vmatrix} 4 & 0 \\ 0 & 0 \end{vmatrix} - 2 \cdot \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 4 \\ 3 & 0 \end{vmatrix} \\ &= 1 \cdot 0 - 2 \cdot 0 + 1 \cdot (1 \cdot 0 - 3 \cdot 4) = -12 \end{aligned}$$

Entwicklung nach 2. Spalte $(**)_2$:

$$\begin{aligned} &- 2 \cdot \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} + 4 \cdot \begin{vmatrix} 1 & 3 \\ 1 & 0 \end{vmatrix} - 0 \cdot |\cdots| \\ &= -2 \cdot 0 + 4 \cdot (1 \cdot 0 - 1 \cdot 3) - 0 = -12 \end{aligned}$$

Entwicklung nach 3. Zeile $(*)_3$:

$$\begin{aligned} &+ 3 \cdot \begin{vmatrix} 2 & 1 \\ 4 & 0 \end{vmatrix} - 0 \cdot |\cdots| + 0 \cdot |\cdots| \\ &= 3 \cdot (2 \cdot 0 - 4 \cdot 1) = -12 \end{aligned}$$

△

Übung 17. Mit dem Entwicklungssatz läßt sich rekursive die Determinante einer Matrix in $\mathbb{K}^{n\times n}$ vollständig berechnen. Ist das entsprechende Verfahren *polynomiell* in dem Sinn, dass die benötigte Anzahl der Rechenoperationen in $\mathbb{K}$ durch ein Polynom in n beschränkt ist?

4.1.6 Die Determinante von linearen Abbildungen

Sei $f\colon V \to V$ eine lineare Abbildung, und $A := M_B^B(f)$ die zugehörige Matrix (bezüglich einer Basis B von V; siehe Abschnitt 3.4.5).

Definition 4.1.16. $\det f := \det A$.

Dies ist wohldefiniert, da für $A_1 = M_B^B(f)$, $A_2 = M_{B'}^{B'}(f)$ nach Satz 3.4.28 eine invertierbare Matrix $S \in \mathbb{K}^{n\times n}$ existiert mit $A_2 = S^{-1}A_1S$. Also

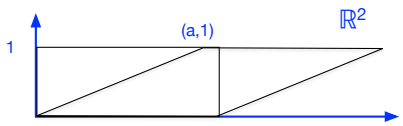
$$\begin{aligned}
\det A_2 &= \det(S^{-1}A_1S) \\
&= (\det S)^{-1} \det A_1 (\det S) && \text{nach Satz 4.1.12} \\
&= \det A_1 && \text{da Multiplikation in } \mathbb{K} \text{ kommutativ.}
\end{aligned}$$

Bemerkung 4.1.17. Wir halten fest: ähnliche Matrizen haben die gleiche Determinante.

Bemerkung 4.1.18. $\det f$ kann als Verzerrungsfaktor für Flächen (in $\mathbb{R}^2$) bzw. Volumina (in $\mathbb{R}^n$) der Abbildung f interpretiert werden (siehe auch späteren Abschnitt **??**):

$$F' := (\det f) \cdot F$$

Beispiel 4.1.19. $V = \mathbb{R}^2$, $A := \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$
beschreibt Scherung in x Richtung mit Faktor a.
$\det(A) = 1$, Flächeninhalt bleibt gleich.



△

4.1.7 Lösung linearer Gleichungssysteme mittels Determinanten

Lineares Gleichungssystem

$$Ax = b$$

mit $A \in \mathbb{K}^{m\times n}$, $b \in \mathbb{K}^m$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Spezialfall $n = m$:

$$\begin{aligned}
& Ax = b \text{ eindeutig lösbar} \\
\Leftrightarrow\ & \operatorname{rg}(A) = n && \text{nach Korollar 3.3.7} \\
\Leftrightarrow\ & \det(A) \neq 0 && \text{nach Satz 4.1.6}
\end{aligned}$$

‘Eisensteinkriterium’ (1844).

Wir wissen also insbesondere, dass $Ax = b$ eine Lösung besitzt, falls $\det(A) \neq 0$. Die Determinante kann aber sogar verwendet werden, um diese explizit auszurechnen! Definieren dazu die Matrix

$$A_j(b) := \begin{pmatrix} a_{11} & \cdots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & & & & \vdots & & \\ a_{n1} & \cdots & a_{n(j-1)} & b_n & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix};$$

$A_j(b)$ entsteht also aus A durch Ersetzung der j-ten Spalte durch b.

Satz 4.1.20 (Cramersche Regel)**.** *Sei $A \in \mathbb{K}^{n\times n}$ mit $\det(A) \neq 0$. Dann berechnet sich die eindeutige Lösung x von $Ax = b$ wie folgt:*

$$x = \frac{1}{\det A}\begin{pmatrix} \det(A_1(b)) \\ \vdots \\ \det(A_n(b)) \end{pmatrix} = \begin{pmatrix} \det(A_1(b))/\det(A) \\ \vdots \\ \det(A_n(b))/\det(A) \end{pmatrix}. \tag{4.4}$$

Beispiel 4.1.21. Das lineare Gleichungssystem

$$\begin{aligned} 2x_1 + 3x_2 &= 1 \\ x_1 - 4x_2 &= 6 \end{aligned}$$

hat die eindeutige Lösung

$$x_1 = \frac{\begin{vmatrix} 1 & 3 \\ 6 & -4 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -4 \end{vmatrix}} = \frac{-22}{-11} = 2$$

$$x_2 = \frac{\begin{vmatrix} 2 & 1 \\ 1 & 6 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -4 \end{vmatrix}} = \frac{11}{-11} = -1 \qquad \triangle$$

Beweis der Cramerschen Regel (4.4).

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} a_{n1} \\ \vdots \\ a_{nn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

b nach links in die i-te Spalte bringen:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} x_i a_{1i} - b_1 \\ \vdots \\ x_i a_{ni} - b_n \end{pmatrix} + \cdots + \begin{pmatrix} a_{n1} \\ \vdots \\ a_{nn} \end{pmatrix} x_n = \mathbf{0}$$

Die Spalten sind also linear abhängig, und nach Satz 4.1.6 ist die Determinante gleich Null:

$$\begin{vmatrix} a_{11} & \cdots & x_i a_{1i} - b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & x_i a_{ni} - b_n & \cdots & a_{nn} \end{vmatrix} = 0$$

Wegen der Linearitätsbedingung (D1):

$$x_i \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix} - \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} = 0$$

Also $x_i = \frac{|A_i(b)|}{|A|}$. □

Folgerung. Wenn $A \in \mathbb{Q}^{n\times n}$ und $b \in \mathbb{Q}^n$, und $x \in \mathbb{R}^n$ die eindeutige Lösung ist von $Ax = b$, dann gilt $x \in \mathbb{Q}^n$.

Wie schnell sind die Algorithmen zum Lösen linearer Gleichungssysteme? Man sieht leicht, dass die Überführung einer Matrix aus $\mathbb{K}^{m\times n}$ in Stufenform aus Abschnitt 3.2.4 höchstens mn viele elementare Zeilenumformungen erfordert. Jede Zeilenumformung benötigt eine lineare Anzahl an arithmetischen Operationen. Damit scheint der gaußsche Algorithmus insgesamt besser zu sein als die Auswertung von (4.4). Wir müssen allerdings bei der exakten Analyse vorsichtig sein, denn eine einzelne arithmetische Operation kann sehr viel Zeit erfordern, wenn die Zahlen sehr groß werden. Bei ungeschickten Folgen von elementaren Zeilenumformungen zur Umwandlung in Stufenform können tatsächlich extrem große Zahlen auftreten; wir betrachten dazu das folgende Beispiel.

Beispiel 4.1.22. Betrachte die folgende Überführung einer Matrix in Stufenform. Sei $x \in \mathbb{Z}$ eine Zahl, z.B. $x = 2$.

$$\begin{pmatrix} 1 & -x & 0 \\ x & 1 & 0 \\ x & x & x+1 \\ x & 0 & 0 \end{pmatrix} \xrightarrow[i\in\{1,2,3\}]{z_i - xz_1 \rightsquigarrow z_i} \begin{pmatrix} 1 & -x & 0 \\ 0 & x^2+1 & 0 \\ 0 & x^2+x & x+1 \\ 0 & x^2 & 0 \end{pmatrix} \xrightarrow{z_3 - z_4 \rightsquigarrow z_3} \begin{pmatrix} 1 & -x & 0 \\ 0 & x^2+1 & 0 \\ 0 & x & x \\ 0 & x^2 & 0 \end{pmatrix}$$

$$\xrightarrow{z_2 - xz_3 \rightsquigarrow z_2} \begin{pmatrix} 1 & -x & 0 \\ 0 & 1 & -x^2 \\ 0 & x & x \\ 0 & x^2 & 0 \end{pmatrix} \xrightarrow[z_4 - x^2 z_2 \rightsquigarrow z_4]{z_3 - xz_2 \rightsquigarrow z_3} \begin{pmatrix} 1 & -x & 0 \\ 0 & 1 & -x^2 \\ 0 & 0 & x^3+x \\ 0 & 0 & (x^2)^2 \end{pmatrix}$$

Wir bemerken, dass jede der 4 Zeilenumformungen *natürlich* ist in dem Sinn, dass sie, angewandt auf eine Matrix der Gestalt

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

mit A in Stufenform, einen Einträg der erste Spalte von C kleiner macht.

Dieses Beispiel läßt sich verallgemeinern. Und zwar sei

$$A(1,x) := \begin{pmatrix} 1 & -x \\ x & -1 \\ x & x \\ x & 0 \end{pmatrix} \text{ und } A(i+1,x) := \begin{pmatrix} & & & & 0 \\ & & A(i,x) & & \vdots \\ & & & & 0 \\ & & & & x+1 \\ & & & & 1 \\ x & 0 & \cdots & 0 & x \\ x & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Dann gibt es für jedes $n \geq 2$ eine natürliche (im obigen Sinn) Umformung der Matrix $A(n,x) \in \mathbb{Z}^{2n\times n}$ in Stufenform, bei der der Eintrag rechts unten in der Stufenform x^{2^n} ist (*'doppelt exponentiell'* [2]).

Es werden exponentiell viele Bits in n benötigt, um eine Zahl der Größenordnung x^{2^n} abzuspeichern. Damit benötigt ein Verfahren, bei dem solche Zahlen erzeugt werden, auch exponentiell viel Zeit. Man kann sich schnell davon überzeugen, dass es dann schon für relativ kleine n zu astronomisch großen Rechenzeiten kommt. △

Das Verfahren zur Berechnung der eindeutigen Lösung eines Gleichungssystems aus Abschnitt 4.1.7 mit Hilfe von Determinanten hat das Problem der zu großen Zahlen *nicht* (insbesondere können die auftretenden Determinanten nie doppelt exponentiell groß werden; siehe Lemma 4.1.24).

Die Umformung in Stufenform aus Beispiel 4.1.22 ist *nicht* die, die der gaußschen Algorithmus vorgenommen hätte: beim Verfahren aus Abschnitt 3.2.4 wird im k-ten Schritt von jeder Zeile z_l mit $l > k$ der Vektor $(a_{l,j_k} a_{k,j_k}^{-1}) z_k$ subtrahiert (wir verwenden die Bezeichnungen aus Abschnitt 3.2.4). Mit Hilfe von Determinanten läßt sich zeigen, dass die Laufzeit des gaußschen Algorithmus polynomiell in der Eingabegröße ist. Essentiell dafür ist, dass alle auftretenden rationalen Zahlen stets *gekürzt* werden; dazu folgende Definition.

Definition 4.1.23. Sei $r = p/q \in \mathbb{Q}$, $p, q \in \mathbb{Z}$ teilerfremd, $q > 0$. Wir definieren

$$\mathrm{Groe}(r) := 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(q+1) \rceil \in \mathbb{N}.$$

Sei nun $b \in \mathbb{Q}^n$ und $A \in \mathbb{Q}^{m\times n}$. Dann definieren wir

$$\mathrm{Groe}(b) := 1 + \mathrm{Groe}(b_1) + \cdots + \mathrm{Groe}(b_n)$$
$$\mathrm{Groe}(A) := mn + \sum_{i,j} \mathrm{Groe}(a_{ij}).$$

Lemma 4.1.24. *Sei $A \in \mathbb{Q}^{m\times m}$. Dann ist* $\mathrm{Groe}(\det A) \leq 2\,\mathrm{Groe}(A)$.

Beweis. Sei $A = (p_{ij}/q_{ij})_{i,j}$ wobei $p_{ij}, q_{ij} \in \mathbb{Z}$ für alle i, j teilerfremd und $q_{ij} > 0$. Seien ausserdem $p, q \in \mathbb{Z}$ mit $p/q = \det A$ so dass p und q teilerfremd und $q > 0$. Dann gilt

$$q \leq \prod_{i,j=1}^{n} q_{ij} = 2^{\log_2(\prod_{i,j=1}^{n} q_{ij})} = 2^{\sum_{i,j=1}^{n} \log_2 q_{ij}} < 2^{\mathrm{Groe}(A)-1} \tag{4.5}$$

und mit der Leibnizschen Formel

$$|\det A| \leq \prod_{i,j=1}^{n} (|p_{ij}| + 1) \,.$$

Damit haben wir

$$|p| = |\det A| \cdot q \leq \prod_{i,j=1}^{n} (|p_{ij}| + 1) q_{ij} = 2^{\prod_{i,j=1}^{n} \log_2(|p_{ij}|+1) + \log_2 q_{ij}} < 2^{\mathrm{Groe}(A)-1} \,. \tag{4.6}$$

Aus (4.5) und (4.6) folgt

$$\mathrm{Groe}(\det A) = 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(q + 1) \rceil < 2\,\mathrm{Groe}(A). \qquad \square$$

Proposition 4.1.25. *Wenn $Ax = b$, mit $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$, eine Lösung hat, dann auch eine der Größe höchstens $2(\mathrm{Groe}(A) + \mathrm{Groe}(b))$.*

Beweis. Wir können annehmen, dass die Zeilen von A linear unabhängig sind (denn da $Ax = b$ eine Lösung hat, können abhängige Zeilen entfernt werden ohne den Lösungsraum zu verändern). Ausserdem können wir durch umsortieren der Spalten von A annehmen, dass A von der Gestalt $[A_1 \; A_2]$ ist für A_1 mit $\det(A_1) \neq 0$. Dann ist

$$\begin{pmatrix} A_1^{-1} b \\ 0 \end{pmatrix}$$

eine Lösung von $Ax = b$, und die Größe dieser Lösung erfüllt nach Lemma 4.1.24 die gewünschte Schranke. $\square$

Wir wollen nun nachweisen, dass der gaußsche Algorithmus polynomielle Laufzeit hat. Es genügt nicht zu zeigen, dass die berechneten Lösungen polynomielle Größe haben, sondern wir müssen dies auch von allen Zahlen nachweisen, die im Laufe der Berechnung auftreten!

Falls $M \in \mathbb{K}^{m \times n}$, $i_1, \ldots, i_k \in \{1, \ldots, m\}$, und $j_1, \ldots, j_l \in \{1, \ldots, n\}$, dann schreiben wir $M_{j_1,\ldots,j_k}^{i_1,\ldots,i_k}$ für die Untermatrix von M, die aus M durch Löschen aller Zeilen ausser $i_1, \ldots, i_k$ und aller Spalten ausser $j_1, \ldots, j_l$ entsteht. Sei

$$A_k = \begin{pmatrix} B_k & C_k \\ \mathbf{0} & D_k \end{pmatrix}$$

die Matrix im k-ten Schritt des Verfahrens aus Abschnitt 3.2.4. Dann gilt für jeden Eintrag d_{ij} von D_k offensichtlicherweise

$$d_{ij} = \frac{\det \begin{pmatrix} B_k & * \\ \mathbf{0} & d_{ij} \end{pmatrix}}{\det(B_k)} = \frac{\det\left((A_k)_{1,\ldots,k,k+j}^{1,\ldots,k,k+i}\right)}{\det\left((A_k)_{1,\ldots,k}^{1,\ldots,k}\right)}. \tag{4.7}$$

Da A_k aus A durch Addition von Vielfachen der ersten k Zeilen zu anderen Zeilen entstanden ist, gilt (4.7) bis auf das Vorzeichen auch für A anstatt A_k (wir erinnern an Proposition 4.1.3 und Bemerkung 3.2.24), d.h.,

$$d_{ij} = \pm \frac{\det\left(A^{1,\dots,k,k+i}_{1,\dots,k,k+j}\right)}{\det\left(A^{1,\dots,k}_{1,\dots,k}\right)}.$$

Also gilt $\operatorname{Groe}(d_{ij}) \leq 4\operatorname{Groe}(M)$ nach Lemma 4.1.24. Auf ähnliche Weise läßt sich auch für den zweiten Teil des gaußschen Algorithmus nachweisen, dass die auftretenden Zahlen nicht zu groß werden [4].

4.1.8 Invertieren einer Matrix mittels Determinanten

Falls eine Matrix $A \in \mathbb{K}^{n\times n}$ invertierbar ist, so lässt sich die Inverse elegant mit Hilfe von Determinanten berechnen. Dazu verwenden wir wieder die Teilmatrizen A_{ij} (Definition 4.1.13) und die Schreibweise $A_i(b)$ aus Abschnitt 4.1.7. Zunächst eine Hilfsaussage, die direkt aus dem Entwicklungssatz (Satz 4.1.14) folgt.

Lemma 4.1.26. *Für $A \in \mathbb{K}^{n\times n}$ und $i, j \in \{1,\dots,n\}$ gilt* $\det(A_j(e_i)) = (-1)^{i+j}\det(A_{ij})$.

Definition 4.1.27. Für $A \in \mathbb{K}^{n\times n}$ und $i, j \in \{1,\dots,n\}$ heisst $a^{\#}_{ij} = (-1)^{i+j}\det(A_{ij})$ ein *Kofaktor* von A. Die Transponierte der Kofaktormatrix $(a^{\#}_{ij})_{i,j\in\{1,\dots,n\}}$ heisst die *Adjunkte* oder *Komplementärmatrix* von A, und wird mit $A^{\#}$ bezeichnet.

Beispiel 4.1.28. Falls $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$, dann ist $a^{\#}_{21} = (-1)^3\det(A_{12}) = -1$. △

Satz 4.1.29. *Sei $A \in \mathbb{K}^{n\times n}$ invertierbar. Dann gilt $A^{-1} = \frac{A^{\#}}{\det A}$.*

Anders geschrieben: es gilt

$$A^{-1} := \frac{1}{\det A}\begin{pmatrix} \det A_{11} & -\det A_{21} & \dots & (-1)^{n+1}\det A_{n1} \\ -\det A_{12} & \det A_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ (-1)^{n+1}\det A_{1n} & \cdots & \cdots & (-1)^{2n}\det A_{nn} \end{pmatrix}$$

Falls $A = (a_{i,j})_{i\in\{1,\dots,m\},j\in\{1,\dots,n\}}$, dann schreiben wir $a_{i,*}$ für die i-te Zeile von A, und $a_{*,j}$ für die j-te Spalte von A.

Beweis. Wir zeigen $A^{\#}A = \det A \cdot E_n$. Tatsächlich gilt für $i, j \in \{1, \dots, n\}$, dass

$$\begin{aligned}
(A^{\#}A)_{ij} &= \sum_{k \in \{1,\dots,n\}(A^{\#})_{ik} \cdot a_{kj}} \\
&= \sum_{k \in \{1,\dots,n\}} (-1)^{i+k} \det(A_{ki}) \cdot a_{kj} \\
&= \sum_{k \in \{1,\dots,n\}} a_{kj} \cdot \det(A_i(e_k)) && \text{(Lemma 4.1.26)} \\
&= \det(a_{1*}, \dots, a_{(i-1)*}, a_{j*}, a_{(i+1)*}, \dots, a_{n*}) && \text{(Linearität in der } j\text{-ten Spalte)} \\
&= \delta_{ij} \cdot \det(A) && \text{(det ist alternierend).}
\end{aligned}$$

□

Beispiel 4.1.30. Die inverse Matrix von $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ ist

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

△

4.2 Polynomringe

Ein Polynom Was ist das?

$$x^2 + 2x + 1$$

Ausführlicher: Vorlesung *Algebra*.
Trennung von Syntax und Semantik, Polynomen und Polynomfunktionen.

4.2.1 Ringe

Vieles (aber nicht alles) in dieser Vorlesung bleibt gültig, wenn man Körper durch *Ringe* ersetzt.

Definition 4.2.1. Eine Menge R mit zwei binären Operationen, $+$ ('Addition') und $\cdot$ ('Multiplikation'), heißt *Ring*, falls gilt

1. $(R, +)$ ist eine abelsche Gruppe: $+$ ist assoziativ, es gibt ein neutrales Element und inverse Elemente bezüglich $+$, und $+$ ist kommutativ (sieh Abschnitt 2.1)
2. $(R, \cdot)$ ist eine *Halbgruppe*, d.h., die Multiplikation ist assoziativ.
3. Es gelten die Distributivitätsgesetze (vergleiche mit Abschnitt 2.2!)

$$\begin{aligned}
x \cdot (y + z) &= x \cdot y + x \cdot z \\
(y + z) \cdot x &= y \cdot x + z \cdot x
\end{aligned}$$

Ein Ring R heißt

- *Ring mit Eins* falls es ein neutrales Element für die Multiplikation gibt. Falls es so ein Element gibt, so ist es eindeutig (siehe Beweis von Lemma 2.1.3), und wird mit 1 bezeichnet.
- *kommutativer Ring* falls die Multiplikation kommutativ ist.

Ein Element $u \in R$ eines Rings R mit Eins heißt *Einheit* falls es ein *multiplikatives Inverses* hat, d.h., falls es Element $v \in R$ gibt, so dass $vu = uv = 1$.

Beispiel 4.2.2. $(\mathbb{Z}; +, \cdot)$: der Ring der ganzen Zahlen (kommutativ, mit Eins; die einzigen Einheiten sind 1 und -1). △

Beispiel 4.2.3. $(\mathbb{Z}_n, +, \cdot)$: der *Restklassenring* (siehe Abschnitt 1.2.10; kommutativ, mit Eins; Körper falls n prim). △

Beispiel 4.2.4. $(\mathbb{K}^{n\times n}, +, \cdot)$: der *Matrizenring* über $\mathbb{K}$ (nicht kommutativ, siehe Beispiel 3.2.6; aber mit Eins E_n). △

Beispiel 4.2.5. Sei V ein Vektorraum. Dann ist

$$\mathrm{End}(V) := \{f\colon V \to V \mid f \text{ lineare Abbildung}\}$$

der *Endomorphismenring*, mit folgenden Operationen

$$(f_1 + f_2)(v) := f_1(v) + f_2(v)$$
$$(f_1 \cdot f_2)(v) := f_1(f_2(v))$$

Das neutrale Element für die Addition ist der Endomorphismus, der ganz V auf $\mathbf{0}$ abbildet, und für den wir $\underline{0}$ schreiben. △

Bemerkung 4.2.6. Die folgenden Definitionen dieser Vorlesung haben eine natürliche Verallgemeinerung von Körpern auf Ringe:

- Matrizen,
- Determinanten,
- Das Analogon zu Vektorräumen über einem Körper ist der Begriff des *Moduls*[1] über einem Ring.

4.2.2 Polynome über $\mathbb{K}$

Weitere wichtige Beispiele für Ringe sind *Polynomringe*. Polynome über $\mathbb{R}$ sind Ihnen bereits aus der Schule bekannt; die Elemente eines solchen Ringes sind *Polynome*, mit einer geeigneten Addition und Multiplikation. Tatsächlich lassen sich Polynome bereits über einem Ring anstatt eines Körpers betrachten, und wir werden Polynome daher gleich in dieser Allgemeinheit definieren.

[1] Im Unterschied zu anderem sprachlichen Gebrauch wird Modul in diesem Kontext mit Betonung auf dem *o* ausgesprochen.

Sei R ein kommutativer Ring mit Eins (z.B. ein Körper). Eine Abbildung $\phi\colon \mathbb{N} \to R$ heißt auch *Folge*. Schreibweise: $\phi = (a_i)_{i\in\mathbb{N}} = (a_0, a_1, \dots)$ mit $a_i := \phi(i) \in \mathbb{K}$. Sei $\mathcal{F}$ die Menge aller Folgen mit der Eigenschaft, dass $a_i = 0$ für *fast alle* $i \in \mathbb{N}$, d.h., mit Ausnahme von endlich vielen. Auf $\mathcal{F}$ werden folgende Operationen definiert:

- Addition:
$$(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} := (a_i + b_i)_{i\in\mathbb{N}}$$
- Multiplikation mit Skalar $c \in R$:
$$c \cdot (a_i)_{i\in\mathbb{N}} := (c \cdot a_i)_{i\in\mathbb{N}}$$

Bemerkung 4.2.7. Falls R sogar ein Körper $\mathbb{K}$ ist, dann wird $\mathcal{F}$ zu einem $\mathbb{K}$-Vektorraum. Eine (unendliche!) Basis ist

$$\begin{aligned}
&(1, 0, 0, \dots),\\
&(0, 1, 0, \dots),\\
&(0, 0, 1, \dots),\\
&\dots
\end{aligned}$$

Es gilt $\mathcal{F} \leq R^{\mathbb{N}}$, d.h., $\mathcal{F}$ ist ein Untervektorraum vom Funktionsraum $R^{\mathbb{N}}$ (siehe Abschnitt 2.3.1).

Neue Bezeichnungen (X ein beliebiges Symbol):

alt		neu
$(1, 0, 0, \dots)$	$=:$	X^0
$(0, 1, 0, \dots)$	$=:$	X^1
$\vdots$	$\vdots$	$\vdots$
$(0, 0, \dots, 0, 1, 0, \dots)$	$=:$	X^n
$\mathcal{F}$	$=:$	$R[X]$

Es folgt:

$$\begin{aligned}
(k, 0, 0, \dots) &= k \cdot (1, 0, 0, \dots) = k \cdot X^0 =: k \cdot 1 \quad (= k \in R)\\
(0, k, 0, \dots) &= k \cdot (0, 1, 0, \dots) = k \cdot X^1 =: k \cdot X\\
(0, \dots, 0, k, 0, \dots) &= k \cdot (0, \dots, 0, 1, 0, \dots) = k \cdot X^n
\end{aligned}$$

Bemerkung 4.2.8. Man kann R als Teilmenge von $R[X]$ auffassen (und das werden wir im Folgenden tun). Insbesondere steht dann $0 \in R$ für das Element $(0, 0, \dots) \in R^{\mathbb{N}}$.

Haben also

$$(a_0, a_1, \dots, a_n, 0, 0, \dots) = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n$$

und insbesondere

$$(0, 0, \dots) = 0 + 0 \cdot X + 0 \cdot X^2 + \dots = 0.$$

Die Elemente von $R[X]$ heißen *Polynome (über R) in der Unbestimmten X*.

4.2.3 Der Polynomring $R[X]$

In $R[X]$ lässt sich eine Multiplikation wie folgt definieren:

- Für Basiselemente:
$$X^i \cdot X^j := X^{i+j}$$
- Für Linearkombinationen gemäß dem Distributivgesetz:
für $\phi = a_0 + a_1X + \cdots + a_nX^n$ und $\psi = b_0 + b_1X + \cdots + b_mX^m$ gelte
$$\phi \cdot \psi = c_0 + c_1X + \cdots + c_{n+m}X^{n+m}$$
mit $c_k = \sum_{i=0}^{k} a_k b_{k-i}$ für $k \in \mathbb{N}$.

Satz 4.2.9. $(R[X], +, \cdot)$ *ist ein kommutativer Ring mit Eins.*

Die Eins 1 ist neutrales Element für die Multiplikation.
Bezeichnung: *Polynomring über R in der Unbekannten X.*

4.2.4 Der Grad eines Polynoms

Sei
$$\phi = a_0 + a_1X + a_2X^2 + \cdots \in R[X]$$
Definieren den *Grad* des Polynoms ϕ wie folgt:
$$\begin{aligned} \operatorname{grad}(\underline{0}) &:= -\infty \\ \operatorname{grad}(\phi) &:= \max(i \in \mathbb{N} \mid a_i \neq 0\} \end{aligned}$$

Dann gilt:
$$\begin{aligned} \operatorname{grad}(\phi + \psi) &\leq \max(\operatorname{grad}(\phi), \operatorname{grad}(\psi)) \\ \operatorname{grad}(\phi \cdot \psi) &\leq \operatorname{grad}(\phi) + \operatorname{grad}(\psi) \end{aligned} \tag{4.8}$$

wobei $\max(a, -\infty) = \max(-\infty, a) := -\infty$, und $-\infty + a = a + (\infty) := -\infty$ für alle $a \in \mathbb{N} \cup \{-\infty\}$. Falls R sogar ein Körper ist, gilt in (4.8) Gleicheit anstatt $\leq$.

4.2.5 Polynomfunktionen

Nun der bereits angekündigte wichtige Übergang von der Syntax zur Semantik.

Sei $\phi \in R[X]$ ein Polynom,
$$\phi = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n .$$

Sei S ein Ring mit $R \subseteq S$ (z.B. $S = R[X]$, siehe Bemerkung 4.2.8) und $s \in S$. Dann ist
$$\phi^S(s) := a_0 + a_1s + a_2s^2 + \cdots + a_ns^n$$

ein Element von S!
Auswertung von ϕ in S an der Stelle s. "Einsetzen" von s in ϕ.

Die Abbildung

$$\phi^S : S \to S : r \mapsto \phi^S(s)$$

heißt die zu ϕ gehörige *Polynomfunktion*. In der Algebra allgemeiner: 'Termfunktion'.

Wichtig: Unterschied

Polynom (Syntax)	Polynomfunktion (Semantik)
ϕ	ϕ^S

Definition 4.2.10. Ein Element $s \in S$ heißt *Nullstelle von ϕ in S* falls $\phi^S(s) = 0$.

Hier steht 0 für das Nullelement des Ringes S = Nullelement von R.

4.2.6 Der Auswertungshomomorphismus

Satz 4.2.11 (Auswertungssatz)**.** *Es sei S ein Ring und $R \subseteq S$ ein kommutativer Ring mit 1. Sei $s \in S$ so dass $s \cdot r = r \cdot s$ für alle $r \in R$. Dann gilt für alle $\phi, \psi \in R[X]$:*

$$(\phi + \psi)^S(s) = \phi^S(s) + \psi^S(s)$$
$$(\phi \cdot \psi)^S(s) = \phi^R(s) \cdot \psi^S(s)$$

Die Voraussetzungen sind z.B. gegeben, wenn S ebenfalls ein kommutativer Ring ist. Algebraischer Hintergrund: die Abbildung

$$h_s : R[X] \to S : \phi \mapsto \phi^S(s)$$

ist ein (Ring-) Homomorphismus.

Beispiel 4.2.12. Sei $S := \mathbb{K}^{2\times 2}$. Ist $\mathbb{K} \subseteq S$? Eigentlich nicht. Aber schon mit 'Trick' über Einbettung von $\mathbb{K}$ in S: ein Körperelement $k \in \mathbb{K}$ wird als Matrix

$$k \cdot \underbrace{E_2}_{\text{Eins-element im Ring } \mathbb{K}^{2\times 2}} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

interpretiert. Dann gilt $\mathbb{K} \subseteq S$ und für $k \in \mathbb{K}$ und $A \in \mathbb{K}^{2\times 2}$ gilt

$$\begin{aligned} k \cdot A &:= (k \cdot E_2) \cdot A \\ &= A \cdot (k \cdot E_2) && (3.1) \\ &= A \cdot k \end{aligned}$$

und damit ist Satz 4.2.11 anwendbar. Seien $\phi_1 = (-2 + X)$ und $\phi_2 = (3 + X)$. Dann ist

$$\begin{aligned} \phi := \phi_1\phi_2 &= (-2 + X)(3 + X) \\ &= -6 + X + X^2 \end{aligned}$$

Nullstellen von ϕ in $\mathbb{K}$: $2, -3$.
Für z.B. $A = \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} \in \mathbb{K}^{2\times 2} = S$ ergibt sich:

$$\phi_1^S(A) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -5 \end{pmatrix}$$
$$\phi_2^S(A) = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 0 & 0 \end{pmatrix}$$
$$\phi^S(A) = \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix}^2 = \cdots = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Satz 4.2.11 (Auswertungssatz) ergibt

$$\phi(A) = \phi_1(A) \cdot \phi_2(A) = \begin{pmatrix} 0 & 1 \\ 0 & -5 \end{pmatrix} \cdot \begin{pmatrix} 5 & 1 \\ 0 & 0 \end{pmatrix} = 0$$

Nullstellen von ϕ in $\mathbb{K}^{2\times 2}$: z.B. die Matrix A. △

Beispiel 4.2.13. Sei V ein $\mathbb{K}$-Vektorraum. Dann ist $\mathrm{End}(V)$ ein (nicht kommutativer) Ring (Beispiel 4.2.5). Wir werden im Folgenden annehmen, dass $\mathbb{K} \subseteq \mathrm{End}(V)$ ist: das Element $\lambda \in \mathbb{K}$ fassen wir auf als $v \mapsto \lambda\,\mathrm{id}_V$. Also können wir Polynome $\phi \in \mathbb{K}[X]$ auswerten in $\mathrm{End}(V)$. △

4.2.7 Polynomdivision

Teilbarkeitslehre für Polynome ähnlich wie für Zahlen (Vorlesung Algebra).

Definition 4.2.14. Seien $\phi, \psi \in \mathbb{K}[X]$. Dann heißt ϕ *ein Vielfaches von* ψ, und ψ *ein Teiler* von ϕ (Schreibweise: $\psi|\phi$), falls es ein $\phi_1 \in \mathbb{K}[X]$ gibt mit $\phi = \phi_1\psi$.

Polynomdivision: Division mit Rest.

Beispiel:

$$\begin{array}{lllllll}
(& X^5 & +3X^4 & +0\cdot X^3 & +X^2 & +6X & -6) : \quad (X^2+X-1) = X^3+2X^2-X+4 \\
-(& X^5 & +X^4 & -X^3) & & & \\
 & 0 & +2X^4 & -X^3 & & & \\
 & -(& 2X^4 & +2X^3 & -2X^2) & & \\
 & & 0 & -X^3 & 3X^2 & & \\
 & & -(& -X^3 & -X^2 & X) & \\
 & & & 0 & 4X^2 & +5X & -6 \\
 & & & -(& 4X^2 & +4X & -4 \\
 & & & & 0 & X & -2 \qquad \text{‘Rest’ } \rho = X-2
\end{array}$$

Also: $\frac{\phi}{\psi} = \phi_1 + \frac{\rho}{\psi}$, d.h.,

$$\phi_1\psi + \rho$$

wobei $\mathrm{grad}(\rho) < \mathrm{grad}(\psi)$.

Lemma 4.2.15. *Sei $\phi \in \mathbb{K}[X]$ und $k \in \mathbb{K}$. Dann ist k genau dann Nullstelle von ϕ, wenn $(X-k)|\phi$.*

Beweis. Sei $\psi := (X-k)$.
'$\Leftarrow$': $\phi = \phi_1 \cdot \psi$ ergibt mit Satz 4.2.11

$$\phi(k) = \phi_1(k)\psi(k) = \phi_1(k) \cdot 0 = 0$$

'$\Rightarrow$': Polynomdivision liefert $\phi = \phi_1 \cdot (X-k) + \rho$ wobei $\operatorname{grad}(\rho) < \operatorname{grad}(\psi) = 1$. Wegen $\phi(k) = 0$ folgt $\rho(k) = 0$. Da $\operatorname{grad}(\rho) = 0$ ist $\rho \in \mathbb{K}$. Also $\rho = 0$, und damit $\psi|\phi$. □

Definition 4.2.16. Die *algebraische Vielfachheit* einer Nullstelle k ist definiert als

$$\max\{m \in \mathbb{N} : (X-k)^m|\phi\}.$$

Eine *mehrfache* Nullstelle ist entsprechende eine Nullstelle mit algebraischer Vielfachheit größer als 1. Wie zeigt man, dass ein Polynom mehrfache Nullstellen hat? Dafür ist das folgende Lemma oft praktisch. Die *Ableitung* eines Polynoms $\phi(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ ist definiert als das Polynom

$$\phi'(X) := a_1 + 2a_2X + \cdots + na_nX^{n-1}.$$

(Siehe auch Beispiel 3.4.19.)

Lemma 4.2.17. *Ein Polynom $\phi \in \mathbb{K}[X]$ hat genau dann $\lambda \in \mathbb{K}$ als mehrfache Nullstelle, wenn λ sowohl eine Nullstelle von ϕ als auch von ϕ' ist.*

Beweis. Wenn λ eine mehrfache Nullstelle von ϕ ist, dann gilt $\phi(X) = (1-X)^m\psi(X)$ mit $m \geq 2$ (Lemma 4.2.15). Also ist

$$\phi'(X) = m(X-\lambda)^{m-1}\psi(X) + (X-\lambda)^m\psi'(X)$$

was ebenfalls λ als Nullstelle hat.

Umgekehrt: nehmen wir an, dass λ Nullstelle von sowohl ϕ als auch von ϕ' ist. Dann können wir schreiben $\phi(X) = (X-\lambda)\psi(X)$, und $\phi'(X) = \psi(X) + (X-\lambda)\psi'(X)$. Also $0 = \phi'(X)(\lambda) = \psi(\lambda) + (\lambda-\lambda)\psi'(\lambda) = \psi(\lambda)$ und damit ist λ Nullstelle von ψ. Also ist λ mehrfache Nullstelle von ϕ. □

4.3 Eigenwerte, Eigenvektoren, Diagonalisierbarkeit

Viele Anwendungen in der Physik, Stochastik, diskreten Mathematik, ...

Definition 4.3.1. Sei V ein $\mathbb{K}$-Vektorraum, und $f: V \to V$ ein Endomorphismus. (D.h., f ist eine lineare Abbildung, siehe Abschnitt 3.4.) Ein Element $\lambda \in \mathbb{K}$ heißt *Eigenwert (EW) von f*, falls es einen Vektor $v \neq \mathbf{0}$ gibt, so dass:

$$f(v) = \lambda v \tag{4.9}$$

Jeder Vektor $v \neq \mathbf{0}$ mit dieser Eigenschaft heißt *Eigenvektor* von f zum Eigenwert λ.

Speziell: *Eigenwert* $\lambda \in \mathbb{K}$ und *Eigenvektor* $u \in \mathbb{K}^{n\times n}$ einer Matrix $A \in \mathbb{K}^{n\times n}$: definiert als EW und Eigenvektor von

$$f_A\colon \mathbb{K}^n \to \mathbb{K}^n : x \mapsto Ax\,,$$

d.h.,

$$Au = \lambda u \qquad\qquad \text{mit } u \neq \mathbf{0}\,.$$

Bemerkung 4.3.2. Der Nullvektor $v = \mathbf{0}$ erfüllt (4.9) trivialerweise, ist aber *kein* Eigenvektor. Der Eigenwert 0 tritt genau dann auf, wenn $\operatorname{Kern}(f) \neq \{\mathbf{0}\}$ (also genau dann, wenn f nicht injektiv ist, bzw. wenn $\det(f) = 0$; Satz 3.2.27).

Definition 4.3.3. Seien $\lambda \in \mathbb{K}$ und $f\colon V \to V$.

$$\operatorname{Eig}_\lambda(f) := \operatorname{Eig}_\lambda := \{v \in V \mid f(v) = \lambda v\} \tag{4.10}$$

heißt *Eigenraum* von f zum Eigenwert λ.

$\operatorname{Eig}_\lambda(f)$ ist Untervektorraum von V, denn

$$\begin{aligned}\operatorname{Eig}_\lambda(f) &= \{v \in V \mid f(v) - \lambda v = \mathbf{0}\}\\ &= \{v \in V \mid (f - \lambda\,\mathrm{id})(v) = \mathbf{0}\}\\ &= \operatorname{Kern}(f - \lambda\,\mathrm{id}) \leq V.\end{aligned}$$

Die Dimension von $\operatorname{Eig}_\lambda$ heißt *geometrische Vielfachheit* von λ.
Spezialfall $f = f_A\colon \mathbb{K}^n \to \mathbb{K}^n : x \mapsto Ax$:

$$\operatorname{Eig}_\lambda(A) := \operatorname{Eig}_\lambda(f_A) = \operatorname{Kern}(A - \lambda E_n)$$

Also:

$$\begin{aligned}\dim(\operatorname{Eig}_\lambda(A)) &= \dim(\operatorname{Kern}(A - \lambda E))\\ &= n - \operatorname{rg}(A - \lambda E)\end{aligned}$$

ist die geometrische Vielfachheit von λ nach der Dimensionsformel (Satz 3.3.6).

Bemerkung 4.3.4. Seien V ein $\mathbb{K}$-Vektorraum mit Basis $B = (v_1, \ldots, v_n)$, $f\colon V \to V$ ein Endomorphismus, und $A := M_B^B(f)$ Darstellungsmatrix von f (siehe Abschnitt 3.4.5). Dann haben A und f die gleichen Eigenwerte, und $\operatorname{Eig}_\lambda(f) \simeq \operatorname{Eig}_\lambda(A)$. Genauer: sei $\Phi_B\colon \mathbb{K}^n \to V$ der kanonische Basisisomorphismus (Abschnitt 2.4.3). Dann gilt

$$\begin{aligned}Ax = \lambda x &\Leftrightarrow \Phi_B(Ax) = \Phi_B(\lambda x)\\ &\Leftrightarrow f(\Phi_B(x)) = \lambda\Phi_B(x)\end{aligned}$$

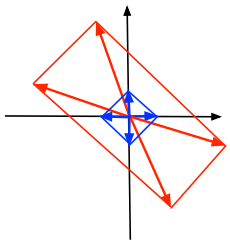
Ist x der Koordinatenvektor von v bzgl. Basis B (das heißt, $\Phi_B(x) = v$) dann gilt:

$$\begin{aligned}&v \text{ ist Eigenvektor von } f \text{ zu EW } \lambda\\ \Leftrightarrow\ &x \text{ ist Eigenvektor von } A \text{ zu EW } \lambda\end{aligned} \tag{4.11}$$

Beispiel 4.3.5. $f\colon \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Ax$ lineare Abbildung mit Darstellungsmatrix

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

Was macht f?



$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto Ae_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$
$$e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto Ae_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

Experimentieren:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad Av_1 = \begin{pmatrix} 3-1 \\ -1+3 \end{pmatrix} = 2v_1$$
$$v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad Av_2 = \begin{pmatrix} 3+1 \\ -1-3 \end{pmatrix} = 4v_2$$

$\lambda_1 = 2$ Eigenwert, $v_1 = \binom{1}{1}$ Eigenvektor.
$\lambda_2 = 4$ Eigenwert, $v_1 = \binom{1}{-1}$ Eigenvektor.
$B := (v_1, v_2)$ ist sogar Basis. Muss nicht immer sein!
Für beliebigen Vektor $v \in \mathbb{R}^2$

$$v = \alpha_1 v_1 + \alpha_2 v_2$$

folgt

$$\begin{aligned} f(v) &= \alpha_1 f(v_1) + \alpha_2 f(v_2) \\ &= 2\alpha_1 v_1 + 4\alpha_2 v_2 \end{aligned}$$

Das heißt Streckung um Faktor 2 in Richtung v_1, und um Faktor 4 in Richtung v_2.
Das kann man aus der Darstellungsmatrix $M_B^B(f)$ direkt ablesen:

$$M_B^B(f) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

Ist Diagonalmatrix (Beispiel 3.2.8) mit Eigenwerten auf der Diagonale. △

Diagonalmatrix erstrebenswert:

- nur n Werte (statt n^2);
- alle Rechnungen (Inverse, Determinante, etc.) einfacher;
- Verhalten der Abbildung ablesbar;
- EW ablesbar.

4.3.1 Anwendung: Pagerank

Webseiten $S := \{1, \dots, n\}$.
Links zwischen Seiten: Teilmenge L von S^2.
Wichtigkeit $0 \le w(1), \dots, w(n) \in \mathbb{R}$ (für Ranking).
Wie könnte sinnvolles Ranking funktionieren?

Idee. Eine Seite ist wichtig, wenn es viele Links von wichtigen Seiten auf diese Seite gibt.

$$w(i) \sim \sum_{j:(j,i)\in L} w(j)$$

Formal:

$$w(i) = \tilde{\lambda} \sum_{j=1}^{n} a_{ji} w(j)$$

wobei $a_{ij} := 1$ falls $(i,j) \in L$ und $a_{ij} := 0$ sonst. Also

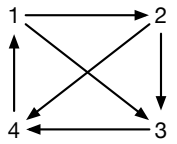
$$\begin{pmatrix} w(1) \\ \vdots \\ w(n) \end{pmatrix} =: x = \tilde{\lambda} A x \qquad \text{für } A = (a_{ji})_{i,j=1,\dots,n}$$

D.h., für Ranking wird gebraucht: ein positiver Eigenwert $\lambda = 1/\tilde{\lambda}$ und ein *positiver* Eigenvektor x (alle Einträge positiv):

$$Ax = \lambda x.$$

Beispiel 4.3.6. (Turnier) Teams $1, 2, 3, 4$.

$$A = (a_{ij}) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$



Interpretation der Matrix:
$a_{ij} = 1$: Team i schlägt Team j, sonst $a_{ij} = 0$.

Eigenvektor

$$x \approx \begin{pmatrix} 0.62 \\ 0.55 \\ 0.32 \\ 0.45 \end{pmatrix}$$

zum Eigenwert $\approx 1,39$ (einziger positiver EW). △

4.3.2 Berechnung von Eigenwerten und das charakteristische Polynom

Sei $A \in \mathbb{K}^{n\times n}$ bzw. $f\colon V \to V$ Endomorphismus mit $A = M_B^B(f)$ (bezüglich Basis B).

Definition 4.3.7. Das *charakteristische Polynom*[2] von A, beziehungsweise von f, ist das folgende Polynom aus $\mathbb{K}[X]$:

$$\chi_f(X) := \chi_A(X) := \det(XE - A) = \begin{vmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & X - a_{nn} \end{vmatrix}.$$

Bemerkung 4.3.8. Definition funktioniert auch, wenn statt Körper $\mathbb{K}$ nur ein Ring R verwendet wird (Definition 4.2.1; $A \in R^{n\times n}$, $\chi_f \in R[X]$).

Proposition 4.3.9. *Ähnliche Matrizen haben dasselbe charakteristische Polynom.*

Beweis. Für $B = S^{-1}AS$ gilt

$$\begin{aligned} \det(XE - B) &= \det(XS^{-1}S - S^{-1}AS) \\ &= \det(S^{-1}(XE - A)S) \\ &= \det S^{-1} \cdot \det(XE - A) \cdot \det S \\ &= \det(XE - A)\,. \end{aligned}$$

□

Also ist $\chi_f(X)$ unabhängig von der Wahl der Basis B. Ansonsten wäre Definition 4.3.7 so gar nicht möglich.

Satz 4.3.10. *Die Eigenwerte sind* genau *die Nullstellen des charakteristischen Polynoms, d.h.,*

$$\lambda \in \mathbb{K} \textit{ ist EW von } A \quad \Leftrightarrow \quad \det(\lambda E - A) = 0.$$

Beweis.

$$\begin{aligned} \lambda \text{ EW } &\Leftrightarrow \exists v \in \mathbb{K}^n \setminus \{\mathbf{0}\} : v \in \operatorname{Kern}(\lambda E - A) && \text{(Definition 4.3.1)} \\ &\Leftrightarrow \det(\lambda E - A) = 0 && \text{(Abschnitt 4.1.7)} \end{aligned}$$

Die zweite Gleichung folgt aus den Beobachtungen aus Abschnitt 4.1.7: ein homogenes LGS $Bx = \mathbf{0}$ (stets lösbar!) hat genau dann eine eindeutige Lösung, wenn $\det B \neq 0$. □

Beispiel 4.3.11. Betrachte

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

[2] Manche Autor:innen definieren das charakteristische Polynom von A als $\det(A - \lambda E)$. Unsere Definition hat den Vorteil, dass der führende Eintrag des Polynoms stets 1 ist. Allerdings macht das keinen großen Unterschied, da sich die eine Variante der Definition durch Multiplikation mit $(-1)^n$ aus der anderen ergibt.

$$\begin{aligned}
\chi_A(\lambda) = \det(XE - A) &= \begin{vmatrix} X & 1 & -1 \\ 3 & X+2 & -3 \\ 2 & 2 & X-3 \end{vmatrix} && \text{(Definition)} \\
&= X(X+2)(X-3) - 6 - 6 + 2(X+2) + 6X - (X-3)3 && \text{(Sarrus, Beispiel 4.1.9)} \\
&= X^3 - 3X^2 + 2X^2 - 6X - 12 + 2X + 4 + 6X - 3X + 9 && \text{(Ausmultiplizieren)} \\
&= X^3 - X^2 - X + 1 && \text{(Vereinfachen)} \\
&= (X-1)^2(X+1) && \text{(Faktorisieren)}
\end{aligned}$$

Also: haben folgende Nullstellen

$$\begin{aligned}
\lambda_1 &= 1 && \text{(algebraische Vielfachheit 2)} \\
\lambda_2 &= -1 && \text{(algebraische Vielfachheit 1)}
\end{aligned}$$

Geometische Vielfachheiten werden später ausgerechnet (Beispiel 4.3.21). △

Beispiel 4.3.12. $V = \mathbb{R}^2$

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Charakteristisches Polynom:

$$\begin{aligned}
\chi_A(X) &= \cos^2\alpha - 2X\cos\alpha + \sin^2\alpha \\
&= X^2 - 2X\cos\alpha + 1
\end{aligned}$$

Eigenwerte: die Nullstellen von $\chi_A(X)$.
Fallunterscheidung:

- $\alpha = 0$.
$$\chi_A(X) = X^2 - 2X + 1 = (X-1)^2$$
Eigenwert 1, algebraische Vielfachheit 2.

- $\alpha = 180^\circ$.
$$\chi_A(X) = X^2 + 2X + 1 = (+1)^2$$
Eigenwert -1, algebraische Vielfachheit 2.

- $\alpha \neq 0$ und $\alpha \neq 180^\circ$.
$$\chi_A(X) = X^2 - (2\cos\alpha)X + 1$$
hat keine Nullstellen in $\mathbb{R}$.
p,q-Formel: $q = 1$, $p/2 = -\cos\alpha$. Haben die Lösungen
$$-\cos\alpha \pm \sqrt{\cos^2\alpha - 1}$$
mit $\cos^2\alpha - 1 < 0$ für $\alpha \neq \{0^\circ, 180^\circ\}$. △

Beispiel 4.3.13. Die Eigenwerte einer Dreiecksmatrix

$$A = \begin{pmatrix} a_{11} & & & * \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

sind die Elemente der Hauptdiagonalen (algebraische Vielfachheit ist dabei schon berücksichtigt), denn

$$\begin{aligned} \chi_A(X) &= \det(XE - A) \\ &= (X - a_{11})(X - a_{22})\cdots(X - a_{nn}). \end{aligned}$$

△

Bemerkung 4.3.14. Für das charakteristische Polynom

$$\begin{aligned} \chi_A(X) &= \det(XE - A) \\ &= a_n X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \end{aligned}$$

einer Matrix $A \in \mathbb{K}^{n\times n}$ gilt

1. $a_0 = \det -A$ Setze $X = 0$
2. $a_n = 1$
3. $a_{n-1} = (-1)^{n-1}(a_{11}+\cdots+a_{nn}) =: (-1)^{n-1}\operatorname{Spur}(A)$ Summe der Hauptdiagonalen

Beweis durch Auswerten der Leibnizschen Formel. Einzig der Summand

$$(X - a_{11})\cdots(X - a_{nn})$$

der Determinante für die Permutation $\mathrm{id} \in S_n$ ist relevant, da es nur dort $n-1$ Auftreten von X geben kann (alle anderen Permutationen unterscheiden sich von id an mindestens zwei Stellen, die entsprechenden Summanden haben also Grad höchstens $n-2$). Ausmultiplizieren dieses Summanden liefert Koeffizienten $(-1)^{n-1}\operatorname{Spur}(A)$ für X^{n-1}.

Übung 18. Beweisen Sie: für $A, B \in \mathbb{K}^n$ gilt $\operatorname{Spur}(AB) = \operatorname{Spur}(BA)$.

Übung 19. Seien $A_1, \ldots, A_n \in \mathbb{K}^n$ und $\pi \in S_n$ eine Permutation. Gilt $\operatorname{Spur}(A_1, \ldots, A_n) = \operatorname{Spur}(A_{\pi(1)}\cdots A_{\pi(n)})$?

Übung 20. Zeigen Sie: für quadratische Matrizen A und B und

$$M := \begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix}$$

gilt $\chi_M = \chi_A \cdot \chi_B$.

Kommentare. (Erinnerung: Nullstellen $\leftrightarrow$ Linearfaktoren, Lemma 4.2.15)
Sätze und Algorithmen zur Faktorisierung univariater Polynome $\phi \in \mathbb{K}[X]$:

- über $\mathbb{K} = \mathbb{C}$: jedes Polynom zerfällt in Linearfaktoren. Wenn man bereits eine Nullstelle a kennt (numerische Verfahren), so führt man Polynomdivision durch $(X - a)$ durch und wendet das Verfahren rekursiv auf den Quotienten an.
- über $\mathbb{K} = \mathbb{R}$: faktorisieren in $\mathbb{C}$, und beobachten, dass mit jeder komplexen Nullstellen $a + b \cdot i$, für $a, b \in \mathbb{R}$, auch die *konjugiert komplexe* $a - i \cdot b$ eine Nullstelle ist.

  Also treten neben den Linearfaktoren auch Faktoren auf der Gestalt

$$((X - (a + b \cdot i))(X - (a - b \cdot i)) = X^2 + \underbrace{2a}_{\in\mathbb{R}} X + (\underbrace{a^2 + b^2}_{\in\mathbb{R}}).$$

- über endlichen Körpern $\mathbb{K} = \mathbb{F}_q$: Berlekamp-Algorithmus [5].
- über $\mathbb{K} = \mathbb{Q}$: Lenstra-Lenstra-Lovász Algorithmus [3].

4.3.3 Diagonalmatrizen

Erinnern uns an Beispiel

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

aus Abschnitt 4.3. Eigenvektoren bilden Basis $B = \{\binom{1}{1}, \binom{1}{-1}\}$,
und Darstellungsmatrix von f_A diagonal:

$$M_B^B(f_A) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

Wann ist das der Fall?

Lemma 4.3.15. *Sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = (v_1, \ldots, v_n)$, und $f\colon V \to V$ ein Endomorphismus. Dann sind äquivalent:*

1. *Die Darstellungsmatrix von f bezüglich B ist Diagonalmatrix:*

$$A := M_B^B(f) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

2. *B ist Basis aus Eigenvektoren von f, genauer $f(v_i) = \lambda_i v_i$ für $i \in \{1, \ldots, n\}$.*

Beweis. (2) $\Rightarrow$ (1): klar (siehe Abschnitt 3.4.5: Merkregel!)
(1) $\Rightarrow$ (2): Offenbar $Ae_i = \lambda_i e_i$ (i-te Spalte von A)
Also ist e_i Eigenvektor von A zu EW λ_i.
Also ist $v_i = \Phi_B(e_i)$ Eigenvektor von f zu EW λ_i (siehe (4.11)).
Das bedeutet, $f(v_i) = \lambda_i v_i$. □

Ziel im Folgenden: möglichst viele linear unabhängige Eigenvektoren finden.
Verschiedene Eigenwerte sichern lineare Unabhängigkeit!

Lemma 4.3.16. *Seien $v_1, \dots, v_r$ Eigenvektoren von $f \in \mathrm{End}(V)$ zu verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$. Dann sind $v_1, \dots, v_r$ linear unabhängig.*

Beweis. Induktion über r. Für $r = 1$ ist $v_1 \neq \mathbf{0}$ linear unabhängig. Sei nun die Aussage richtig für $r = k \geq 1$; zu zeigen ist die Aussage für $r = k + 1$. Seien $v_1, \dots, v_k, v_{k+1}$ Eigenvektoren zu EW $\lambda_1, \dots, \lambda_k, \lambda_{k+1}$ (paarweise verschieden). O.B.d.A $\lambda_{k+1} \neq 0$ (sonst andere Nummerierung). Sei

$$\alpha_1 v_1 + \cdots + \alpha_{k+1} v_{k+1} = \mathbf{0} \tag{4.12}$$

Dann gilt:

$$\alpha_1 \lambda_{k+1} v_1 + \cdots + \alpha_k \lambda_{k+1} v_k + \alpha_{k+1} \lambda_{k+1} v_{k+1} = \mathbf{0} \quad (\lambda_{k+1} \cdot (4.12)) \tag{4.13}$$

$$\underbrace{\alpha_1 \underbrace{\lambda_1 v_1}_{=f(v_1)}}_{=f(\alpha_1 v_1)} + \cdots + \alpha_k \lambda_k v_k + \alpha_{k+1} \lambda_{k+1} v_{k+1} = \mathbf{0} \quad (\text{Anwenden von } f \text{ auf } (4.12)) \tag{4.14}$$

$$\alpha_1 (\lambda_1 - \lambda_{k+1}) v_1 + \cdots + \alpha_k (\lambda_k - \lambda_{k+1}) v_k = \mathbf{0} \quad (\text{Subtraktion } (4.14) \text{ - } (4.13)$$

Nach Induktionsvoraussetzung sind $v_1, \dots, v_k$ linear unabhängig, also

$$\alpha_1 (\lambda_1 - \lambda_{k+1}) = \cdots = \alpha_k (\lambda_k - \lambda_{k+1}) = 0 \,.$$

Wegen $\lambda_i \neq \lambda_{k+1}$ ist $\lambda_i - \lambda_{k+1} \neq 0$, für alle $i \in \{1, \dots, k\}$, und daher

$$\alpha_1 = \cdots = \alpha_k = 0$$

Aus (4.12) folgt nun

$$\alpha_{k+1} \underbrace{v_{k+1}}_{\neq \mathbf{0}} = \mathbf{0}$$

also auch $\alpha_{k+1} = 0$. □

Definition 4.3.17. Sei V in $\mathbb{K}$-Vektorraum, $f \in \mathrm{End}(V)$, und $A \in \mathbb{K}^{n \times n}$.

- f heißt *diagonalisierbar*, wenn es eine Basis von V gibt, die aus Eigenvektoren von f besteht (Motivation: Lemma 4.3.15);
- A heißt *diagonalisierbar*, wenn es eine invertierbare Matrix $S \in \mathrm{GL}(\mathbb{K}, n)$ gibt, so dass $A' := S^{-1} A S$ eine Diagonalmatrix ist. In anderen Worten: A ist genau dann diagonalisierbar, wenn A ähnlich ist zu einer Diagonalmatrix A'.

Bemerkung 4.3.18. Definition sinnvoll, denn für jede Basis B von V ist f genau dann diagonalisierbar, wenn $M_B^B(f)$ diagonalisierbar.

Satz 4.3.19 (Diagonalisierbarkeitskriterium)**.** *Es seien:*

- *V ein n-dimensionaler $\mathbb{K}$-Vektorraum,*
- *$f \in \mathrm{End}(V)$,*
- *$A = M_B^B(f)$ Darstellungsmatrix von f bezüglich einer Basis B von V,*
- *$\lambda_1, \dots, \lambda_r$ alle paarweise verschiedenen Eigenwerte von f (bzw. von A),*
- *$n_1, \dots, n_r$ die zugehörigen geometrischen Vielfachheiten, $n_i = \dim \mathrm{Kern}(A - \lambda_i E)$,*
- *$(v_1^{(i)}, \dots, v_{n_i}^{(i)})$ sei Basis des Eigenraums $\mathrm{Eig}_{\lambda_i}(f) = \mathrm{Kern}(f - \lambda_i \,\mathrm{id})$,*
- *$m_1, \dots, m_r$ die algebraischen Vielfachheiten von $\lambda_1, \dots, \lambda_r$, das heißt,*

$$m_i = \max\{m \in \mathbb{N} \mid \exists \psi \in \mathbb{K}[\lambda] \colon \chi_f(X) = (X - \lambda_i)^m \psi\}\,.$$

Dann gilt

1. *$(v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(r)}, \dots, v_{n_r}^{(r)})$ ist linear unabhängig.*
2. *$n_i \le m_i$ und $\sum_{i=1}^r n_i \le \sum_{i=1}^r m_i \le n$.*
3. *Die folgenden Aussagen sind äquivalent:*
   a) *f ist diagonalisierbar;*
   b) *Es gibt eine Basis von $\mathbb{K}^n$, die nur aus Eigenvektoren von A besteht.*
   c) *A ist diagonalisierbar; in diesem Fall ist für jede Matrix S, deren Spalten $u_1, \dots, u_n$ linear unabhängige Eigenvektoren von A sind, $S^{-1}AS$ eine Diagonalmatrix.*
   d) *Das charakteristische Polynom χ_A zerfällt in Linearfaktoren*
   $$\chi_A(X) = (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$$
   und $n_i = m_i$ für $i \in \{1, \dots, r\}$.
   e) *$\sum_{i=1}^r n_i = n = \dim V$.*

Bemerkung 4.3.20. Unmittelbare Folgerung aus $(e) \Rightarrow (a)$: Falls $r = n$, also wenn es n verschiedene Eigenwerte gibt, dann ist f diagonalisierbar. Dies ist selbstverständlich nur ein hinreichendes, nicht aber ein notwendiges Kriterium: denke an die Diagonalmatrix E_2, die nur einen Eigenwert hat.

Beweis. Zu 1.:

$$\underbrace{\alpha_1^{(1)} v_1^{(1)} + \dots + \alpha_{n_1}^{(1)} v_{n_1}^{(1)}}_{=:w_1 \in \mathrm{Eig}_{\lambda_1}} + \dots + \underbrace{\alpha_1^{(r)} v_1^{(r)} + \dots + \alpha_{n_r}^{(r)} v_{n_r}^{(r)}}_{=:w_r \in \mathrm{Eig}_{\lambda_r}} = \mathbf{0} \tag{4.15}$$

Definiere $S := \{i \in \{1, \dots, r\} \mid w_i \neq \mathbf{0}\}$.

- 1. Fall: $S \neq \varnothing$. Die Menge $\{w_i \mid i \in S\}$ besteht aus Eigenvektoren zu paarweise verschiedenen Eigenwerten. Nach Lemma 4.3.16 sind $w_1, \dots, w_r$ linear unabhängig. Dann gilt $\sum_{i \in S} w_i \neq 0$, im Widerspruch zu (4.15) und der Definition von S.

- 2. Fall: $S = \varnothing$. Dann gilt für jedes $i \in \{1, \dots, n\}$

$$w_i = \alpha_1^{(i)} v_1^{(i)} + \cdots + \alpha_{n_i}^{(i)} v_{n_1}^{(i)} = \mathbf{0}$$

  und daher $\alpha_1^{(i)} = \cdots = \alpha_{n_i}^{(1)} = 0$ da $v_1^{(i)}, \dots, v_{n_i}^{(i)}$ Basis bilden.

Zu 2.: Die Basis $(v_1^{(i)}, \dots, v_{n_i}^{(i)})$ von Eig_{λ_i} lässt sich zu Basis $\tilde{B} := (v_1^{(i)}, \dots, v_{n_i}^{(i)}, \dots, v_n{}^{(i)})$ von V ergänzen (Steinitz'scher Austauschsatz: Satz 2.4.13). Die Darstellungsmatrix hat dann die Form

$$M := M_{\tilde{B}}^{\tilde{B}}(f) = \left(\begin{array}{ccc|c} \lambda_i & \cdots & 0 & \\ \vdots & \ddots & \vdots & * \\ 0 & \cdots & \lambda_i & \\ \hline & \mathbf{0} & & * \\ & & & \end{array}\right)$$

Denn: die ersten n_i Spalten: Koordinatenvektoren der Bilder von $v_1^{(i)}, \dots, v_{n_i}^{(i)}$ nach Merkregel, da $f(v_j^{(i)}) = \lambda_i v_j^{(i)}$.
Also

$$\chi_f = \det(\lambda E - M) = (X - \lambda_i)^{n_i} \cdot \text{Restpolynom}$$

d.h., $n_i \leq m_i$. Wegen

$$\mathrm{grad}(\phi \cdot \psi) = \mathrm{grad}(\phi) + \mathrm{grad}(\psi)$$

folgt

$$\sum_{i=1}^{r} m_i \leq \mathrm{grad}(\chi_f) = n\,.$$

Zu 3.: Wir zeigen $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (a)$.

$(a) \Rightarrow (b)$: Da f diagonalisierbar, hat V eine Basis $C = (w_1, \dots, w_n)$ aus Eigenvektoren von f (Lemma 4.3.15). Die Koordinatenvektoren $u_1 = \Phi_B^{-1}(w_1), \dots, u_n = \Phi_B^{-1}(w_n)$ bilden Basis von $\mathbb{K}^n$ aus Eigenvektoren von A.

$(b) \Rightarrow (c)$: Sei $u_1, \dots, u_n$ eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A, und sei

$$S = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}.$$

Dann gilt für eine Diagonalmatrix $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, dass

$$SD = \begin{pmatrix} \lambda_1 u_1 & \cdots & \lambda_n u_n \end{pmatrix} \qquad \text{(Matrizenmultiplikation)}.$$

Also gilt

$$SD = AS = \begin{pmatrix} Au_1 & \cdots & Au_n \end{pmatrix}$$

genau dann wenn $Au_i = \lambda u_i$ für $i = 1, \ldots, n$, d.h., genau dann, wenn die Spalten Eigenvektoren sind. Da $(u_1, \ldots, u_n)$ Basis von $\mathbb{K}^n$ ist, folgt dass

$$\operatorname{rg}(S) = n \Rightarrow S \in \operatorname{GL}(n, \mathbb{K})$$

und

$$SD = AS \Leftrightarrow D = S^{-1}AS.$$

Also $(b) \Rightarrow (c)$.

$(c) \Rightarrow (d)$: A und

$$D = S^{-1}AS =: \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

sind ähnlich, haben also das gleiche charakteristische Polynom (Proposition 4.3.9)

$$\begin{aligned} \chi_A(X) = \chi_D(X) &= (X - d_1)(X - d_2)\cdots(X - d_n) \\ &=: (X - \lambda_1)^{m_1}\cdots(X - \lambda_r)^{m_r} \qquad \text{(zusammenfassen)} \end{aligned}$$

d.h., zu λ_i gibt es genau m_i verschiedene Indizes mit $\lambda_i = d_{t_1} = \cdots = d_{t_{m_i}}$. Die zugehörigen Spalten $u_{t_1}, \ldots, u_{t_{m_i}}$ von A sind linear unabhängige (nach Voraussetzung) Eigenvektoren von λ_i, d.h., $m_i \leq n_i$. Daher $n_i = m_i$.

$(d) \Rightarrow (e)$: Mit (d) gilt $n = \operatorname{grad}(\chi_f) = \sum_i m_i = \sum_i n_i$ und daher (e).

$(e) \Rightarrow (a)$: Wenn $\sum_i n_i = n$, dann bilden die Eigenvektoren in 1. eine Basis (wegen $n = \dim V$). □

4.3.4 Wie diagonalisiert man eine Matrix?

Sei $A \in \mathbb{K}^{n \times n}$. Bestimmung einer Basis aus Eigenvektoren.

- 1. Schritt: Bestimmung aller Eigenwerte von A (Verfahren aus Abschnitt 4.3.2).
- 2. Schritt: Zu jedem Eigenwert λ wird eine Basis des Eigenraums $\operatorname{Eig}_\lambda(A)$ bestimmt.

- 3. Schritt: Ergeben alle Basen aus Schritt 2 insgesamt n Vektoren, so bilden diese eine Basis von V aus Eigenvektoren und A ist diagonalisierbar, sonst nicht. Nimmt man diese Eigenvektoren von A als Spalten einer Matrix S, so liefert diese die Diagonalisierung $S^{-1}AS$.

Beispiel 4.3.21. $n = 3$.

$$f\colon \mathbb{R}^3 \to \mathbb{R}^3 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} -y + z \\ -3x - 2y + 3z \\ -2x - 2y + 3z \end{pmatrix}$$

lineare Abbildung $f = f_A : u \mapsto Au$ mit

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

$A = M_B^B(f)$ für $B = (e_1, e_2, e_3)$ Standardbasis von $\mathbb{K}^3$.

Diagonalisierbar? $D = S^{-1}AS$?

- 1. Schritt. Bestimmung der Eigenwerte.
  Beispiel 4.3.11: Eigenwerte
  - $\lambda_1 = 1$ mit algebraischer Vielfachheit $m_1 = 2$, und
  - $\lambda_2 = -1$ mit algebraischer Vielfachheit $m_2 = 1$.

$$\chi_A(X) = (X - 1)^2(X + 1)$$

- 2. Schritt.
  - Bestimmung einer Basis von $\mathrm{Eig}_{\lambda_1}(A) = \mathrm{Kern}(A - \lambda_1 E_3) = \mathrm{Lös}(A - \lambda_1 E_3, \mathbf{0})$.

$$A - \lambda_1 E_3 = \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}$$

$\Rightarrow \quad \mathrm{rg}(A - \lambda_1 E_3) = 1$

$\Rightarrow \quad \dim(\mathrm{Kern}(A - \lambda_1 E_3)) = 3 - 1 = 2$ ist geometrische Vielfachheit von λ_1.

Gesucht: Lösungen des Gleichungssystems

$$(A - \lambda_1 E_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Im Allgemeinen mit dem Gaußschen Algorithmus (Abschnitt 3.3.4), hier auch direkt klar:

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

Rang 1, also dim(Lös) = 3 − 1 = 2
Lösung:

$$\begin{aligned} x_3 &= \mu_2 && \text{freier Parameter } \mu_2 \in \mathbb{R} \\ x_2 &= \mu_1 && \text{freier Parameter } \mu_2 \in \mathbb{R} \\ x_1 &= -\mu_1 + \mu_2 \end{aligned}$$

Basis für Lösungsraum: Einsetzen einer Basis für die Parameter $\binom{\mu_1}{\mu_2}$, z.B. Einheitsvektoren.

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow u_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow u_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

(u_1, u_2) ist Basis für Eigenraum $\mathrm{Eig}_{\lambda_1}(A)$.

– Bestimmung einer Basis von $\mathrm{Eig}_{\lambda_2}(A) = \mathrm{Kern}(A - \lambda_2 E_3) = \mathrm{Lös}(A - \lambda_2 E_3, \mathbf{0})$.

$$A - \lambda_2 E_3 = \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix}$$

Gaußscher Algorithmus:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -3 & -1 & 3 & 0 \\ -2 & -2 & 4 & 0 \end{array}\right) \xrightarrow{2z_1 + z_3 \rightsquigarrow z_3} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -3 & -1 & 3 & 0 \\ 0 & -4 & 6 & 0 \end{array}\right)$$

$$\xrightarrow{3z_1 + z_2 \rightsquigarrow z_2} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -4 & 6 & 0 \\ 0 & -4 & 6 & 0 \end{array}\right)$$

$$\xrightarrow{z_3 - z_2 \rightsquigarrow z_3} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -4 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Rang = 2, also dim(Lös) = 3 − 2 = 1.
Lösung:

$$\begin{aligned} x_3 &= \mu && \text{freier Parameter } \mu \in \mathbb{R} \\ x_2 &= 3/2\mu \\ x_1 &= x_2 - x_3 = \mu/2 \end{aligned}$$

Basis für Eigenraum $\mathrm{Eig}_{\lambda_2}(A)$: setze μ beliebig, z.B. $\mu = 2$, erhalten

$$u_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.$$

- 3. Schritt. Die Basen von Eig_{λ_1} und Eig_{λ_2} ergeben zusammen 3 Vektoren, also Basis von $V = \mathbb{R}^3$. Also ist A diagonalisierbar. Die Matrix S ist gegeben durch

$$S = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

△

Übung 21. Ist

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

diagonalisierbar? Ist A in $\mathbb{C}^{2\times 2}$ diagonalisierbar?

Übung 22. Ist

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

diagonalisierbar? Ist A in $\mathbb{C}^{2\times 2}$ diagonalisierbar?

Übung 23. Stimmen Sie der folgenden Aussage zu: *'die meisten Matrizen in $\mathbb{R}^{2\times 2}$ sind diagonalisierbar'*? Falls ja, warum?

Bemerkung 4.3.22. Diagonalisierbarkeit $D = S^{-1}AS$ ist nützlich auch für Berechnung von Potenzen von A:

$$\begin{aligned} A &= SDS^{-1} \\ A^2 &= SDS^{-1}SDS^{-1} = SD^2S^{-1} \\ &\vdots \\ A^m &= SD^mS^{-1} \end{aligned}$$

Weiterhin:

$$D = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \quad \Rightarrow \quad D^m = \begin{pmatrix} (a_1)^m & & 0 \\ & \ddots & \\ 0 & & (a_n)^m \end{pmatrix}$$

leicht berechenbar.

4.3.5 Anwendung: Lineares Wachstum

Population: t_m Löwenzahnpflanzen im Jahr m.
Wachstum entsprechend der Gleichung

$$t_{m+1} = w_1 t_m + w_2 t_{m-1} + w_3 \tag{4.16}$$

Konkret: Für $t_0 = 0, t_1 = 1, w_1 = w_2 = 1, w_3 = 0$, d.h.,

$$t_{m+1} = t_m + t_{m-1}$$

erhält man die Fibonacci-Folge

$$0, 1, 1, 2, 3, 5, 8, \ldots$$

Beschreibung von (4.16) als lineare Abbildung:

$$\underbrace{\begin{pmatrix} t_{m+1} \\ t_m \end{pmatrix}}_{x_m} = \begin{pmatrix} w_1 & w_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t_m \\ t_{m-1} \end{pmatrix} + \begin{pmatrix} w_3 \\ 0 \end{pmatrix} \tag{4.17}$$

Setze $x_0 = \binom{t_1}{t_0} = \binom{1}{0}$ und $x_{m+1} = Ax_m + b$, d.h., $x_{m+1} = \phi(x_m)$ für lineare Abbildung

$$\phi\colon \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Ax + b$$

Damit lässt sich x_m aus Anfangszustand x_0 berechnen:

$$x_m = \phi(x_{m-1}) = \phi^2(x_{m-2}) = \cdots = \phi^m(x_0) = A^m x_0 + (A^{m-1} + \cdots + A + E)b\,.$$

Beachte: $(A^{m-1} + \cdots + A + E)(A - E) = A^m - E$
Falls $(A - E)$ invertierbar:

$$x_m = A^m x_0 + (A^m - E)(A - E)^{-1} b$$

Falls A diagonalisierbar: $\exists S$ invertierbar mit $S^{-1}AS = D$ Diagonalmatrix, und

$$x_m = SD^m S^{-1} x_0 + (SD^m S^{-1} - E)(A - E)^{-1} b \tag{4.18}$$

Für $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ und $|\lambda_1|, \ldots, |\lambda_n| < 1$ konvergiert

$$D^m = \begin{pmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{pmatrix} \quad \text{gegen} \quad \mathbf{0} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

und die Folge $(x_m)_{m \in \mathbb{N}}$ konvergiert gegen

$$-E(A - E)^{-1} b = (E - A)^{-1} b \qquad \text{(“stabile Folge”)}$$

Für Fibonacci-Folge:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Charakteristisches Polynom:

$$X(X-1)-1 = X^2 - X - 1$$

Eigenwerte:

$$\lambda_1 = (1+\sqrt{5})/2 \approx 1,6180339887\ldots \qquad \text{mit Eigenvektor } \begin{pmatrix}\lambda_1\\1\end{pmatrix}$$

$$\lambda_2 = (1-\sqrt{5})/2 \approx -0,618\ldots \qquad \text{mit Eigenvektor } \begin{pmatrix}\lambda_2\\1\end{pmatrix}$$

$\lambda_1 > 1$ ("Goldener Schnitt"), unbegrenztes Wachstum: aus (4.18) folgt

$$\begin{aligned} x_m = \begin{pmatrix} t_{m+1} \\ t_m \end{pmatrix} &= \frac{1}{\lambda_1-\lambda_2} \underbrace{\begin{pmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{pmatrix}}_{S} \underbrace{\begin{pmatrix}\lambda_1^m & 0\\ 0 & \lambda_2^m\end{pmatrix}}_{D} \underbrace{\begin{pmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{pmatrix}}_{S^{-1}} \underbrace{\begin{pmatrix}1\\0\end{pmatrix}}_{x_0} \\ &= \frac{1}{\lambda_1-\lambda_2}\begin{pmatrix}\lambda_1^{m+1} & \lambda_2^{m+1}\\ \lambda_1^m & \lambda_2^m\end{pmatrix}\begin{pmatrix}1 & -\lambda_2\\ -1 & \lambda_1\end{pmatrix} \end{aligned}$$

Also

$$\underbrace{t_m}_{\text{ganze Zahl}} = \underbrace{(\lambda_1^m - \lambda_2^m)}_{\text{irrationale Zahlen!}} \quad /(\lambda_1-\lambda_2)$$

4.3.6 Trigonalisierbarkeit

Und wenn A nicht diagonalisierbar?

Definition 4.3.23. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *trigonalisierbar*, wenn sie zu einer (oberen) Dreiecksmatrix D ähnlich ist, d.h., $\exists S \in \mathrm{GL}(n,\mathbb{K})$:

$$S^{-1}AS = D = \begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$$

Für Trigonalisierbarkeit reicht ein Teil des Kriteriums für Diagonalisierbarkeit.

Satz 4.3.24. *Eine Matrix $A \in \mathbb{K}^{n\times n}$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom χ_A in Linearfaktoren zerfällt, d.h., $\exists \lambda_1,\ldots,\lambda_n \in \mathbb{K}$ (müssen nicht verschieden sein) so dass*

$$\chi_A(X) = (X-\lambda_1)\cdots(X-\lambda_n).$$

Bemerkung 4.3.25. Jede Matrix $A \in \mathbb{C}^{n\times n}$ ist trigonalisierbar, da jedes Polynom über dem Körper $\mathbb{C}$ der komplexen Zahlen in Linearfaktoren zerfällt (*Fundamentalsatz der Algebra*, kommt später im Studium).

Beweis von Satz 4.3.24. "$\Rightarrow$": Sei

$$S^{-1}AS = D = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Dann (Abschnitt 4.3.2)

$$\chi_A = \chi_D = (X - \lambda_1)\cdots(X - \lambda_n)$$

"$\Leftarrow$": per Induktion über n.
Zeigen Aussage für untere Dreiecksmatrix; ist äquivalent da $\chi_A = \chi_{A^\top}$.
Aussage sicher wahr für $n = 1$.
Sei u_{n+1} Eigenvektor von A zu Eigenwert λ_{n+1}. Existiert, da χ_A in Linearfaktoren zerfällt. Ergänzen u_{n+1} zu einer Basis $B = (u_1, u_2, \ldots, u_{n+1})$ von $\mathbb{K}^{n+1}$.
Sei R die Matrix mit den Spalten $u_1, u_2, \ldots, u_{n+1}$.
Dann ist $M = R^{-1}AR$ von der Gestalt

$$\begin{pmatrix} \tilde{M} & \mathbf{0} \\ * & \lambda_{n+1} \end{pmatrix}$$

Es gilt

$$\chi_A = \chi_M = (X - \lambda_1)\chi_{\tilde{M}}\,,$$

also zerfällt auch $\chi_{\tilde{M}}$ in Linearfaktoren.
Dann ist $\tilde{M}$ nach Induktionsannahme trigonalisierbar, d.h., es existiert eine invertierbare Matrix $\tilde{S}$ so dass $(\tilde{S})^{-1}\tilde{M}\tilde{S}$ eine untere Dreiecksmatrix. Definiere

$$S := \begin{pmatrix} \tilde{S} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \in \mathbb{K}^{n+1\times n+1}$$

Da $\det(S) = \det(\tilde{S}) \neq 0$ ist S invertierbar, und es gilt

$$S^{-1} = \begin{pmatrix} (\tilde{S})^{-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}$$

Dann ist

$$\begin{aligned} S^{-1}R^{-1}ARS &= S^{-1}\begin{pmatrix} \tilde{M} & \mathbf{0} \\ * & \lambda_{n+1} \end{pmatrix} S \\ &= \begin{pmatrix} (\tilde{S})^{-1}\tilde{M}\tilde{S} & \mathbf{0} \\ * & \lambda_{n+1} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ * & & & \lambda_n \end{pmatrix} \end{aligned}$$

untere Dreiecksmatrix. Mit $Q := RS$ ist also $Q^{-1}AQ$ untere Dreiecksmatrix, d.h., A ist trigonalisierbar. $\square$

Beispiel 4.3.26. Eine Matrix, die trigonalisierbar, aber nicht diagonalisierbar ist:

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

Eigenwert $\lambda_1 = 1$ mit algebraischer Vielfachheit 2, geometrische Vielfachheit ist

$$\begin{aligned} \dim(\mathrm{Kern}(A - \lambda_1 E)) &= \dim(\mathrm{Kern}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) \\ &= \dim(\{\begin{pmatrix} x \\ y \end{pmatrix} \mid y = 0\}) = 1. \end{aligned}$$

△

4.3.7 Anwendung: Stochastische Matrizen

Sei $A \in \mathbb{R}^{n\times n}$ und $s \in \mathbb{R}^n$. Wann existiert

$$\lim_{m\to\infty} A^m s\,?$$

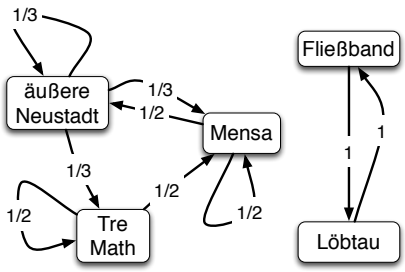
Spezialfall: sei s Eigenvektor von A zum Eigenwert 1, d.h.:

$$As = s \qquad\qquad \text{“stationäre Verteilung”} \quad s$$

Eine Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ heißt

- *zeilenstochastisch* falls $0 \le a_{ij} \le 1$ und Zeilensummen Eins betragen.
- *spaltenstochastisch*: analog.
- *doppelt stochastisch*: sowohl als auch.
- *stochastisch*: zeilen- oder spaltenstochastisch.

Beispiel 4.3.27. Betrachten das folgende Beispiel.



Beschreibung durch Matrix (‘Übergangsmatrix’):

$$\begin{pmatrix} 1/3 & 1/2 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/2 & 0 & 0 \\ 1/3 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{array}{c} \text{äußere Neustadt} \\ \text{Mensa} \\ \text{Tre Math} \\ \text{Fließband} \\ \text{Löbtau} \end{array}$$

△

Satz 4.3.28. *Sei $A \in \mathbb{R}^{n\times n}$ stochastisch, aperiodisch und irreduzibel, und $s \in \mathbb{R}^n$. Dann existiert $\lim_{m\mapsto\infty} A^m s$, ist unabhängig von s, und gleich dem Eigenvektor zum Eigenwert 1 von A.*

Können den Grenzwert also berechnen, indem wir ein lineares Gleichungssystem lösen!

Reverse Engineering: was könnte hier 'aperiodisch' heissen? Und was 'irreduzibel'?

Definition 4.3.29. Eine Matrix heißt *irreduzibel* wenn sie nicht geschrieben werden kann in der Form

$$\begin{pmatrix} M & 0 \\ P & N \end{pmatrix}$$

für quadratische Matrizen M und N.

$\rightsquigarrow$ Dämpfungsfaktor bei Google PageRank.

$\rightsquigarrow$ Weiterführende Frage: Wie schnell ist die Konvergenz?

Allgemeinerer Fall: A nicht mehr notwendigerweise stochastisch.

A heißt *positiv* falls für alle $i, j \in \{1, \ldots, n\}$ gilt $a_{ij} > 0$. Positive Vektoren: analog.

Satz 4.3.30 (Perron(-Frobenius), positiver Fall)**.** *Falls $A \in \mathbb{R}^{n \times n}$ positiv und irreduzibel, so so gibt es einen positiven (also insbesonderen reellen) Eigenwert λ der algebraischen Vielfachheit 1 so dass alle anderen Eigenwerte betragsmäßig strikt kleiner sind, und und einen positiven Eigenvektor zu λ.*

Anmerkungen.

- Wenn wir statt der Positivität von A nur fordern, dass A nicht-negativ ist, so kann es andere Eigenvektoren geben, die betragsmässig größtmöglich sind: zum Beispiel hat die Matrix
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
die Eigenwerte 1 und -1.
- Ausserdem muss der maximale Eigenwert nicht die algebraische Vielfachheit 1 haben, kann auch 0 sein, und der entsprechende Eigenwert muss nicht positiv sein: zum Beispiel hat
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
den einzigen Eigenwert 0 der Vielfachheit 2 mit zugehörigem Eigenvektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Was aber für den nicht-negativen Fall bleibt:

Satz 4.3.31 ((Perron-) Frobenius, nicht-negativer Fall)**.** *Falls $A \in \mathbb{R}^{n \times n}$ nicht-negativ und irreduzibel, so gibt es einen positiven (reellen) betragsmäßig größten Eigenwert λ mit nicht-negativen Eigenvektor. Die Anzahl der betragsmäßig größten Eigenwerte ist genau die Periodizität von A.*

Zu diesem Satz sind verschiedene Beweise bekannt, die allerdings über den Stoff der Vorlesung hinausgehen. Einer der Beweise verwendet den Fixpunktsatz von Brouwer.

Literaturverzeichnis

[1] M. Bodirsky. Introduction to mathematical logic, 2023. Course notes, TU Dresden, `https://wwwpub.zih.tu-dresden.de/~bodirsky/Logic.pdf`.

[2] G. Havas and C. Wagner. Matrix reduction algorithms for euclidean rings. In *Proc. 1998 Asian Symposium on Computer Mathematics*, pages 65–70. Lanzhou University Press, 1998.

[3] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[4] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley - Interscience Series in Discrete Mathematics and Optimization, 1998.

[5] J. von zur Gathen and D. Panario. Factoring polynomials over finite fields: A survey. *Journal of Symbolic Computation*, 31(1):3–17, 2001.