Lineare Algebra
Skript* zur Vorlesung LA10, LA20

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

21. Januar 2026

*Wintersemester 2015/2016; aufbauend auf einem handschriftlichen Skript von Reinhard Pöschel, mit Änderungen von Andreas Thom im Wintersemester 2016/2017. Vollständige Überarbeitung fürs Wintersemester 2023/2024. Ich freue mich über Emails mit Kommentaren und Verbesserungswünschen.

Inhaltsverzeichnis

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 (Empfehlung).

Zum Skript

- Blau markiert: Kommentare, Hervorhebungen, Literaturverweise.
- Rot markiert: nachträglich geändert oder hinzugefügt, bzw. Hyperlink.
- *Kursiv gedruckt:* Begriff wird definiert, oder sonst herausgehoben.
- Grün markiert (ab Kapitel 8): die komplexen/unitären Varianten der Aussagen.
- Das Symbol $\square$ markiert das Ende eines Beweises und das Symbol $\triangle$ markiert das Ende eines Beispiels.
- *Bemerkungen* dienen zum Vertiefen und vernetzen, und können gelegentlich übersprungen werden.
- Die *Übungen* ersetzen nicht die Übungen in den Tutorien, sondern sind zum Vertiefen und Vernetzen des Stoffes gedacht.

Literatur

- *Linear Algebra*, von Peter Petersen, Springer Verlag, 2012. Auf Englisch [9].
- *Lineare Algebra: Eine Einführung für Studienanfänger*, von Gerd Fischer, Springer Verlag, 18te Auflage, 2013 [5].
- *Lineare Algebra: Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen*, von Albrecht Beutelspacher, Vieweg+Teubner Verlag, 6te Auflage, 2003 [1].

Vorbemerkungen

- Großer Fortschritt in der Geometrie: Einführung von Koordinaten.
- Übersicht LA10 (Kapitel 1 bis Kapitel 4):
  - Grundlegendes zur Sprache der Mathematik
  - Gruppen und Körper
  - Vektorräume, Basen und Dimension,
  - Lineare Gleichungssysteme, Matrizen und lineare Abbildungen,
  - Polynome und Determinanten
  - Eigenwerte und Eigenvektoren, Diagonalisierbarkeit
- Übersicht LA20 (Kapitel 5 bis Kapitel 8):
  - Dualität
  - Analytische Geometrie
  - Normalformen (Reduzierte Stufennormalform, Frobenius-Normalform, . . .)
  - Euklidische und unitäre Vektorräume

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. [2]).

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 $A \cap B = \varnothing$.

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 $M_i := \{1, 2, \dots, n\}$; also $M_0 = \varnothing$, $M_1 = \{0\}$, $M_2 = \{0, 1\}$, $M_3 = \{0, 1, 2\}$, und so weiter. Dann ist

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

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* $(a_1, \dots, a_n)$. Das Element a_i, für $i \in \{1, \dots, n\}$, wird der *i-te Eintrag des Tupels* genannt.

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

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:

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

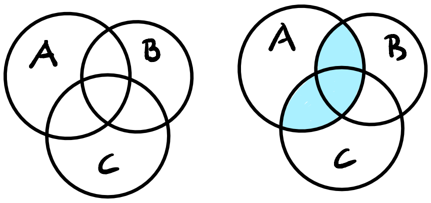
Beispiel: $\mathbb{R}^3$.

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 Venndiagrammen 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.

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

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 1.2.3. Für eine formale Definition des Begriffs *definierbar* verweisen wir auf die Vorlesung *Einführung in die mathematische Logik* [2].
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* (siehe Abschnitt 1.2.2), mit Kante zwischen a und b falls $(a, b) \in R$.
- Spezialfall $A = B$: man spricht von einer *Relation auf* A. Darstellung durch *gerichtete Graphen* (siehe Abschnitt 1.2.2): 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.

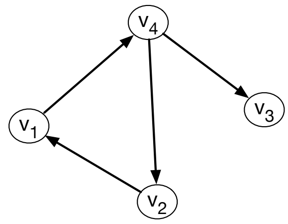


Abbildung 1.1: Eine Illustration des gerichteten Graphen G mit Knotenmenge $\{v_1, v_2, v_3, v_4\}$ und Kantenmenge $\{(v_1, v_4), (v_4, v_2), (v_2, v_1), (v_4, v_3)\})$.

1.2.2 Graphen

Ein *(ungerichteter) Graph* ist ein geordnetes Paar $G = (V, E)$ bestehend aus einer Menge V von *Knoten* und einer Menge E von zwei-elementigen Teilmengen von V, den *Kanten* des Graphen G.

Ein *gerichteter Graph* ist ein geordnetes Paar (V, A) bestehend aus einer Menge V von *Knoten* und einer zweistelligen Relation $A \subseteq V^2$ auf V, den *Kanten* des gerichteten Graphen. Graphen können wie in Abbildung 1.1 illustriert werden.

Jeder ungerichtete Graph $G = (V, E)$ kann als gerichteter Graph (V, A) aufgefasst werden: wir setzen $A = \{(a, b) \mid \{a, b\} \in E\}$. Die Kantenrelation A ist dann eine symmetrische Relation im Sinne von Abschnitt 1.2. Aus (V, A) gewinnen wir $G = (V, E)$ zurück durch $E = \{\{a, b\} \mid (a, b) \in A\}$.

1.2.3 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*, und a ein *Urbild* von $f(a)$ unter f. Weitere häufige Schreibweise: $x \mapsto f(x)$.

1.2.4 Spezielle Eigenschaften von Funktionen

- $f\colon A \to B$ heißt *surjektiv* falls für alle $b \in B$ ein $a \in A$ existiert mit $(a, b) \in G_f$. In anderen Worten, jedes $b \in B$ hat (mindestens) ein Urbild unter f.
- $f\colon A \to B$ 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, jedes $b \in B$ hat höchstens ein Urbild.
- $f\colon A \to B$ 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

$$\{(a, a') \in A^2 \mid f(a) = f(a')\}. \tag{1.1}$$

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. △

1.2.5 Komposition von Abbildungen

Seien $f\colon A \to B$ und $g\colon B \to C$ Funktionen. Dann bezeichnet $g \circ f$ die *Komposition* (Hintereinanderausführung) von f und g, nämlich die Abbildung von A nach C die definiert wird durch $(g \circ f)(x) := g(f(x))$ für alle $x \in A$.

1.2.6 Umkehrabbildung

Wenn $f\colon A \to B$ eine Funktion ist, dann definiert $(G_f)^{-1} := \{(b, a) \mid (a, b) \in G_f\}$ 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. Falls $f\colon A \to B$ bijektiv ist, dann gilt $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.

1.2.7 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 $\mathcal{P}(A)$. △

1.2.8 Gleichmächtige Mengen

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

Schreibweise:

- $|A| \leq |B|$ gdw es eine injektive Abbildung $f\colon A \to B$ gibt.
- $|A| < |B|$ falls $|A| \leq |B|$ 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 $u \notin U$.

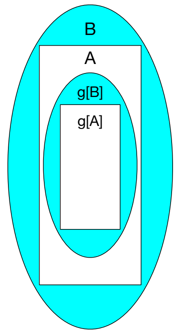
Wäre $u \in U$, so $u \in f(u)$, also $u \notin U$ nach Def. von U, Widerspruch.

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

Satz 1.2.10 (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.

[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.



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 $x = h(x)$. $\square$

1.2.9 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.

(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!

1.2.10 Die natürlichen Zahlen

Der Aufbau der natürlichen Zahlen als Mengen:

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

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

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

Die Relation $\leq$ auf $\mathbb{N}$ ist 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 1.2.12. 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: $n, m \in \mathbb{N}$

$$\begin{aligned} n \cdot 0 &:= 0 \\ n \cdot m^+ &:= n \cdot m + n \end{aligned}$$

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.

1.2.11 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. [2].

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:

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

Induktionsanfang $n = 1$.

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

Induktionsschritt: 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 $(\mathbb{R} \setminus \{0\}, \cdot)$. $\triangle$

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

$$\begin{aligned} +: & K \times K \to K && \text{Addition} \\ \cdot: & K \times K \to K && \text{Multiplikation} \end{aligned}$$

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 $x, y \in K$:

- $0 \cdot x = 0 = x \cdot 0$. Denn $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$, also $0 = 0 \cdot x$.
- $(-x) \cdot y = -(x \cdot y) = x \cdot (-y)$. Denn $0 = 0 \cdot y = (x + (-x))y = x \cdot y + (-x) \cdot y$, also $(-x) \cdot y = -(x \cdot y)$. Die Gleichung $-(x \cdot y) = x \cdot (-y)$ folgt analog.
- $x \cdot y = 0 \Leftrightarrow x = 0$ 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 1.2.11).

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

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

(der *Addition*) und

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

(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

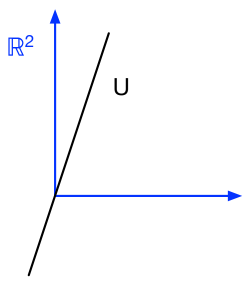
Sei V ein $\mathbb{K}$-Vektorraum.

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 $U \neq \emptyset$. 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 $U \neq \emptyset$ 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:

$$U_1, U_2 \leq V \;\Rightarrow\; U_1 \cap U_2 \leq V$$

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 $\langle \emptyset \rangle = \mathbf{0}$.

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 $X, Y \subseteq V$:

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

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 $U := \langle v_1, \ldots, v_n \rangle$.

1. $\mathbf{0} = 0 \cdot v_1 + \cdots + 0 \cdot v_n \in U$.

2. Seien $u, v \in U$, d.h., $u = \lambda_1 v_1 + \cdots + \lambda_n v_n$ und $v = \mu_1 v_1 + \cdots + \mu_n v_n$. Dann ist $u + v = (\lambda_1 + \mu_1) v_1 + \cdots (\lambda_n + \mu_n) v_n \in U$.

3. Seien $u \in U, \lambda \in K, u = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Dann ist $\lambda u = (\lambda\lambda_1) v_1 + \cdots + (\lambda\lambda_n) v_n \in U$.

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

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

2.4.2 Lineare Unabhängigkeit

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

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

Ansonsten: $(v_1, \ldots, v_n)$ *linear abhängig.* Eine Menge $U = \{v_1, \ldots, v_n\} \subseteq V$ heißt *linear unabhängig* wenn jedes n-Tupel $(v_1, \ldots, v_n)$ 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 $(v_1, \ldots, v_n)$ 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*

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

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

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

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 $u = \lambda_1 v_1 + \dots + \lambda_n v_n$. Wenn nun gilt

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

so folgt

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

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

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

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

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

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

- *(Ergänzung) 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).*

- *(Auswahl) 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. Ergänzung: '$\Rightarrow$' 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 $\lambda_1, \dots, \lambda_n \in K$ und $v_1, \dots, v_n \in B$ mit $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, 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 $\lambda_0, \lambda_1, \dots, \lambda_n \in K$ und $v_1, \dots, v_n \in B$ so dass $\lambda_0 v + \lambda_1 v_1 + \dots + \lambda_n b_n = \mathbf{0}$, aber $(\lambda_0, \lambda_1, \dots, \lambda_n) \neq (0, 0, \dots, 0)$. Falls $\lambda_0 = 0$ so sind $b_1, \dots, b_n$ linear abhängig, Widerspruch. Falls $\lambda_0 \neq 0$ so ist $v = \lambda_0^{-1}(-\lambda_1 b_1 - \dots - \lambda_n b_n) \in \langle B \rangle$ im Widerspruch zu $v \notin \langle B \rangle$. □

Übung 4. Beweisen Sie den zweiten Teil von 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 $V = \langle v_1, \dots, v_n \rangle$. Falls die Menge $M = \{v_1, \dots, v_n\}$ linear unabhängig ist, dann ist $\{v_1, \dots, v_n\}$ 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 (Abschnitt 5.1). □

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 $B \subseteq \langle B' \rangle$ und

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

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

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. $\square$

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)). $\triangle$

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) *$|C| \le |B|$, d.h., $m \le n$:*
'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 $|B| \ge 1 = m$. 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') $m - 1 \le n$, 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 $n \ge m$, fertig.
2. Fall: $n \le m - 1$ (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. $w_m \in \left\langle B' \right\rangle$.

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

Wäre $\lambda_i = 0$ für alle $i \ge m$, 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) $|B_1| \le |B_2|$ und analog $|B_2| \le |B_1|$.

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 $U \le V$ 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 $B = \{v_1, \ldots, v_n\}$ 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:

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

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:

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

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 $U_i := \langle u_i \rangle = K v_i$.

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 $\mathbb{K}$. Schreiben $\mathbb{K}^{m \times n}$ für die Menge aller $(m \times n)$-Matrizen.

Falls $B \in \mathbb{K}^{m \times n}$, so geben wir B häufig in folgender Schreibweise an:

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \ldots & \ldots & \ldots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

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;
- (gerichteten und ungerichteten) Graphen, und vielem anderen mehr.

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

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

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

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

Beschreibung linearer Gleichungssysteme

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

Übersichtlichere Schreibweise für das lineare Gleichungssystem:

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

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

Beschreibung von Graphen

Sei $G = (V, A)$ ein gerichteter Graph mit Knotenmenge $V = \{v_1, \dots, v_n\}$. Dann ist die *Adjazenzmatrix* von G die Matrix $B \in \{0,1\}^{n \times n}$ mit $b_{i,j} = 1$ falls $\{v_i, v_j\} \in E$ und $b_{i,j} = 0$ sonst. Der gerichtete Graph aus Abbildung 1.1 hat beispielsweise folgende Adjazenzmatrix:

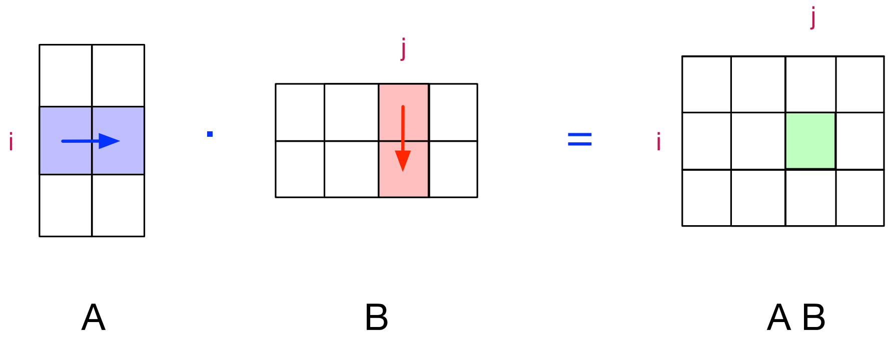
$$B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Die Adjazentmatrix eines ungerichteten Graphen ist definiert als die Adjazenzmatrix des zugehörigen symmetrischen Graphen (siehe Abschnitt 1.2.2).

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

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

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

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*
$$A(B + C) = AB + AC$$
$$(B + C)A = BA + CA$$

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

$$e_i = (0, \dots, 0, \underbrace{1}_{i}, 0, \dots, 0) \in \mathbb{K}^n$$

und wird *Einheitsvektor* genannt.

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

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

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \triangle$$

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 $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ 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, so dass $DA = AD$ für alle $A \in \mathbb{K}^{n\times n}$. Dann gilt $D = \lambda E_n$ für ein $\lambda \in \mathbb{K}$ (diese Matrizen werden auch *Skalarmatrizen* genannt).

Übung 11. Zeigen Sie: Falls $B \in \mathbb{K}^{n\times n}$ so, dass $BA = AB$ für alle $A \in \mathbb{K}^{n\times n}$, dann gilt $B = \lambda E_n$ für ein $\lambda \in K$. **Hinweis:** Betrachte die Matrix E_{ij}, die eine 1 an der Stelle i, j hat, und sonst Null ist. Was bedeutet $E_{ij}B = BE_{ij}$?

Übung 12. Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt genau dann $Ax = \mathbf{0}$ für alle $x \in \mathbb{R}^n$, wenn $A = \mathbf{0}$.

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

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

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 Darstellung: Zeilenrang (nämlich 2) sofort ablesbar. △

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

Beweis. $\langle z_1, \dots, z_n \rangle$ 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):

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

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 $a_{1j_1}, \ldots, a_{rj_r} \in K \setminus \{0\}$.

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 $i \in \{1, \ldots, r\}$, und $\operatorname{rg}(A) = r$.

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

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

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]

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

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.

[1] Sterne in Matrizen stehen für beliebige Einträge.

<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

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

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

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

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

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

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 $w_i = \varphi_B(u_i)$. 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

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

Insbesondere: $(b_1, \dots, b_s)$ ist genau dann Basis von $\langle u_1, \dots, u_m \rangle$, wenn $(\varphi(b_1), \dots, \varphi(b_s))$ 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 $\lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0}$, 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 $v_1, \ldots, v_n$ von A linear unabhängig, also eine Basis von $\mathbb{K}^n$, und $\langle v_1, \ldots, v_n\rangle = \mathbb{K}^n$. 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.

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

△

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 4.1.9)*

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

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

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 $s_1, \dots, s_n$ für die Spalten von A.

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

□

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

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 1.2.4):

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

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

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

Kern von A (ist der Kern von f_A, allerdings etwas anders definiert als in (1.1) in Abschnitt 1.2.4):

$$\operatorname{Kern} A := \{x \in \mathbb{K}^n \mid Ax = \mathbf{0}\} = \operatorname{Lös}(A, \mathbf{0}). \tag{3.8}$$

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

Satz 3.3.6. *Es gilt*

$$\begin{aligned} \dim(\operatorname{Kern} A) + \dim(\operatorname{Bild} A) &= n && \text{‘Dimensionsformel’} \\ \operatorname{rg}(A) &= n - d && \text{‘Rangformel’} \end{aligned}$$

wobei $d := \dim(\operatorname{Kern} A)$ *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 $(v_1, \dots, v_d)$ Basis von $\operatorname{Lös}(A, \mathbf{0})$, so ist

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

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

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* $\operatorname{rg}(A) = n$*.*

Beweis.

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

[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.

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

Bemerkung 3.3.9. Für $n = m$ und $\mathrm{rg}(A) = n$ 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 $\mathrm{rg}(A) \leq m < n$).

Bemerkung 3.3.11. Für $n < m$: alle 3 Fälle sind möglich (siehe Abschnitt 3.3.2). Wegen $\mathrm{rg}(A) \leq n < m$ (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., $\mathrm{Lös}(A, b) = \mathrm{Lös}(A', b')$, 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.

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

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

1. Fall: $b'_{r+1}, \dots, b'_n$ nicht alle 0. Dann ist $\mathrm{rg}(A|b) > r = \mathrm{rg}(A)$, 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. $\mathrm{Lös}(A, b) = \mathrm{Lös}(A', b')$.

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

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

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 $\mathrm{rg}(A) = 2 = \mathrm{rg}(A|b)$ 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:

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

Bemerkung 3.3.14. 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. $\dim \langle z_1, \dots, z_m \rangle$ berechnen (Abschnitt 3.2.5);
6. Basis von $\langle z_1, \dots, z_m \rangle$ 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

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

Im Beispiel (aus Abschnitt 3.3.4):

$$\begin{array}{cc} & \begin{array}{ccccc} x_1 & \lambda_1 & x_3 & \lambda_2 & \lambda_3 \end{array} \\ A' = & \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 0 \\ 0 & 0 & 1 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array}$$

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

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 $s_1, \dots, s_n$ Spalten von $A \in \mathbb{K}^{m \times n}$. Dann: $\operatorname{Bild} A = \langle s_1, \dots, s_n \rangle$ (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 3.3.15. 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 3.3.16.

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

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

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

Bemerkung 3.3.17. 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 3.3.18. 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 $A = \mathrm{rg}(A) = 2$ (Abschnitt 3.3.4) braucht man nur zwei linear unabhängige Spalten von A zu finden, z.B.

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

Übung 13. Sei $a \in \mathbb{K}^{n\times 1}$ (ein Spaltenvektor). Was ist die Dimension der Matrix $a^\top a$? Was ist die Dimension der Matrix $aa^\top$?

Eine Matrix heißt *symmetrisch*, wenn $A^\top = A$. Zum Beispiel sind Adjazentmatrizen von ungerichteten Graphen stets symmetrisch (siehe Abschnitt 3.2).

Übung 14. Zeigen Sie: für alle $A \in \mathbb{K}^{n\times n}$ ist $A^\top A$ symmetrisch.

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 3.3.19 (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 ('duale') System*

$$(A|b)^\top y = (\mathbf{0}|-1)^\top \tag{3.9}$$

lösbar ist.

Beweis. Seien $z_1, \ldots, z_m$ 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 $(\mathbf{0}|-1) \in \langle z_1, \ldots, z_m\rangle$. 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 3.3.19 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 $r := \mathrm{rg}(C) < \mathrm{rg}(C|d)$ 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 $(\mathbf{0}| - 1) \in \langle z_1, \ldots, z_m\rangle$. 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 $\mathrm{id}_V\colon V \to V$ ist linear (ein Endomorphismus).

2. Die Nullabbildung $V \to V\colon v \mapsto \mathbf{0}$ ist stets linear.

3. Allgemeiner: für $\lambda \in K$ ist $x \mapsto \lambda x$ linear.

4. Die Abbildung $f\colon \mathbb{C} \to \mathbb{C} : z \mapsto i \cdot z$ ist linear (betrachten $\mathbb{C}$ als $\mathbb{R}$-Vektorraum).

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

   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*

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

eine lineare Abbildung.

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

$$f_A\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \qquad \triangle$$

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

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

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

Übung 15. Es sei $\mathbb{K}$ ein Körper. Beweisen Sie, dass eine Abbildung $f\colon \mathbb{K} \to \mathbb{K}$ 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 $v_1, \ldots, v_n \in V$ und $f\colon V \to W$ eine lineare Abbildung. Dann gilt:

$$f[\langle\{v_1, \ldots, v_n\}\rangle] = \langle\{f(v_1), \ldots, f(v_n)\}\rangle \tag{3.10}$$

Denn:

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

Insbesondere:

$$f[V] \leq W \qquad \text{(Proposition 2.4.2)} \tag{3.11}$$

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

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 $v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in V$ (Satz 2.4.6). Wir definieren dann (wohldefiniert!)

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

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

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

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 1.2.3).
- $\operatorname{rg} f := \dim(\operatorname{Bild} f)$ (Definition sinnvoll wegen $\operatorname{Bild} f \leq W$, siehe (3.11))
- $\operatorname{dfkt} f := \dim(\operatorname{Kern} f)$ 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 $\iff$ $\operatorname{Kern} f = \{\mathbf{0}\}$;*
2. *Falls W endlich dimensional, so gilt*
$$f \text{ surjektiv} \iff \operatorname{Bild} f = W \iff \dim \operatorname{Bild} f = \dim W;$$
3. *Falls $\dim V = \dim W = n < \infty$, so gilt f injektiv $\iff$ f surjektiv $\iff$ f bijektiv (also Isomorphismus)*

Beweis. Zu 1.
$$\begin{aligned} f(v) = f(v') &\iff f(v) - f(v') = \mathbf{0} \\ &\iff f(v - v') = \mathbf{0} \\ &\iff v - v' \in \operatorname{Kern} f \end{aligned}$$

Zu 2. Siehe Abschnitt 2.4.5.
Zu 3. (Gilt nur in endlich dimensionalen Vektorräumen.)
$$\begin{aligned} f \text{ injektiv} &\iff \operatorname{Kern} f = \{\mathbf{0}\} \\ &\iff \dim \operatorname{Kern} f = 0 \\ &\iff \dim \operatorname{Bild} f = n - 0 = n && \text{(nach Dimensionsformel, Satz 3.3.6)} \\ &\iff \operatorname{Bild} f = W && \text{(siehe 2.)} \\ &\iff f \text{ surjektiv} && \text{(nach Definition).} \end{aligned}$$
□

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

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

Beweis. 1, '$\Rightarrow$': Sei f injektiv und $\lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = \mathbf{0}$. (Z.z.: $\lambda_1 = \dots = \lambda_n = 0$). Dann:
$$\begin{aligned} & \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = \mathbf{0} \\ \Rightarrow\ & f(\lambda_1 v_1 + \dots + \lambda_n v_n) = \mathbf{0} && \text{(Linearität von } f) \\ \Rightarrow\ & \lambda_1 v_1 + \dots + \lambda_n v_n \in \operatorname{Kern} f \\ \Rightarrow\ & \lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0} && \text{(Satz 3.4.6 (1.), da } f \text{ injektiv)} \\ \Rightarrow\ & \lambda_1 = \dots = \lambda_n = 0 && \text{(da } v_1, \dots, v_n \text{ linear unabhängig)} \end{aligned}$$

'$\Leftarrow$': 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 ist 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

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

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.

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

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

$$\begin{aligned} w \in v + U &\Leftrightarrow (w = v + u \text{ für ein } u \in U) \\ &\Leftrightarrow (v - w = u' \text{ für ein } u' \in U) \\ &\Leftrightarrow v - w \in U \\ &\Leftrightarrow v \in w + U \\ &\Leftrightarrow v + U = w + U \end{aligned} \tag{3.12}$$

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 16. 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 $p, q \in \mathbb{R}^2$.

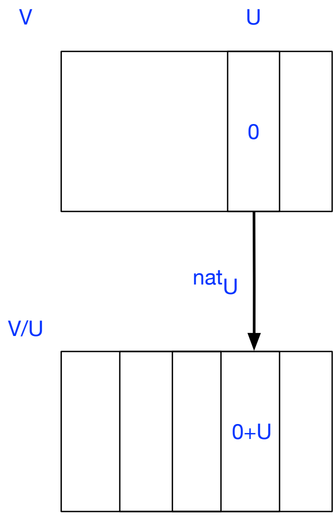
Satz 3.4.11. *Sei V ein $\mathbb{K}$-Vektorraum und $U \leq V$. Dann ist V/U zusammen mit den folgenden Operationen ein $\mathbb{K}$-Vektorraum (der* Faktorraum*):*

- *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 (3.12) 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 (3.12) 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 $U :=$ Kern f. Dann gilt*

$$\mathrm{Bild}\, f \simeq V/U$$

Ein Isomorphismus ist gegeben durch

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

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 $u \in U = \mathrm{Kern}\, f$ 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.

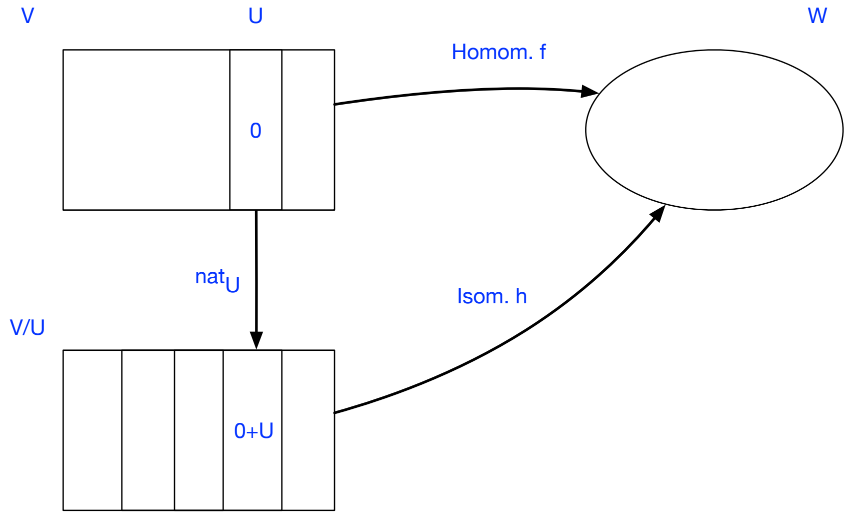
- h ist surjektiv: Nach Definition gilt für beliebiges $f(v) \in \mathrm{Bild}\, f$

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

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

$$\begin{aligned} h(v+U) = \mathbf{0} &\Longleftrightarrow f(v) = \mathbf{0} \\ &\Longleftrightarrow v \in \mathrm{Kern}\, f = U \\ &\Longleftrightarrow v + U = U = \mathbf{0}_{V/U} \end{aligned}$$ □

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:

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

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:

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

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 $U \le V$ 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

$$\dim(\mathrm{Bild}\, f) = \dim V - \dim(\mathrm{Kern}\, f) \tag{3.13}$$

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*

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

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

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

Auch gilt

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

Übung 17*.* Es sei $f\colon V \to W$ eine lineare Abbildung. Sei $U \le V$ 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:

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

Das heißt, f wird eindeutig festgelegt durch die Matrix

$$A := M_C^B(f) := \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \tag{3.14}$$

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:

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ \uparrow{\scriptstyle \varphi_B} & & \uparrow{\scriptstyle \varphi_C} \\ \mathbb{K}^n & \xrightarrow{A := M_C^B(f)} & \mathbb{K}^m \end{array}$$

Ein sogenanntes *kommutatives Diagram*. Konkret:

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

Denn:

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

Beispiel 3.4.17. $V := \mathbb{K}^n$, $W := \mathbb{K}^m$. Standardbasen $B_n := (e_1, \dots, e_n)$ von V und B_m von W. Der kanonische Basisisomorphismus $\varphi_{B_n} \colon \mathbb{K}^n \to \mathbb{K}^n : v \mapsto v$ 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

$$a_0 + a_1 X + a_2 X^2 + a_3 X^3$$

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:

$$\varphi_B \colon \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto a_0 v_0 + a_1 v_1 + a_2 v_2 + a_3 v_3 = a_0 + a_1 X + x_2 X^2 + a_3 X^3$$

Das heißt $\varphi_B(e_i) = v_{i-1}$ für $i \in \{1, \ldots, 4\}$. Das Differenzieren

$$\mathrm{diff} \colon \mathbb{R}^{(\leq 4)}[X] \to \mathbb{R}^{(\leq 4)}[X]$$

ist lineare Abbildung mit Matrix

$$A = M_B^B(\mathrm{diff}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{tikzcd} a_0 + a_1 X + a_2 X^2 + a_3 X^3 \arrow[r, "\mathrm{diff}"] & a_1 + 2a_2 X + 3a_3 X^3 = \varphi_B(Au) \arrow[d, "\varphi_B^{-1}"] \\ u = \begin{pmatrix} a_0 \\ \vdots \\ a_3 \end{pmatrix} \arrow[u, "\varphi_B"] \arrow[r, "f_A"] & Au = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{pmatrix} \end{tikzcd}$$

△

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.*

$$\begin{tikzcd} V_1 \arrow[r, "f_1"] & V_2 \arrow[r, "f_2"] & V_3 \\ \mathbb{K}^n \arrow[u, "\varphi_{B_1}"] \arrow[r, "f_{A_1}"] & \mathbb{K}^m \arrow[u, "\varphi_{B_2}"] \arrow[r, "f_{A_2}"] & \mathbb{K}^r \arrow[u, "\varphi_{B_3}"] \end{tikzcd}$$

Dann gilt

$$M^{B_1}_{B_3}(f_2 \circ f_1) = M^{B_2}_{B_3}(f_2) M^{B_1}_{B_2}(f_1)$$

(Funktionenkomposition) *(Matrizenmultiplikation)*

Folgerung für Inverse:

$$M^C_B(f^{-1}) = (M^B_C(f))^{-1}$$

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

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \varphi_B^{-1}(v)$$

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

$$x' = Tx \text{ und } x = Sx'$$

$$
\begin{tikzcd}
\mathbb{K}^n \arrow[dr, "\varphi_B"] \arrow[dd, bend left, "T"] & \\
& V \\
\mathbb{K}^n \arrow[uu, bend left, "S"] \arrow[ur, "\varphi_{B'}"'] &
\end{tikzcd}
$$

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, (3.14))

$$
\begin{aligned}
S &= \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = M_B^{B'}(\mathrm{id}_V) \\
T = S^{-1} &= \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} = M_{B'}^{B}(\mathrm{id}_V)
\end{aligned}
$$

Koordinatentransformation: Koordinatenvektor von $v \in V$

$$
\begin{gathered}
\text{bzgl. Basis } B = (e_1, e_2) \qquad \text{bzgl. Basis } B = (w_1, w_2) \\
x \mapsto Tx = S^{-1}x \\
e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\
e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}
\end{gathered}
$$

△

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.
$f: V \to W$ 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

$$M_{C_2}^{B_2}(f) = M_{C_2}^{C_1}(\mathrm{id}_w) \circ M_{C_1}^{B_1}(f) \circ M_{B_1}^{B_2}(\mathrm{id}_w) \tag{3.15}$$

$$A_2 = \tilde{S}^{-1} A_1 S \tag{3.16}$$

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).

$$\begin{tikzcd}
\mathbb{K}^n \arrow[dr, "\varphi_{B_1}"] \arrow[dd, bend left, "T"] & & & \mathbb{K}^m \arrow[dl, "\varphi_{C_1}"'] \arrow[dd, bend left, "\tilde{T}"] \\
& V \arrow[r, "f"] & W & \\
\mathbb{K}^n \arrow[ur, "\varphi_{B_2}"'] \arrow[uu, bend left, "S"] & & & \mathbb{K}^m \arrow[ul, "\varphi_{C_2}"] \arrow[uu, bend left, "\tilde{S}"]
\end{tikzcd}$$

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 18. 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

$$\begin{aligned} A_2 = \tilde{S}^{-1}A_1S &= M_{C_2}^{C_1}(\mathrm{id}_V)M_{C_1}^{B_1}(f_{A_1})M_{B_1}^{B_2}(\mathrm{id}_V) \\ &= M_{C_2}^{B_2}(f_{A_1}) \qquad \text{(siehe (3.15))} \end{aligned}$$

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

$$\begin{aligned} v \in U &\iff v \in \langle u_1, \dots, u_m \rangle \\ &\iff Tv \in \langle Tu_1, \dots, Tu_m \rangle = \langle e_1, \dots, e_m \rangle \\ &\iff (Tv)_{m+1} = \cdots = (Tv)_n = 0 \\ &\iff Rv = \mathbf{0} \\ &\iff v \in \text{Lös}(R, \mathbf{0}). \end{aligned}$$

□

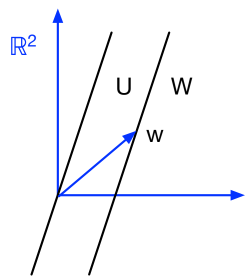
3.4.10 Gleichungssysteme und affine Unterräume

Idee: die Lösungsmenge eines allgemeinen linearen Gleichungssystems ist ein "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.

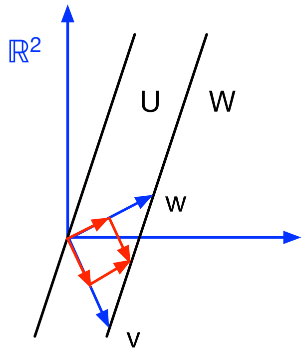


Abbildung 3.1: Illustration einer Affinkombination von w und v in $\mathbb{R}^2$.

(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 $\alpha_1 + \dots + \alpha_l = 1$. 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 $w + U$. $\square$

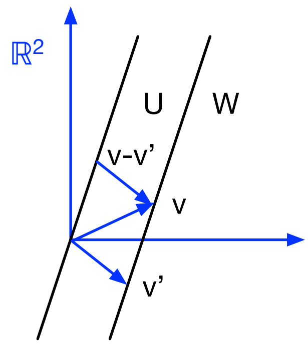


Abbildung 3.2: Illustration zum Beweis der Implikation (3) ⇒ (1).

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

$$(a_1 a_2) \circ (a_2 a_3) \circ \cdots \circ (a_{r-1} a_r)$$ □

Proposition 4.1.2. *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*). Proposition 4.1.2 zeigt, dass das Signum wohldefiniert ist.

Um Proposition 4.1.2 zu beweisen, benötigen wir folgende Definition.

Definition 4.1.3. Sei $\sigma \in S_n$. Ein *Fehlstand* von σ ist ein Paar (i, j) mit $1 \le i < j \le n$ und $\sigma(i) > \sigma(j)$. Wir schreiben $f(\sigma)$ für die Anzahl der Fehlstände von σ.

Proposition 4.1.2 folgt unmittelbar aus folgendem Lemma.

Lemma 4.1.4. *Es gilt* $\operatorname{sign}(\sigma) = (-1)^{f(\sigma)}$.

Beweis. Sei $\sigma = \pi \circ \tau$ für ein $\pi \in S_n$ und eine Transposition $\tau \in S_n$. Offenbar ist $f(\sigma) = f(\pi) + 1$ oder $f(\sigma) = f(\pi) - 1$. Da $f(\mathrm{id}) = 0$ folgt die Aussage nun aus Satz 4.1.1. □

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

$$S_n = A_n \cup \rho A_n \qquad \text{wobei } \tau A_n := \{\rho \circ \sigma \mid \sigma \in A_n\}.$$

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 4.1.5. 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 $z_1, \ldots, z_n, z_i' \in \mathbb{K}^n$ mit $i \in \{1, \ldots, n\}$ 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) $\det E_n = 1$.

Gibt es solche Funktionen?

Beispiel 4.1.6. $n = 2$.

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} := a_1 b_2 - a_2 b_1 \tag{4.1}$$

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

$$\begin{aligned} &(a_1 + b_1)(a_2 + b_2) - 2F_1 - 2F_2 - 2F_3 \\ &= a_1 a_2 + a_1 b_2 + b_1 a_2 + b_1 b_2 - 2 b_1 a_2 - a_1 a_2 - b_1 b_2 \\ &= a_1 b_2 - a_2 b_1. \end{aligned}$$

Vorzeichen ist negativ falls $180^\circ < \alpha = \measuredangle(u_1, u_2) < 360^\circ$. △

Bemerkung 4.1.7. Für $n = 3$ sind Determinanten als Volumina interpretierbar (siehe Abschnitt 6.4.2).

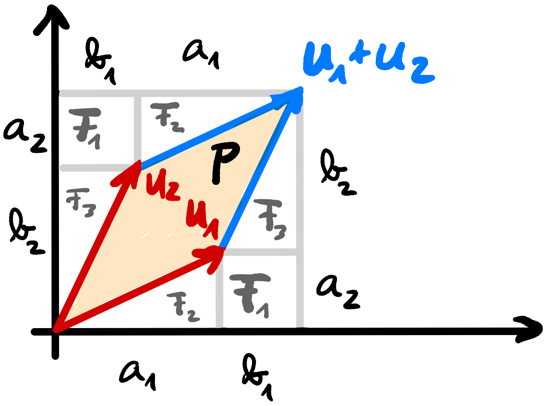


Abbildung 4.1: Die Determinante misst den vorzeichenbehafteten Flächeninhalt eine Parallelograms.

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 4.1.8. *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

$$\det A \overset{(3)}{=} \det B \overset{(*)}{=} -\det B' \overset{(3)}{=} -\det A'.$$ □

(D2) lässt sich verschärfen:

Satz 4.1.9. *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:

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_{n-1} \\ z_n \end{pmatrix} \overset{(D1)}{=} \sum_{i=1}^{n-1} \lambda_i \det \begin{pmatrix} z_1 \\ \vdots \\ z_{n-1} \\ z_i \end{pmatrix} \overset{(D2)}{=} 0$$

"$\Leftarrow$": 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 4.1.8, im Widerspruch zu (D3). Also $\det A \neq 0$. □

Rückführung auf E_n stets möglich bei $\mathrm{rg}(A) = n$, also ist Determinante eindeutig berechenbar. Wenn es sie überhaupt gibt!

Lemma 4.1.10. *Es gibt höchstens eine Determinantenfunktion* $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$.

Beweis. Seien $\det, \det'$ Determinantenfunktionen. Wegen Satz 4.1.9 gilt $\det A = \det' A = 0$ falls $\mathrm{rg}\, A < n$. Sei also $\mathrm{rg}\, A = n$. Nach Satz 3.2.27 erhalten wir E_n aus A durch elementare Zeilenumformungen. Dabei ändert sich die Determinante gemäß Proposition 4.1.8 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 4.1.11 (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*

$$\det(A) = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \tag{4.2}$$

Verallgemeinerung vom Fall $n = 2$ aus Abschnitt 4.1.2.

Beweis. Wegen Lemma 4.1.10 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

$$\begin{aligned} \det B &= \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) b_{1\sigma(1)} \cdots b_{i,\sigma(i)} \cdots b_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdots (a_{i,\sigma(i)} + a'_{i,\sigma(i)}) \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i,\sigma(i)} \cdots a_{n\sigma(n)} + \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdots a'_{i,\sigma(i)} \cdots a_{n\sigma(n)} \\ &= \det A + \det A' \end{aligned}$$

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:

$$a_{1j} = a_{2j} \qquad \text{für } j \in \{1, \dots, n\}$$

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 $n \in \{3, \dots, n\}$, und die vierte, da $a_{1j} = a_{2j}$ für alle $j \in \{1, \dots, n\}$, 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 19. Beweisen Sie, dass für quadratische Matrizen A und B gilt

$$\begin{vmatrix} A & 0 \\ * & B \end{vmatrix} = \det(A) \cdot \det(B).$$

Beispiel 4.1.12. 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):

$$\begin{array}{ccc} +a_{11}a_{22}a_{33} & +a_{12}a_{23}a_{31} & +a_{13}a_{21}a_{32} \\ \sigma = \mathrm{id} & \sigma = (123) & \sigma = (132) \end{array}$$

Ungerade Permutationen (Nebendiagonalen):

$$\begin{array}{ccc} -a_{31}a_{22}a_{13} & -a_{32}a_{23}a_{11} & -a_{33}a_{21}a_{12} \\ \sigma = (13) & \sigma = (23) & \sigma = (12) \end{array}$$

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 $a_{1\sigma(1)} \cdot \cdots \cdot a_{n\sigma(n)}$ 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 4.1.8 in Abschnitt 4.1.3.
- Determinante der Stufenform ist Produkt der Hauptdiagonalelemente (wie oben erklärt).

Beispiel 4.1.13.

$$\begin{aligned} \begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 2 \\ 0 & 1 & 0 \end{vmatrix} &= \begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 1 & 0 \end{vmatrix} && (z_2 - 2z_1 \rightsquigarrow z_2) \\ &= \begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{vmatrix} && (z_3 - z_2 \rightsquigarrow z_3) \\ &= 1 \cdot 1 \cdot 2 = 2 \end{aligned}$$

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 20. 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 4.1.14. *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 4.1.8).

Satz 4.1.15. *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 $\mathrm{rg}(B) < n$ dann $\mathrm{rg}(AB) < n$ (siehe Korollar 3.4.16), und mit Satz 4.1.9

$$\det(AB) = 0 = \det A \cdot \det B.$$

Also muss (1) nur für invertierbares B bewiesen werden ($\mathrm{rg}(B) = n$, $\det B \neq 0$). 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 $f(A) = 0$:

$$\begin{aligned} \mathrm{rg}(A) < 0 \Rightarrow \mathrm{rg}(AB) < n & \qquad \text{(Korollar 3.4.16)} \\ \Rightarrow \det AB = 0 & \qquad \text{wie oben für } B \text{ statt } A. \end{aligned}$$

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 4.1.10) 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.:

$$\begin{aligned} 1 = \det E_n &= \det(AA^{-1}) \\ &= \det A \cdot \det(A^{-1}) \qquad \text{nach Teil 1.} \end{aligned}$$

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 4.1.16. 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.

$$A = \begin{pmatrix} & & a_{1j} & & \\ & B_1 & \vdots & B_2 & \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ & B_3 & \vdots & B_4 & \\ & & a_{nj} & & \end{pmatrix}, \quad A_{ij} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$$

Satz 4.1.17 (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 4.1.14: $\det A^\top = \det A$. □

Beispiel 4.1.18. $n = 3$:

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & 4 & 0 \\ 3 & 0 & 0 \end{vmatrix} = \quad ? \qquad \text{Vorzeichen } (-1)^{i+j} : \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

Entwicklung nach 1. Zeile $(*)_1$:

$$+ 1 \cdot \begin{vmatrix} 4 & 0 \\ 0 & 0 \end{vmatrix} - 2 \cdot \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 4 \\ 3 & 0 \end{vmatrix}$$
$$= 1 \cdot 0 - 2 \cdot 0 + 1 \cdot (1 \cdot 0 - 3 \cdot 4) = -12$$

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 21. 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 $f: V \to V$ eine lineare Abbildung, und $A := M_B^B(f)$ die zugehörige Matrix (bezüglich einer Basis B von V; siehe Abschnitt 3.4.5).

Definition 4.1.19. $\det f := \det A$.

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 $A_2 = S^{-1} A_1 S$. Also

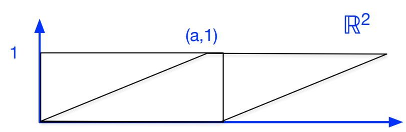
$$\begin{aligned} \det A_2 &= \det(S^{-1} A_1 S) \\ &= (\det S)^{-1} \det A_1 (\det S) && \text{nach Satz 4.1.15} \\ &= \det A_1 && \text{da Multiplikation in } \mathbb{K} \text{ kommutativ.} \end{aligned}$$

Bemerkung 4.1.20. Wir halten fest: ähnliche Matrizen haben die gleiche Determinante.

Bemerkung 4.1.21. $\det f$ kann als Verzerrungsfaktor für Flächen (in $\mathbb{R}^2$) bzw. Volumina (in $\mathbb{R}^n$) der Abbildung f interpretiert werden (siehe auch späteren Abschnitt 6.4.2):

$$F' := (\det f) \cdot F$$

Beispiel 4.1.22. $V = \mathbb{R}^2$, $A := \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ 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 $n = m$:

$$\begin{aligned} & Ax = b \text{ eindeutig lösbar} & \\ \iff & \operatorname{rg}(A) = n & \text{nach Korollar 3.3.7} \\ \iff & \det(A) \neq 0 & \text{nach Satz 4.1.9} \end{aligned}$$

'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 4.1.23 (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 4.1.24. Das lineare Gleichungssystem

$$\begin{aligned} 2x_1 + 3x_2 &= 1 \\ x_1 - 4x_2 &= 6 \end{aligned}$$

hat die eindeutige Lösung

$$x_1 = \frac{\begin{vmatrix} 1 & 3 \\ 6 & -4 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -4 \end{vmatrix}} = \frac{-22}{-11} = 2$$

$$x_2 = \frac{\begin{vmatrix} 2 & 1 \\ 1 & 6 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -4 \end{vmatrix}} = \frac{11}{-11} = -1$$

△

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 4.1.9 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 $x_i = \frac{|A_i(b)|}{|A|}$. $\square$

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 die folgenden Beispiele.

Beispiel 4.1.25. Betrachte die Matrix

$$\begin{pmatrix} 2 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 1 & 2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 1 & \cdots & 1 & 2 \end{pmatrix}$$

Diese läßt sich mit dem Verfahren aus Abschnitt 3.2.4 wie folgt in Diagonalform bringen: mit den elementaren Zeilenumformungen $2z_i - z_1 \rightsquigarrow z_i$ für $i \in \{2, \ldots, n\}$ erhalten wir

$$\begin{pmatrix} 2 & 0 & 0 & \cdots & 0 \\ 0 & 4 & 0 & \cdots & 0 \\ 0 & 2 & 4 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 2 & \cdots & 2 & 4 \end{pmatrix}$$

Wir fahren fort mit $2z_i - z_2 \rightsquigarrow z_i$ für $i \in \{3, \ldots, n\}$ und erhalten

$$\begin{pmatrix} 2 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 4 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 8 & 0 & & \vdots \\ 0 & 0 & 4 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 8 & 0 \\ 0 & 0 & 4 & \cdots & 4 & 8 \end{pmatrix}$$

und so weiter, bis wir schließlich mit der Umformung $2z_n - z_{n-1} \rightsquigarrow z_n$ die Diagonalmatrix

$$\begin{pmatrix} 2 & 0 & 0 & \cdots & 0 \\ 0 & 2^2 & 0 & \cdots & 0 \\ 0 & 0 & 2^3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 2^n \end{pmatrix}$$

Die Zahl 2^n ist zwar groß, aber noch nicht extrem groß: sie kann noch ohne Probleme abgespeichert und manipuliert werden. Es können aber auch noch sehr viel größere Zahlen auftreten; betrachte dazu das nächste Beispiel. △

Beispiel 4.1.26. 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, den Betrag eines Eintrags der ersten 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'* [6]).

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. $\triangle$

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 4.1.28).

Die Umformung in Stufenform aus Beispiel 4.1.26 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 4.1.27. 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 4.1.28. *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

$$|\det A| \le \prod_{i,j=1}^{n} (|p_{ij}| + 1) .$$

Damit haben wir

$$|p| = |\det A| \cdot q \le \prod_{i,j=1}^{n} (|p_{ij}| + 1) q_{ij} = 2^{\prod_{i,j=1}^{n} \log_2(|p_{ij}|+1) + \log_2 q_{ij}} < 2^{\mathrm{Groe}(A)-1} . \tag{4.6}$$

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 4.1.29. *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 4.1.28 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 $M \in \mathbb{K}^{m \times n}$, $i_1, \dots, i_k \in \{1, \dots, m\}$, und $j_1, \dots, j_l \in \{1, \dots, n\}$, dann schreiben wir $M^{i_1,\dots,i_k}_{j_1,\dots,j_k}$ für die Untermatrix von M, die aus M durch Löschen aller Zeilen ausser $i_1, \dots, i_k$ und aller Spalten ausser $j_1, \dots, j_l$ 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

$$d_{ij} = \frac{\det \begin{pmatrix} B_k & * \\ \mathbf{0} & d_{ij} \end{pmatrix}}{\det(B_k)} = \frac{\det\left((A_k)^{1,\dots,k,k+i}_{1,\dots,k,k+j}\right)}{\det\left((A_k)^{1,\dots,k}_{1,\dots,k}\right)}. \tag{4.7}$$

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 $\mathrm{Groe}(d_{ij}) \leq 4\,\mathrm{Groe}(M)$ nach Lemma 4.1.28. Auf ähnliche Weise läßt sich auch für den zweiten Teil des gaußschen Algorithmus nachweisen, dass die auftretenden Zahlen nicht zu groß werden [10].

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 4.1.16) und die Schreibweise $A_i(b)$ aus Abschnitt 4.1.7. Zunächst eine Hilfsaussage, die direkt aus dem Entwicklungssatz (Satz 4.1.17) folgt.

Lemma 4.1.30. *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 4.1.31. 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 $(a^{\#}_{ij})_{i,j\in\{1,\dots,n\}}$ heiss die *Adjunkte* oder *Komplementärmatrix* von A, und wird mit $A^{\#}$ bezeichnet.

Beispiel 4.1.32. Falls $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$, dann ist $a^{\#}_{21} = (-1)^3\det(A_{12}) = -1$. △

Satz 4.1.33. *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 $i, j \in \{1, \ldots, n\}$, dass

$$
\begin{aligned}
(A^{\#}A)_{ij} &= \sum_{k \in \{1,\ldots,n\}} (A^{\#})_{ik} \cdot a_{kj} \\
&= \sum_{k \in \{1,\ldots,n\}} (-1)^{i+k} \det(A_{ki}) \cdot a_{kj} \\
&= \sum_{k \in \{1,\ldots,n\}} a_{kj} \cdot \det(A_i(e_k)) && \text{(Lemma 4.1.30)} \\
&= \det(a_{1*}, \ldots, a_{(i-1)*}, a_{j*}, a_{(i+1)*}, \ldots, a_{n*}) && \text{(Linearität in der } j\text{-ten Spalte)} \\
&= \delta_{ij} \cdot \det(A) && \text{(det ist alternierend).}
\end{aligned}
$$

□

Beispiel 4.1.34. 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?

$$x^2 + 2x + 1$$

Ausführlich behandelt in: Vorlesung *Algebra (AL10)*.
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 · (‘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 (siehe 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.

Definition 4.2.2. 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 4.2.3. $(\mathbb{Z}; +, \cdot)$: der Ring der ganzen Zahlen (kommutativ, mit Eins; die einzigen Einheiten sind 1 und -1). △

Beispiel 4.2.4. $(\mathbb{Z}_n, +, \cdot)$: der *Restklassenring* (siehe Abschnitt 1.2.11; kommutativ, mit Eins; Körper falls n prim). △

Beispiel 4.2.5. $(\mathbb{K}^{n \times n}, +, \cdot)$: der *Matrizenring* über $\mathbb{K}$ (nicht kommutativ, siehe Beispiel 3.2.6; aber mit Eins E_n). △

Beispiel 4.2.6. Sei V ein Vektorraum. Dann ist

$$\mathrm{End}(V) := \{f : V \to V \mid f \text{ lineare Abbildung}\}$$

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 4.2.7. 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 $\varphi\colon \mathbb{N} \to R$ heißt auch *Folge*. Schreibweise: $\varphi = (a_i)_{i\in\mathbb{N}} = (a_0, a_1, \dots)$ mit $a_i := \varphi(i) \in \mathbb{K}$. 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 4.2.8. Falls R sogar ein Körper $\mathbb{K}$ ist, dann wird $\mathcal{F}$ zu einem $\mathbb{K}$-Vektorraum. Eine (unendliche!) Basis ist

$$\begin{gathered}(1,0,0,\dots),\\(0,1,0,\dots),\\(0,0,1,\dots),\\\dots\end{gathered}$$

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):

$$\begin{array}{lcr}
\text{alt} & & \text{neu}\\
(1,0,0,\dots) & =: & X^0\\
(0,1,0,\dots) & =: & X^1\\
\vdots & \vdots & \vdots\\
(0,0,\dots,0,1,0,\dots) & =: & X^n\\
\mathcal{F} & =: & R[X]
\end{array}$$

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 4.2.9. 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

$$(a_0, a_1, \dots, a_n, 0, 0, \dots) = a_0 + a_1X + a_2X^2 + \dots + a_nX^n$$

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 $\varphi = a_0 + a_1 X + \cdots + a_n X^n$ und $\psi = b_0 + b_1 X + \cdots + b_m X^m$ gelte
$$\varphi \cdot \psi = c_0 + c_1 X + \cdots + c_{n+m} X^{n+m}$$
mit $c_k = \sum_{i=0}^{k} a_k b_{k-i}$ für $k \in \mathbb{N}$.

Satz 4.2.10. $(R[X], +, \cdot)$ *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
$$\varphi = a_0 + a_1 X + a_2 X^2 + \cdots \in R[X]$$
Definieren den *Grad* des Polynoms φ wie folgt:
$$\begin{aligned} \operatorname{grad}(\underline{0}) &:= -\infty \\ \operatorname{grad}(\varphi) &:= \max(i \in \mathbb{N} \mid a_i \neq 0\} \end{aligned}$$
Dann gilt:
$$\operatorname{grad}(\varphi + \psi) \leq \max(\operatorname{grad}(\varphi), \operatorname{grad}(\psi))$$
$$\operatorname{grad}(\varphi \cdot \psi) \leq \operatorname{grad}(\varphi) + \operatorname{grad}(\psi) \tag{4.8}$$
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) Gleichheit anstatt $\leq$.

4.2.5 Polynomfunktionen

Nun der bereits angekündigte wichtige Übergang von der Syntax zur Semantik.

Sei $\varphi \in R[X]$ ein Polynom,
$$\varphi = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n .$$
Sei S ein Ring mit $R \subseteq S$ (z.B. $S = R[X]$, siehe Bemerkung 4.2.9) und $s \in S$. Dann ist
$$\varphi^S(s) := a_0 + a_1 s + a_2 s^2 + \cdots + a_n s^n$$

ein Element von S!
Auswertung von φ in S an der Stelle s. "Einsetzen" von s in φ.

Die Abbildung

$$\varphi^S : S \to S : r \mapsto \varphi^S(s)$$

heißt die zu φ gehörige *Polynomfunktion*. In der Algebra allgemeiner: 'Termfunktion'.

Wichtig: Unterschied

Polynom	Polynomfunktion
(Syntax)	(Semantik)
φ	φ^S

Definition 4.2.11. Sei $\varphi \in R[X]$ ein Polynom. Ein Element $s \in S$ heißt *Nullstelle von φ in S* falls $\varphi^S(s) = 0$. Für $a \in R$ heißen die Nullstellen des Polynoms $X^n - a$ *(n-te) Wurzeln von a.*

Hier steht 0 für das Nullelement des Ringes S = Nullelement von R.

4.2.6 Der Auswertungshomomorphismus

Satz 4.2.12 (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 $\varphi, \psi \in R[X]$:*

$$(\varphi + \psi)^S(s) = \varphi^S(s) + \psi^S(s)$$
$$(\varphi \cdot \psi)^S(s) = \varphi^R(s) \cdot \psi^S(s)$$

Die Voraussetzungen sind z.B. gegeben, wenn S ebenfalls ein kommutativer Ring ist. Algebraischer Hintergrund: die Abbildung

$$h_s : R[X] \to S : \varphi \mapsto \varphi^S(s)$$

ist ein (Ring-) Homomorphismus.

Beispiel 4.2.13. 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 4.2.12 anwendbar. Seien $\varphi_1 = (-2+X)$ und $\varphi_2 = (3+X)$. Dann ist

$$\begin{aligned}\varphi := \varphi_1\varphi_2 &= (-2+X)(3+X)\\ &= -6 + X + X^2\end{aligned}$$

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:

$$\begin{aligned}
\varphi_1^S(A) &= \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -5 \end{pmatrix}\\
\varphi_2^S(A) &= \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 0 & 0 \end{pmatrix}\\
\varphi^S(A) &= \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix}^2 = \cdots = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Satz 4.2.12 (Auswertungssatz) ergibt

$$\varphi(A) = \varphi_1(A)\cdot\varphi_2(A) = \begin{pmatrix} 0 & 1 \\ 0 & -5 \end{pmatrix}\cdot\begin{pmatrix} 5 & 1 \\ 0 & 0 \end{pmatrix} = 0$$

Nullstellen von φ in $\mathbb{K}^{2\times 2}$: z.B. die Matrix A. △

Beispiel 4.2.14. Sei V ein $\mathbb{K}$-Vektorraum. Dann ist $\operatorname{End}(V)$ ein (nicht kommutativer) Ring (Beispiel 4.2.6). Wir werden im Folgenden annehmen, dass $\mathbb{K} \subseteq \operatorname{End}(V)$ ist: das Element $\lambda \in \mathbb{K}$ fassen wir auf als $v \mapsto \lambda\,\mathrm{id}_V$. Also können wir Polynome $\varphi \in \mathbb{K}[X]$ auswerten in $\operatorname{End}(V)$. △

4.2.7 Polynomdivision

Teilbarkeitslehre für Polynome ähnlich wie für Zahlen (Vorlesung Algebra).

Definition 4.2.15. Seien $\varphi, \psi \in \mathbb{K}[X]$. Dann heißt φ *ein Vielfaches von* ψ, und ψ *ein Teiler* von φ (Schreibweise: $\psi|\varphi$), falls es ein $\varphi_1 \in \mathbb{K}[X]$ gibt mit $\varphi = \varphi_1\psi$.

Polynomdivision: Division mit Rest.

Beispiel:

$$\begin{array}{lllllll}
(& X^5 & +3X^4 & +0\cdot X^3 & +X^2 & +6X & -6) : \ (X^2+X-1) = X^3+2X^2-X+4 \\
-(& X^5 & +X^4 & -X^3) & & & \\
 & 0 & +2X^4 & -X^3 & & & \\
 & -(& 2X^4 & +2X^3 & -2X^2) & & \\
 & & 0 & -X^3 & 3X^2 & & \\
 & & -(& -X^3 & -X^2 & X) & \\
 & & & 0 & 4X^2 & +5X & -6 \\
 & & & -(& 4X^2 & +4X & -4 \\
 & & & & 0 & X & -2 \qquad \text{'Rest' } \rho = X-2
\end{array}$$

Also: $\frac{\varphi}{\psi} = \varphi_1 + \frac{\rho}{\psi}$, d.h.,

$$\varphi_1\psi + \rho$$

wobei $\mathrm{grad}(\rho) < \mathrm{grad}(\psi)$.

Lemma 4.2.16. *Sei $\varphi \in \mathbb{K}[X]$ und $k \in \mathbb{K}$. Dann ist k genau dann Nullstelle von φ, wenn $(X-k)|\varphi$.*

Beweis. Sei $\psi := (X-k)$.

'$\Leftarrow$': $\varphi = \varphi_1 \cdot \psi$ ergibt mit Satz 4.2.12

$$\varphi(k) = \varphi_1(k)\psi(k) = \varphi_1(k) \cdot 0 = 0$$

'$\Rightarrow$': Polynomdivision liefert $\varphi = \varphi_1 \cdot (X-k) + \rho$ wobei $\mathrm{grad}(\rho) < \mathrm{grad}(\psi) = 1$. Wegen $\varphi(k) = 0$ folgt $\rho(k) = 0$. Da $\mathrm{grad}(\rho) = 0$ ist $\rho \in \mathbb{K}$. Also $\rho = 0$, und damit $\psi|\varphi$. $\square$

Definition 4.2.17. Die *algebraische Vielfachheit* einer Nullstelle k ist definiert als

$$\max\{m \in \mathbb{N} : (X-k)^m|\varphi\}.$$

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 $\varphi(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ ist definiert als das Polynom

$$\varphi'(X) := a_1 + 2a_2X + \cdots + na_nX^{n-1}.$$

(Siehe auch Beispiel 3.4.19.)

Lemma 4.2.18. *Ein Polynom $\varphi \in \mathbb{K}[X]$ 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 $\varphi(X) = (1-X)^m\psi(X)$ mit $m \geq 2$ (Lemma 4.2.16). Also ist

$$\varphi'(X) = m(X-\lambda)^{m-1}\psi(X) + (X-\lambda)^m\psi'(X)$$

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 $\varphi(X) = (X-\lambda)\psi(X)$, und $\varphi'(X) = \psi(X) + (X-\lambda)\psi'(X)$. Also $0 = \varphi'(X)(\lambda) = \psi(\lambda) + (\lambda-\lambda)\psi'(\lambda) = \psi(\lambda)$ und damit ist λ Nullstelle von ψ. Also ist λ mehrfache Nullstelle von φ. $\square$

4.3 Eigenwerte, Eigenvektoren, Diagonalisierbarkeit

Viele Anwendungen in der Physik, Stochastik, diskreten Mathematik, ...

Definition 4.3.1. Sei V ein $\mathbb{K}$-Vektorraum, und $f\colon V \to V$ 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 (im Englischen *eigenspace*).

$\operatorname{Eig}_\lambda(f)$ 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 $B = (v_1, \dots, v_n)$, $f\colon V \to V$ 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 $\mathrm{Eig}_\lambda(f) \simeq \mathrm{Eig}_\lambda(A)$. Genauer: sei $\varphi_B\colon \mathbb{K}^n \to V$ der kanonische Basisisomorphismus (Abschnitt 2.4.3). Dann gilt

$$\begin{aligned} Ax = \lambda x &\iff \varphi_B(Ax) = \varphi_B(\lambda x) \\ &\iff f(\varphi_B(x)) = \lambda \varphi_B(x) \end{aligned}$$

Ist x der Koordinatenvektor von v bzgl. Basis B (das heißt, $\varphi_B(x) = v$) dann gilt:

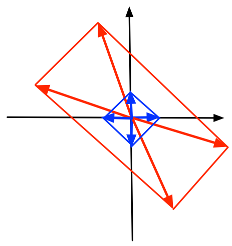
$$\begin{aligned} &v \text{ ist Eigenvektor von } f \text{ zu EW } \lambda \\ \iff &x \text{ ist Eigenvektor von } A \text{ zu EW } \lambda \end{aligned} \tag{4.11}$$

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?

$$\begin{aligned} e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\mapsto Ae_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \\ e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} &\mapsto Ae_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \end{aligned}$$



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 $0 \leq w(1), \dots, w(n) \in \mathbb{R}$ (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

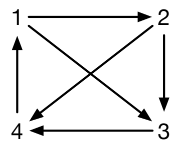
$$\begin{pmatrix} w(1) \\ \vdots \\ w(n) \end{pmatrix} =: x = \tilde{\lambda} A x \qquad \text{für } A = (a_{ji})_{i,j=1,\dots,n}$$

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. $f : V \to V$ 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

$$\begin{aligned} \det(XE - B) &= \det(XS^{-1}S - S^{-1}AS) \\ &= \det(S^{-1}(XE - A)S) \\ &= \det S^{-1} \cdot \det(XE - A) \cdot \det S \\ &= \det(XE - A) . \end{aligned}$$ □

[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.

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

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

$$\begin{aligned} \chi_A(\lambda) = \det(XE - A) &= \begin{vmatrix} X & 1 & -1 \\ 3 & X+2 & -3 \\ 2 & 2 & X-3 \end{vmatrix} && \text{(Definition)} \\ &= X(X+2)(X-3) - 6 - 6 + 2(X+2) + 6X - (X-3)3 && \text{(Sarrus, Beispiel 4.1.12)} \\ &= X^3 - 3X^2 + 2X^2 - 6X - 12 + 2X + 4 + 6X - 3X + 9 && \text{(Ausmultiplizieren)} \\ &= X^3 - X^2 - X + 1 && \text{(Vereinfachen)} \\ &= (X-1)^2(X+1) && \text{(Faktorisieren)} \end{aligned}$$

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. $a_{n-1} = (-1)^{n-1}(a_{11} + \cdots + a_{nn}) =: (-1)^{n-1}\,\mathrm{Spur}(A)$ 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 $(-1)^{n-1}\,\mathrm{Spur}(A)$ für X^{n-1}.

Übung 22. Beweisen Sie: für $A, B \in \mathbb{K}^{n\times n}$ gilt $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$.

Übung 23. Seien $A_1, \ldots, A_n \in \mathbb{K}^{n\times n}$ und $\pi \in S_n$ eine Permutation. Gilt $\mathrm{Spur}(A_1, \ldots, A_n) = \mathrm{Spur}(A_{\pi(1)}\cdots A_{\pi(n)})$?

Übung 24. 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 $\leftrightsquigarrow$ Linearfaktoren, Lemma 4.2.16)
Sätze und Algorithmen zur Faktorisierung univariater Polynome $\varphi \in \mathbb{K}[X]$:

- 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 [12].

- über $\mathbb{K} = \mathbb{Q}$: Lenstra-Lenstra-Lovász Algorithmus [8].

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 $B = (v_1, \dots, v_n)$, und $f: V \to V$ 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 $i \in \{1, \dots, n\}$.*

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 $v_i = \varphi_B(e_i)$ 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

$$\alpha_1 v_1 + \dots + \alpha_{k+1} v_{k+1} = \mathbf{0} \tag{4.12}$$

Dann gilt:

$$\alpha_1 \lambda_{k+1} v_1 + \dots + \alpha_k \lambda_{k+1} v_k + \alpha_{k+1} \lambda_{k+1} v_{k+1} = \mathbf{0} \quad (\lambda_{k+1} \cdot (4.12)) \tag{4.13}$$

$$\underbrace{\alpha_1 \underbrace{\lambda_1 v_1}_{=f(v_1)}}_{=f(\alpha_1 v_1)} + \dots + \alpha_k \lambda_k v_k + \alpha_{k+1} \lambda_{k+1} v_{k+1} = \mathbf{0} \quad (\text{Anwenden von } f \text{ auf } (4.12)) \tag{4.14}$$

$$\alpha_1(\lambda_1 - \lambda_{k+1}) v_1 + \dots + \alpha_k(\lambda_k - \lambda_{k+1}) v_k = \mathbf{0} \quad (\text{Subtraktion } (4.14) \text{ - } (4.13))$$

Nach Induktionsvoraussetzung sind $v_1, \dots, v_k$ linear unabhängig, also

$$\alpha_1(\lambda_1 - \lambda_{k+1}) = \dots = \alpha_k(\lambda_k - \lambda_{k+1}) = 0\,.$$

Wegen $\lambda_i \neq \lambda_{k+1}$ ist $\lambda_i - \lambda_{k+1} \neq 0$, für alle $i \in \{1, \dots, k\}$, und daher

$$\alpha_1 = \dots = \alpha_k = 0$$

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, $f \in \operatorname{End}(V)$, 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 $A' := S^{-1}AS$ 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,*
- $f \in \operatorname{End}(V)$,
- *$A = M_B^B(f)$ 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, $n_i = \dim \operatorname{Kern}(A - \lambda_i E)$,*
- *$(v_1^{(i)}, \dots, v_{n_i}^{(i)})$ sei Basis des Eigenraums $\operatorname{Eig}_{\lambda_i}(f) = \operatorname{Kern}(f - \lambda_i \operatorname{id})$,*
- *$m_1, \dots, m_r$ 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. *$n_i \leq m_i$ und $\sum_{i=1}^r n_i \leq \sum_{i=1}^r m_i \leq n$.*
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 $i \in \{1, \ldots, r\}$.

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.:

$$\underbrace{\alpha_1^{(1)} v_1^{(1)} + \cdots + \alpha_{n_1}^{(1)} v_{n_1}^{(1)}}_{=:w_1 \in \mathrm{Eig}_{\lambda_1}} + \cdots + \underbrace{\alpha_1^{(r)} v_1^{(r)} + \cdots + \alpha_{n_r}^{(r)} v_{n_r}^{(r)}}_{=:w_r \in \mathrm{Eig}_{\lambda_r}} = \mathbf{0} \tag{4.15}$$

Definiere $S := \{i \in \{1, \ldots, r\} \mid w_i \neq \mathbf{0}\}$.

- 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 $w_1, \ldots, w_r$ 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 $i \in \{1, \ldots, n\}$

$$w_i = \alpha_1^{(i)} v_1^{(i)} + \cdots + \alpha_{n_i}^{(i)} v_{n_1}^{(i)} = \mathbf{0}$$

  und daher $\alpha_1^{(i)} = \cdots = \alpha_{n_i}^{(1)} = 0$ da $v_1^{(i)}, \ldots, v_{n_i}^{(i)}$ Basis bilden.

Zu 2.: Die Basis $(v_1^{(i)}, \ldots, v_{n_i}^{(i)})$ von Eig_{λ_i} lässt sich zu Basis $\tilde{B} := (v_1^{(i)}, \ldots, v_{n_i}^{(i)}, \ldots, v_n{}^{(i)})$ von V ergänzen (Steinitz'scher Austauschsatz: Satz 2.4.13). Die Darstellungsmatrix hat dann die Form

$$M := M_{\tilde{B}}^{\tilde{B}}(f) = \left(\begin{array}{ccc:c} \lambda_i & \cdots & 0 & \\ \vdots & \ddots & \vdots & * \\ 0 & \cdots & \lambda_i & \\ \hdashline & & & \\ & \mathbf{0} & & * \\ & & & \end{array}\right)$$

Denn: die ersten n_i Spalten sind die Koordinatenvektoren der Bilder von $v_1^{(i)}, \ldots, v_{n_i}^{(i)}$ 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

$$\mathrm{grad}(\varphi \cdot \psi) = \mathrm{grad}(\varphi) + \mathrm{grad}(\psi)$$

folgt

$$\sum_{i=1}^{r} m_i \leq \operatorname{grad}(\chi_f) = n\,.$$

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 $u_1 = \varphi_B^{-1}(w_1), \dots, u_n = \varphi_B^{-1}(w_n)$ 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 $i = 1, \dots, n$, d.h., genau dann, wenn die Spalten Eigenvektoren sind. Da $(u_1, \dots, u_n)$ Basis von $\mathbb{K}^n$ ist, folgt dass

$$\operatorname{rg}(S) = n \Rightarrow S \in \mathrm{GL}(n, \mathbb{K})$$

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 $n = \mathrm{grad}(\chi_f) = \sum_i m_i = \sum_i n_i$ 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 $\mathrm{Eig}_\lambda(A)$ 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 $\dim(\mathrm{Lös}) = 3 - 1 = 2$

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 $\dim(\text{Lös}) = 3 - 2 = 1$.
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 25. 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 26. 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 27. Stimmen Sie der folgenden Aussage zu: *'die meisten Matrizen in* $\mathbb{C}^{2\times 2}$ *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. Wir nehmen an, dass Löwenzahn einjährig oder zweijährig ist (in seltenen Fällen ist er auch dreijährig, aber das vernachlässigen wir hier – es liesse sich aber analog behandeln). Das Wachstum verhält sich 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

$$0, 1, 1, 2, 3, 5, 8, \dots$$

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., $x_{m+1} = f(x_m)$ für lineare Abbildung

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Ax + b$$

Damit lässt sich x_m aus Anfangszustand x_0 berechnen:

$$x_m = f(x_{m-1}) = f^2(x_{m-2}) = \cdots = f^m(x_0) = A^m x_0 + (A^{m-1} + \cdots + A + E)b\,.$$

Beachte: $(A^{m-1} + \dots + A + E)(A - E) = A^m - E$ ('Teleskopsumme').
Falls $(A - E)$ invertierbar gilt also:

$$x_m = A^m x_0 + (A^m - E)(A - E)^{-1} b$$

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 $|\lambda_1|, \dots, |\lambda_n| < 1$ 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:

$$\lambda_1 = (1 + \sqrt{5})/2 \approx 1{,}6180339887\dots \qquad \text{mit Eigenvektor } \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = (1 - \sqrt{5})/2 \approx -0{,}618\dots \qquad \text{mit Eigenvektor } \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$$

$\lambda_1 > 1$ (“Goldener Schnitt”), unbegrenztes Wachstum.

$$\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & \lambda_1 - \lambda_2 \end{pmatrix}$$

Aus (4.18) folgt

$$\begin{aligned} x_m = \begin{pmatrix} t_{m+1} \\ t_m \end{pmatrix} &= \underbrace{\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}}_{S} \underbrace{\begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix}}_{D} \underbrace{\frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}}_{S^{-1}} \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{x_0} \\ &= \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1^{m+1} & \lambda_2^{m+1} \\ \lambda_1^m & \lambda_2^m \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} \end{aligned}$$

Also

$$\underbrace{t_m}_{\text{ganze Zahl}} = \underbrace{(\lambda_1^m - \lambda_2^m)}_{\text{irrationale Zahlen!}} / (\lambda_1 - \lambda_2)$$

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., $\exists \lambda_1, \dots, \lambda_n \in \mathbb{K}$ (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* oder *Hauptsatz 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.
Wir zeigen die Aussage für untere Dreiecksmatrix; dies ist äquivalent, da $\chi_A = \chi_{A^\top}$.

Die Aussage ist 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 $B = (u_1, u_2, \dots, u_{n+1})$ von $\mathbb{K}^{n+1}$. Sei R die Matrix mit den Spalten $u_1, u_2, \dots, u_{n+1}$. 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. □

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

Der Inhalt dieses Abschnitts ist als Ausblick zu verstehen. 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.:

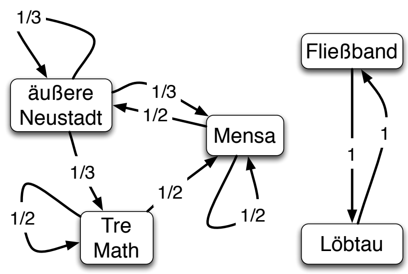
$$As = s \qquad \text{„\textit{stationäre Verteilung}“} \quad s$$

Eine Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ heißt

- *zeilenstochastisch* falls $0 \leq a_{ij} \leq 1$ 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'):

$$\begin{array}{cc} \begin{pmatrix} 1/3 & 1/2 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/2 & 0 & 0 \\ 1/3 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} & \begin{array}{c} \text{äußere Neustadt} \\ \text{Mensa} \\ \text{Tre Math} \\ \text{Fließband} \\ \text{Löbtau} \end{array} \end{array}$$

Wir zeichnen also genau dann einen gerichteten Pfeil von Knoten j nach Knoten i mit der Beschriftung ℓ, wenn in der i-ten Zeile und j-ten Spalte der Übergangsmatrix der Eintrag $\ell > 0$ steht. △

Lemma 4.3.28. *Jede stochastische Matrix $A \in \mathbb{R}^{n \times n}$ hat den Eigenwert 1.*

Beweis. Betrachten $e := (1, \ldots, 1)^\top \in \mathbb{R}^n$. Falls A zeilenstochastisch ist, gilt $Ae = e$, d.h., e ist Eigenvektor von A zum Eigenwert 1. Der spaltenstochastische Fall geht analog. □

Satz 4.3.29. *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 4.3.30. 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.

⇝ Dämpfungsfaktor bei Google PageRank.

⇝ 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 4.3.31 (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 4.3.32 ((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.

Kapitel 5

Dualität

Duale Räume in der linearen Algebra sind ein Beispiel für wichtiges Prinzip in der Mathematik, das Dualitätsprinzip. Manche Aussagen in diesem Zusammenhang gelten für beliebige Vektorräume, andere nur für endlich erzeugte. Wir behandeln also zunächst nochmal Satz 2.4.10: jeder Vektorraum besitzt eine Basis. Wir haben bisher nur einen Beweis für endlich erzeugte Vektorräume kennengelernt. Für den allgemeinen Fall müssen wir etwas ausholen.

5.1 Das Zornsche Lemma

Sei A eine Menge. Eine Relation $R \subseteq A^2$ heißt

- *Halbordnung auf A* (oder *partielle Ordnung*) falls sie reflexiv, antisymmetrisch, und transitiv ist (siehe Abschnitt 1.2);
- *total* falls für alle $a, b \in A$ gilt, dass $(a, b) \in R$ oder $(b, a) \in R$.
- *lineare Ordnung (oder Totalordnung) auf A* falls sie eine totale Halbordnung ist.

Beispiel 5.1.1. Die Ordnung $\leq$ der natürlichen Zahlen aus Abschnitt 1.2.10 ist eine lineare Ordnung. △

Wir schreiben $a < b$ falls $a \leq b$ und $a \neq b$.

Definition 5.1.2. Sei $\leq$ eine Halbordnung auf einer Menge A. Ein Element $a \in A$ heißt *minimal* falls es kein $b \in A$ gibt mit $b < a$. Eine Teilmenge $B \subseteq A$ heißt *Kette*, wenn die Einschränkung von $\leq$ auf B, also die Relation $\leq \cap B^2$, eine lineare Ordnung von B ist. Ein Element $a \in A$ heißt *untere Schranke* von $B \subseteq A$, falls $a \leq b$ für alle $b \in B$ gilt. Eine untere Schranke von B, die in B liegt, heißt ein *kleinstes Element* von B. Analog definiert man die Begriffe *maximales Element*, *obere Schranke*, und *größtes Element*.

Bemerkung 5.1.3. Besitzt eine Halbordnung ein kleinstes Element, so ist dieses eindeutig bestimmt und auch minimal. In einer Totalordnung ist ein minimales Element auch stets das kleinste Element. Jede endliche Halbordnung besitzt minimale und maximale Elemente, und jede endliche lineare Ordnung besitzt ein kleinstes und ein größtes Element.

Beispiel 5.1.4. Die Ordnung $\leq$ der natürlichen Zahlen $\mathbb{N}$ besitzt ein kleinstes Element 0, aber kein größtes Element und auch keine maximalen Elemente. △

Beispiel 5.1.5. Die Menge der echten Untervektorräume von $\mathbb{R}^3$, geordnet durch Inklusion, ist eine Halbordnung, aber keine lineare Ordnung. Sie besitzt ein kleinestes Element, nämlich $\{\mathbf{0}\}$, kein größtes Element, aber mehrere maximale Elemente, nämlich alle 2-dimensionalen Untervektorräume von $\mathbb{R}^3$. △

Der folgende Satz hat axiomatischen Charakter: er ist (in ZF) äequivalent zum Auswahlaxiom (siehe Abschnitt 1.2.9). Für einen Beweis des Satzes mit Hilfe des Auswahlaxioms verweisen wir z.B. auf das Skript zur Logikeinführung an der TU Dresden [2].

Satz 5.1.6 (Zornsches Lemma)**.** *Sei A eine nicht-leere Menge und $\leq$ eine Halbordnung auf A. Falls jede nicht-leere Kette eine obere Schranke in A besitzt, so hat A ein maximales Element.*

Wir verwenden diesen Satz, um Satz 2.4.10 in voller Allgemeinheit zu zeigen. Wir zeigen eine etwas allgemeinere Aussage (eine schwächere Form des Basisergänzungssatzes haben wir bereits in Satz 2.4.9 kennengelernt).

Satz 5.1.7 (Starker Basisergänzungssatz)**.** *Jede linear unabhängige Teilmenge A eines Vektorraumes V ist in einer Basis von V enthalten.*

Beweis. Sei $\mathcal{U}$ die Menge aller linear unabhängigen Teilmengen von V, die A enthalten, geordnet durch Inklusion. Dann ist $A \in \mathcal{U}$, also gilt $\mathcal{U} \neq \varnothing$.

Ist $\mathcal{V}$ eine nichtleere Kette in $\mathcal{U}$, so ist auch $W := \bigcup \mathcal{V}$ linear unabhängig: denn wenn $v_1, \ldots, v_n \in W$ paarweise verschieden sind, so gibt es $V_1, \ldots, V_n \in \mathcal{V}$ mit $v_i \in V_i$ für alle $i \in \{1, \ldots, n\}$. Da $\mathcal{V}$ linear geordnet ist, besitzt $\{V_1, \ldots, V_n\}$ ein größtes Element; ohne Beschränkung der Allgemeinheit sei dies V_1. Also sind $v_1, \ldots, v_n \in V_1$, und somit linear unabhängig. Folglich ist W eine obere Schranke von $\mathcal{V}$.

Nach dem Lemma von Zorn besitzt $\mathcal{U}$ ein maximales Element B. Das heißt, B ist eine maximale linear unabhängige Teilmenge von V, nach Satz 2.4.9 also eine Basis. □

5.2 Duale Räume

Seien V, W $\mathbb{K}$-Vektorräume. Dann ist

$$\mathrm{Hom}_{\mathbb{K}}(V, W) = \{f \mid f \colon V \to W \text{ ist lineare Abbildung}\}$$

selbst ein $\mathbb{K}$-Vektorraum (Bemerkung 3.4.21). Spezialfall $W = \mathbb{K}$:

Definition 5.2.1. Eine lineare Abbildung $f \colon V \to \mathbb{K}$ heißt auch *Linearform von V*, und

$$V^* := \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$$

heißt *Dualraum* von V.

Auf V^* sind wie folgt Vektorraumoperationen definiert:

- Addition: für $f, g \in V^*$ und $v \in V$
$$(f + g)(v) := f(v) + g(v).$$
- Multiplikation mit Skalar: für $\alpha \in \mathbb{K}$, $f \in V^*$, und $v \in V$
$$(\alpha f)(v) := \alpha f(v).$$

Der Nullvektor von V^* ist die Nullabbildung:

$$\mathbf{0}: V \to \mathbb{K} : v \mapsto 0.$$

Beispiel 5.2.2. Ist $V = \mathbb{K}^n$, so wird $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ beschrieben durch $\mathbb{K}^{1\times n} \simeq \mathbb{K}^n$. Wir können die Elemente von V als Spaltenvektoren und die Linearformen auf V als Zeilenvektoren auffassen. △

Übung 28. Sei V der unendlichdimensionale Vektorraum $\mathbb{R}^{\mathbb{N}}$ mit komponentenweiser Addition und Skalarmultiplikation (siehe Abschnitt 2.3.1). Definiere $t: V \to V$ durch $t(a_0, a_1, a_2, \dots) := (0, a_0, a_1, a_2, \dots)$. Zeige, dass t eine lineare Abbildung ist ohne Eigenvektoren.

5.3 Duale Basen

Sei $B = (b_i)_{i\in I}$ eine Basis von V. Dann gibt es zu jedem $i \in I$ genau ein $b_i^* \in V^*$ mit $b_i^*(b_j) = \delta_{ij}$ für alle $j \in I$. Hier ist δ_{ij}, das *Kroneckersymbol*, wie folgt definiert

$$\delta_{ij} := \begin{cases} 1 \in \mathbb{K} & \text{falls } i = j \\ 0 \in \mathbb{K} & \text{sonst.} \end{cases} \tag{5.1}$$

Offensichtlicherweise ist b_i^* eine lineare Abbildung, und falls $c^* \in V^*$ so, dass $c^*(b_j) = \delta_{ij}$ für alle $j \in I$, dann gilt $c^* = b_i^*$. Denn jedes $v \in V$ lässt sich schreiben als $\lambda_1 b_{i_1} + \dots + \lambda_n b_{i_n}$ für $i_1, \dots, i_n \in I$, $n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. Dann gilt

$$\begin{aligned} c^*(v) &= \lambda_1 c^*(b_{i_1}) + \dots + \lambda_n c^*(b_{i_n}) \\ &= \lambda_i \\ &= \lambda_1 b_i^*(b_{i_1}) + \dots + \lambda_n b_i^*(b_{i_n}) = b_i^*(v). \end{aligned}$$

Lemma 5.3.1. *Ist $B = (b_i)_{i\in I}$ eine Basis von V, dann ist $B^* := (b_i^*)_{i\in I}$ linear unabhängig. Ist I endlich, so ist B^* eine Basis von V^*.*

Beweis. Es sei $n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, und $i_1, \dots, i_n \in I$ beliebig. Wir definieren $v^* := \lambda_1 b_{i_1}^* + \dots + \lambda_n b_{i_n}^*$. Falls nun $v^* = 0$, dann gilt insbesondere $v^*(b_{i_j}) = 0$ für jedes $j \in \{1, \dots, n\}$. Da $v^*(b_{i_j}) = \lambda_j$ folgt $\lambda_j = 0$ für alle $j \in \{1, \dots, n\}$.

Falls nun I endlich ist, dann lässt sich jedes Element $f \in V^*$ schreiben als $\sum_{i\in I} f(b_i)b_i^*$: denn für jedes $j \in I$ gilt

$$\left(\sum_{i\in I} f(b_i)b_i^*\right)(b_j) = \sum_{i\in I} f(b_i)\underbrace{b_i^*(b_j)}_{\delta_{ij}} = f(b_j).$$

Die Aussage folgt, da zwei lineare Funktionen genau dann gleich sind, wenn sie die gleichen Werte auf der Basis B annehmen. □

Die Basis B^* von V^* heißt die zu B *duale Basis*.

Korollar 5.3.2. *Falls V endlichdimensional, so gilt $V \simeq V^*$.*

Beweis. Folgt aus Lemma 5.3.1. Folgt ebenfalls aus dem Fundamentalsatz der endlichdimensionalen Vektorräume (Satz 3.4.8). Ein konkreter Isomorphismus ist gegeben durch

$$\iota_B\colon V \to V^* : \sum_i^n \lambda_i b_i \mapsto \sum_i^n \lambda_i b_i^*.$$ □

5.4 Die natürliche Isomorphie $V \cong V^{**}$

Der *Bidualraum* zu V ist definiert als

$$V^{**} := (V^*)^* = \mathrm{Hom}_{\mathbb{K}}(V^*, \mathbb{K}).$$

Falls V endlich-dimensional, so gilt $V \cong V^* \cong V^{**}$ nach Korollar 5.3.2. Wie wir gleich sehen werden, sind die Vektorräume V und V^{**} auf besondere Weise ("natürlich") isomorph.

Für beliebige Vektorräume V und $v \in V$ sei

$$v^{**}\colon V^* \to \mathbb{K}$$

definiert durch $v^{**}(f) := f(v)$.

Satz 5.4.1. *Sei V ein Vektorraum. Die Abbildung*

$$\varphi\colon V \to V^{**} : v \mapsto v^{**}$$

ist ein injektiver Homomorphismus. Falls V endlichdimensional ist, so ist φ ein Isomorphismus – der natürliche Isomorphismus *zwischen V und V^{**}.*

Bemerkung 5.4.2. Wenn V abzählbar unendlich aber nicht endlich erzeugt ist, wie zum Beispiel $V = \mathbb{F}_2[X]$, dann gilt

$$\begin{aligned}
|V^{**}| &\geq |V^*| && \text{(nach Satz 5.4.1)}\\
&\geq |2^{\mathbb{N}}| && \text{(da } V \text{ unendliche Basis besitzt, Satz 2.4.10)}\\
&> |\mathbb{N}| && \text{(nach Satz 1.2.9)}\\
&= |V| && \text{(da } V \text{ abzählbar).}
\end{aligned}$$

Also kann φ nicht surjektiv sein (Satz 1.2.10).

Beweis von Satz 5.4.1. Zeigen zuerst, dass v^{**} linear für jedes $v \in V$; d.h., $v^{**} \in V^{**}$. Seien dazu $f, f' \in V^*$ und $\lambda \in \mathbb{K}$ beliebig. Dann gilt

$$\begin{aligned}
v^{**}(f_1 + \lambda f_2) &= (f_1 + \lambda f_2)(v) && \text{(Definition von } v^{**}\text{)}\\
&= f_1(v) + \lambda f_2(v) && \text{(Rechnen in } V^*\text{)}\\
&= v^{**}(f_1) + \lambda v^{**}(f_2) && \text{(Definition von } v^{**}\text{).}
\end{aligned}$$

Zeigen als nächstes, dass φ linear. Seien $v_1, v_2 \in V$ und $\lambda \in \mathbb{K}$. Sei $f \in V^*$ beliebig.

$$\begin{aligned}
\varphi(v_1 + \lambda v_2)(f) &= f(v_1 + \lambda v_2) && \text{(Definition von } \varphi\text{)}\\
&= f(v_1) + \lambda f(v_2) && \text{(Linearität von } f\text{)}\\
&= \varphi(v_1)(f) + \lambda \varphi(v_2)(f) && \text{(Definition von } \varphi\text{)}\\
&= (\varphi(v_1) + \lambda \varphi(v_2))(f)
\end{aligned}$$

Injektivität: Ist $v \in V \setminus \{\mathbf{0}\}$, so ist $\varphi(v) = v^{**} \neq \mathbf{0}$. Dazu ergänzen wir v zu einer Basis B von V und definieren $f \in V^*$ durch $f(u) := 1$ für alle $u \in B$. Dann gilt $v^{**}(f) = f(v) = 1$. Falls V endlichdimensional, so folgt Bijektivität aus Satz 3.4.6. □

Bemerkung 5.4.3. Sei V endlichdimensional. Im Gegensatz zu den Isomorphismen zwischen V und V^*, die von der Basis B abhängen, ist der Isomorphismus $\varphi\colon V \to V^{**}$ *natürlich* (oder *kanonisch*), soll heissen, unabhängig von der Wahl einer Basis.

5.5 Annulatoren

Sei V endlichdimensional, $S \subseteq V$.

Definition 5.5.1. Der *Annulator von S in V^** ist die Menge

$$S^0 := \{f \in V^* \mid f(s) = 0 \text{ für alle } s \in S\}.$$

Bemerkung 5.5.2. $S^0 \leq V^*$ (direktes Nachrechnen der Definition).

Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $U \leq V$ ein Untervektorraum.

Proposition 5.5.3. *Es gilt* $\dim U + \dim U^0 = \dim V$. *Wenn* $(u_1, \dots, u_k)$ *Basis von* U *und* $(u_1, \dots, u_k, u_{k+1}, \dots, u_n)$ *Basis von* V, *dann ist* $(u_{k+1}^*, \dots, u_n^*)$ *Basis von* U^0.

Beweis. $(u_{k+1}^*, \dots, u_n^*)$ Basis von U^0: zunächst gilt für $v \in U$ und $j \in \{k+1, \dots, n\}$ dass $u_j^*(v) = 0$ und damit dass $u_j^* \in U^0$. Denn $v = \sum_{i=1}^k \alpha_i u_i$ und damit ist $u_j^*(v) = \sum_{i=1}^k \alpha_i u_j^*(u_i) = 0$. Lineare Unabhängigkeit und $\langle u_{k+1}^*, \dots, u_n^* \rangle = U^0$: nachrechnen wie im Beweis von Lemma 5.3.1. Also: $\dim U + \dim U^0 = k + (n-k) = n = \dim V$. □

Bemerkung 5.5.4. Nach Satz (Satz 3.4.15) gilt $\dim V/U + \dim U = \dim V$ und damit

$$(V/U)^* \cong V/U \cong U^0 \qquad \text{(siehe Abschnitt 3.4.4).}$$

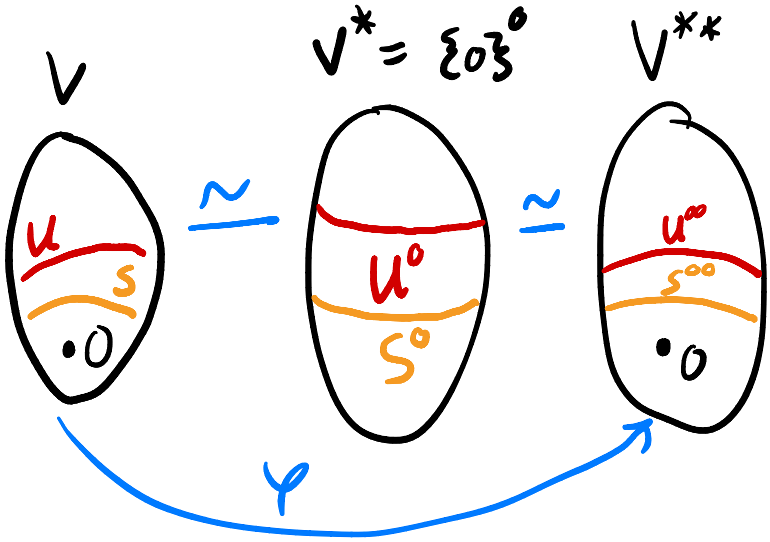


Abbildung 5.1: Illustration zu Dualraum und Annulator.

Betrachten nun $S \subseteq V$ beliebig und $S^{00} := (S^0)^0 \subseteq V^{**} \cong V$. Siehe Abbildung 5.1.

Proposition 5.5.5. *Sei φ der natürliche Isomorphismus zwischen V und V^{**}. Dann gilt $S^{00} = \varphi(S)$ genau dann wenn $S \leq V$.*

Beweis. Es gelte zunächst $S^{00} = \varphi(S)$. Mit der Bemerkung haben wir $S^{00} \leq V^{**}$. Also ist $S = \varphi^{-1}(S^{00}) \leq \varphi^{-1}(V^{**}) = V$.

Umgekehrt sei $S \leq V$. Sei $(u_1, \dots, u_k)$ Basis von S, und $(u_1, \dots, u_k, u_{k+1}, \dots, u_n)$ Basis von V. Zeigen zuerst $S^{00} \subseteq \varphi(S)$. Sei $w = \sum_{i=1}^n \alpha_i \varphi(u_i) \in S^{00}$ und sei $j \geq k+1$. Nach Proposition 5.5.3 ist $u_j^* \in S^0$. Dann gilt

$$
\begin{aligned}
0 &= w(u_j^*) && (w \in S^{00} \text{ und } u_j^* \in S^0) \\
&= \sum_{i=1}^n \alpha_i \varphi(u_i)(u_j^*) && (\text{Definition von } w) \\
&= \sum_{i=1}^n \alpha_i u_j^*(u_i) && (\text{Definition von } \varphi) \\
&= \alpha_j && (\text{Definition des Kroneckersymbols}).
\end{aligned}
$$

Also folgt

$$w = \sum_{i=1}^k \alpha_i \varphi(u_i) \in \varphi(S).$$

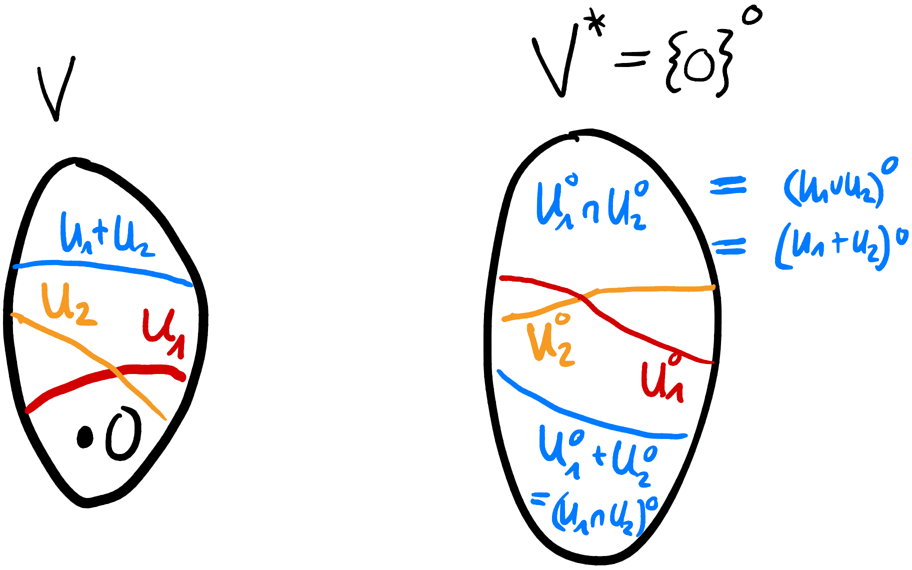


Abbildung 5.2: Illustration zu $U_1^0 \cap U_2^0 = (U_1 \cup U_2)^0 = (U_1 + U_2)^0$.

$\varphi(S) \subseteq S^{00}$: Sei $v \in S$. Zu zeigen: $\varphi(v) \in S^{00}$. Sei $f \in S^0$. Dann

$$\begin{aligned} \varphi(v)(f) &= f(v) && \text{(Definition von } \varphi) \\ &= 0 && \text{(da } v \in S \text{ und } f \in S^0). \end{aligned}$$ □

5.6 Dualitätssatz der linearen Algebra

Seien V, W endlichdimensionale $\mathbb{K}$-Vektorräume, und $f\colon V \to W$ eine lineare Abbildung. Dann ist

$$f^*\colon W^* \to V^* : g \mapsto g \circ f$$

eine lineare Abbildung, die *zu f duale Abbildung.*

Satz 5.6.1. *Sei V ein endlichdimensionaler Vektorraum. Dann ist $U \mapsto U^0$ eine bijektive Abbildung von der Menge der Untervektorräume von V auf die Menge der Untervektorräume von V^*. Dabei gelten:*

1. $\{\mathbf{0}\}^0 = V^*$
2. $U_1 \subseteq U_2 \implies U_1^0 \supseteq U_2^0$
3. $(U_1 \cap U_2)^0 = U_1^0 + U_2^0$ *(siehe Definition 2.4.17)*
4. $(U_1 \cup U_2)^0 = (U_1 + U_2)^0 = U_1^0 \cap U_2^0$

Für lineare Abbildung $f\colon V \to W$ gilt

5. $\operatorname{Kern}(f^*) = (\operatorname{Bild} f)^0$
   f^* *genau dann injektiv wenn* f *surjektiv.*

6. $\operatorname{Bild}(f^*) = (\operatorname{Kern} f)^0$
   f^* *genau dann surjektiv wenn* f *injektiv.*

7. $\operatorname{rg}(f^*) = \operatorname{rg}(f)$
   f^* *genau dann bijektiv, wenn* f *bijektiv.*

Beweis. Zu 1: $\{\mathbf{0}\}^0 = \{f \in V^* \mid f(\mathbf{0}) = \mathbf{0}\} = V^*$.
Zu 2: Sei $f \in U_2^0$. Dann gilt für alle $u \in U_2^0$, dass $f(u) = 0$. Wenn $U_1 \subseteq U_2$, dann gilt auch für alle $u \in U_1^0$, dass $f(u) = 0$. Also $f \in U_1^0$.
3.-7.: Übung! □

Korollar 5.6.2. *Sei* V *endlichdimensionaler Vektorraum und* $S \subseteq V$*. Dann gilt*

$$\langle S\rangle = \varphi^{-1}(S^{00})$$

Beweis. Aus $S \subseteq \langle S\rangle$ folgt durch zweimaliges Anwenden von Satz 5.6.1 (2.) dass $S^{00} \subseteq \langle S\rangle^{00}$. Da $S^{00} \leq V$ haben wir also

$$S^{00} = \langle S\rangle^{00} = \varphi(\langle S\rangle)$$

nach Proposition 5.5.5, und damit die Aussage des Korollars. □

Übung 29. Satz 3.3.19 trägt den Namen *Dualität*. Diskutieren Sie, ob dieser Titel gerechtfertigt ist. Wie sieht das duale System vom dualen System aus? Gibt es einen Zusammenhang zur Dualität, wie sie in diesem Kapitel betrachtet wurde?

Bemerkung 5.6.3. Es gibt noch andere Kontexte in der Mathematik, in denen der Begriff *Dualität* verwendet wird, z.B.

- Komplementbildung bezüglich der Teilmengen einer Menge A (Schnitt ist dann dual zu Vereinigung, siehe Abschnitt 1.1.3)
- Negation in der Aussagenlogik (Konjunktion ist dann dual zur Disjunktion, der Allquantor ist dual zum Existenzquantor, siehe Abschnitt 1.3.1).
- Eine ganz andere Form der Dualität kommt aus der Graphentheorie: dort ist der *Dualgraph* eines ebenen Graphen G definiert als der Graph G^*, dessen Knoten die Flächen von G sind, und in dem zwei Knoten mit einer Kante verbunden sind, wenn sich entsprechenden Flächen in G eine Kante teilen. Hier läßt sich zeigen, dass $(G^*)^*$ isomorph ist zu G^*. Siehe Abbildung 5.1.

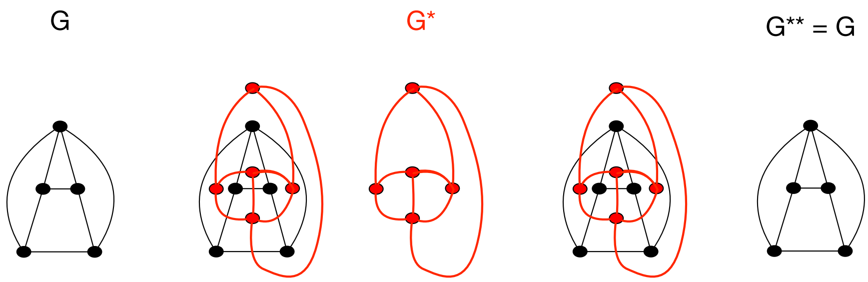


Abbildung 5.3: Der Dualgraph G^* eines ebenen Graphen G, und der Dualgraph des Dualgraphen $(G^*)^*$.

Kapitel 6

Analytische Geometrie

Bisher: abstrakte $\mathbb{K}$-Vektorräume V; allerdings: $V \simeq \mathbb{K}^n$.
Dieses Kapitel: Spezialisierung $\mathbb{K} = \mathbb{R}$ und oft $n = 2, 3$ ($\rightsquigarrow$ zusätzliche Eigenschaften).

6.1 Das Skalarprodukt

Nicht zu verwechseln mit "Produkt mit einem Skalar" (skalare Multiplikation).

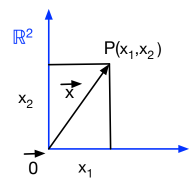
6.1.1 Wiederholung und Bezeichnungen

Verschiedene Interpretationen der Paare $(x_1, x_2) \in \mathbb{R}^2$:

- Punkte $P(x_1, x_2)$ der *euklidischen Ebene* mit Koordinaten x_1, x_2 (bezüglich festgelegtem Koordinatensystem)
- *Translation* der Ebene: $(y_1, y_2) \mapsto (y_1 + x_1, y_2 + x_2)$.
- Zeilenvektoren $(x_1, x_2) \in \mathbb{R}^{1\times 2}$
- Spaltenvektoren $\binom{x_1}{x_2} \in \mathbb{R}^{2\times 1}$

Entsprechende Verallgemeinerungen auf $\mathbb{R}^3$, $\mathbb{R}^4$, ...

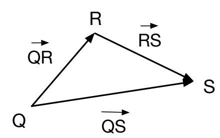
Interpretation als komplexe Zahle $x_1 + x_2 i$: Spezialität von $\mathbb{R}^2 \simeq \mathbb{C}$.



- Darstellung eines Punktes durch 'Ortsvektor':
  Pfeil von Punkt $P(0, 0)$ zu Punkt $P(x_1, x_2)$.

- Darstellung von Translation durch 'freien Vektor' (x_1, x_2):
  beschreibt Pfeil $\overrightarrow{RS}$ von Punkt R nach Punkt S.

  Addition: Komposition von Translationen.
  Aneinandersetzen der Pfeile: $\overrightarrow{QR} + \overrightarrow{RS} = \overrightarrow{QS}$.



6.1.2 Länge (Norm) eines Vektors

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n\times 1}$$

Länge (oder *Norm*) von $\vec{x}$:

$$\|\vec{x}\| := \sqrt{x_1^2 + \cdots + x_n^2}$$

$n = 2$: Pythagoras!
Abstand zweier Punkte: Seien $P : \vec{x} = (x_1, x_2)$ und $Q : \vec{y} = (y_1, y_2)$ zwei Punkte in $\mathbb{R}^2$, dann gilt

$$\|\overrightarrow{PQ}\| = \|\vec{y} - \vec{x}\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2}$$

6.1.3 Das Skalarprodukt

Seien $\vec{x}, \vec{y} \in \mathbb{R}^n$. Betrachten $\vec{x}, \vec{y}$ als Elemente von $\mathbb{R}^{n\times 1}$.

$$\vec{x} * \vec{y} := (\vec{x})^\top \vec{y} = (x_1, \ldots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n$$

heißt *inneres* oder *Skalarprodukt* (andere Schreibweise: $\langle \vec{x}, \vec{y} \rangle$). Was für eine Abbildung ist $*$?

$$*\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : (\vec{x}, \vec{y}) \mapsto \vec{x} * \vec{y}$$

1. $*$ ist *bilinear*, d.h., für alle $\vec{v} \in \mathbb{R}^n$ sind die Abbildungen $f_{\vec{v}}\colon \mathbb{R}^n \to \mathbb{R} : \vec{x} \mapsto \vec{x} * \vec{v}$ und die Abbildung $f_{\vec{v}}\colon \mathbb{R}^n \to \mathbb{R} : \vec{x} \mapsto \vec{v} * \vec{x}$ lineare Abbildungen (siehe Abschnitt 3.4). Sind eh die gleiche Abbildung.

2. $*$ ist symmetrisch (d.h., kommutativ):

$$\vec{x} * \vec{y} = \vec{y} * \vec{x}$$

3. $*$ ist *positiv definit*, d.h.,

$$\begin{aligned} \vec{x} \neq \mathbf{0} &\Rightarrow \vec{x} * \vec{x} > 0 \\ \vec{x} = \mathbf{0} &\Rightarrow \mathbf{0} * \mathbf{0} = 0 \qquad \text{(folgt bereits aus Bilinearität).} \end{aligned}$$

Allgemein ist *ein Skalarprodukt* eines $\mathbb{R}$-Vektorraumes V eine bilineare, symmetrische, und positiv definite Abbildung von V^2 nach $\mathbb{R}$. Deswegen spricht man im Fall vom oben eingeführten Skalarprodukt $*$ des $\mathbb{R}^n$ auch vom *üblichen* oder *Standard*-Skalarprodukt. Dies ist nur das nächstliegende Skalarprodukt; z.B. ist für jede invertierbare $n\times n$-Matrix A auch $(\vec{x}, \vec{y}) \mapsto A\vec{x} * A\vec{y}$ ein Skalarprodukt.

Definition 6.1.1. Ein *euklidischer Vektorraum* ist ein $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt auf V.

Beispiel 6.1.2. Der $\mathbb{R}$-Vektorraum aller stetigen reellen Funktionen auf dem Interval $[-\pi, +\pi]$ ist euklidisch mit Skalarprodukt

$$g * h := \frac{1}{\pi}\int_{-\pi}^{+\pi} g(t)h(t)dt. \qquad \triangle$$

6.1.4 Die Ungleichung von Cauchy-Schwarz

Ist V ein euklidischer Vektorraum mit Skalarprodukt $*$, und $\vec{x} \in V$, so versteht man unter der *Norm* von x die reelle Zahl $||\vec{x}|| := \sqrt{\vec{x} * \vec{x}} \geq 0$.

Satz 6.1.3. *Es sei V ein euklidischer Vektorraum mit Skalarprodukt $*$. Dann gilt*

$$\vec{x} * \vec{y} \leq ||\vec{x}||\cdot||\vec{y}||$$

Äquivalent dazu ist:

$$(\vec{x} * \vec{y})(\vec{y} * \vec{x}) \leq (\vec{x} * \vec{x})(\vec{y} * \vec{y})$$

(Rückrichtung durch Wurzelziehen, da $||\vec{x}||, ||\vec{y}|| \geq 0$.)

Beweis. Falls $\vec{y} = \mathbf{0}$ ist die Aussage trivial. Sei nun $\vec{y} \neq \mathbf{0}$. Setzen $\alpha := \frac{\vec{x} * \vec{y}}{||y||^2}$. Nun gilt

$$\begin{aligned} 0 \leq (\vec{x} - \alpha\vec{y}) * (\vec{x} - \alpha\vec{y}) &= \vec{x} * \vec{x} - 2\alpha(\vec{x} * \vec{y}) + \alpha^2(\vec{y} * \vec{y}) \\ &= ||\vec{x}||^2 - 2\frac{(\vec{x} * \vec{y})^2}{||\vec{y}||^2} + \frac{(\vec{x} * \vec{y})^2}{||\vec{y}||^2} \\ &= ||\vec{x}||^2 - \frac{(\vec{x} * \vec{y})^2}{||\vec{y}||^2}. \end{aligned}$$

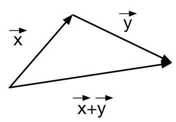
Damit ist $(\vec{x} * \vec{y})^2 \leq ||\vec{x}||^2||\vec{y}||^2$. □

Bemerkung 6.1.4. Falls $\vec{y}$ ein positives Vielfaches von $\vec{x}$, also wenn $\vec{y} = \lambda\vec{x}$ für ein $\lambda \in \mathbb{R}_{\geq 0}$, dann gilt sogar Gleichheit: denn

$$\vec{x} * \lambda\vec{x} = \lambda \cdot ||\vec{x} * \vec{x}|| = \lambda \cdot ||\vec{x}|| \cdot ||\vec{x}|| = ||\vec{x}|| \cdot ||\lambda\vec{x}||.$$

6.1.5 Die Dreiecksungleichung

Es sei V ein euklidischer Vektorraum. Dann gilt



$$||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$$

Beweis.

$$\begin{aligned} ||\vec{x} + \vec{y}||^2 &= (\vec{x} + \vec{y}) * (\vec{x} + \vec{y}) \\ &= \vec{x} * \vec{x} + 2(\vec{x} * \vec{y}) + \vec{y} * \vec{y} && \text{Bilinearität} \\ &\leq ||\vec{x}||^2 + 2||\vec{x}||\cdot||\vec{y}|| + ||\vec{y}||^2 && \text{Cauchy-Schwarz (Satz 6.1.3)} \\ &= (||\vec{x}|| + ||\vec{y}||)^2 \end{aligned}$$

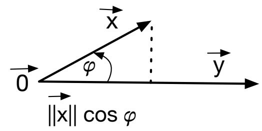
Da Norm nicht-negativ kann man Wurzel ziehen und erhält das gewünschte. □

6.1.6 Geometrische Interpretation des Skalarproduktes im $\mathbb{R}^2$

Es gilt für $n = 2$:

$$\vec{x} * \vec{y} = ||\vec{x}|| \cdot ||\vec{y}|| \cdot \cos\varphi \tag{6.1}$$

wobei $\varphi = \measuredangle(\vec{x}, \vec{y})$ Winkel zwischen $\vec{x}$ und $\vec{y}$ (im Bogenmaß).



Produkt wird negativ, falls $\frac{\pi}{2} < \varphi < \frac{3\pi}{2}$. (stumpfe Winkel)
Produkt wird Null, falls $\varphi = \frac{\pi}{2}$ oder $\varphi = \frac{3\pi}{2}$.
Produkt wird positiv, falls $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$. (spitze Winkel)

Projektion von Vektor $\vec{x}$ auf Gerade g mit Richtung $\vec{y}$ hat Länge

$$||\vec{x}|| \cdot \cos\varphi = \frac{\vec{x} * \vec{y}}{||\vec{y}||}$$

Projektion von Vektor $\vec{x}$ auf g ist also (genaueres zu Projektionen: siehe Abschnitt 6.2.4):

$$\frac{\vec{x} * \vec{y}}{||\vec{y}||^2}\vec{y}.$$

Bemerkung 6.1.5. Die Gleichung (6.1) kann als *Definition* für den Winkel φ zwischen $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ benutzt werden.

Spezialfall: $\vec{x} \perp \vec{y} \;:\Leftrightarrow\; \measuredangle(\vec{x}, \vec{y}) = \{\pi/2, -\pi/2\} \;\Leftrightarrow\; \vec{x} * \vec{y} = 0$ (Orthogonalität)

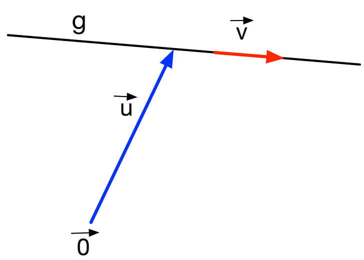
6.2 Geradendarstellungen

Geraden im $\mathbb{R}^n$ können auf verschiedene Arten dargestellt werden.

6.2.1 Parameterdarstellung

Die *Parameterdarstellung* einer Geraden g im $\mathbb{R}^n$ (durch den Punkt $\vec{u}$ und mit dem Richtungsvektor $\vec{v}$):

$$g = \vec{u} + \mathbb{R}\vec{v} = \{\vec{u} + \lambda\vec{v} \mid \lambda \in \mathbb{R}\}$$



In Koordinatendarstellung:

$$\begin{aligned} x_1 &= u_1 + \lambda v_1 \\ &\vdots \\ x_n &= u_n + \lambda v_n \end{aligned}$$

Alternativ: Statt mit Richtungsvektor $\vec{v}$ kann die Gerade g auch mit einem weiteren Punkt $\vec{w}$ auf g dargestellt werden ($\vec{v} = \vec{w} - \vec{u}$).

6.2.2 Hessesche Normalform

Sei g Gerade in $\mathbb{R}^2$. Sei $\vec{u} \in g$, und $\vec{n}$ *Normalenvektor* von g: $\vec{n} \perp g$ (soll heissen $\vec{n} \perp \vec{v}$ für Richtungsvektor $\vec{v}$ von g). Wir fordern zusätzlich $||\vec{n}|| = 1$ ($\vec{n}$ ist *Normaleneinheitsvektor*). Dann gilt:

$$\begin{aligned} \vec{x} \in g &\Longleftrightarrow (\vec{x} - \vec{u}) \perp \vec{n} && (\Longleftarrow \text{ nur für } n = 2!) \\ &\Longleftrightarrow \vec{n} * (\vec{x} - \vec{u}) = 0 \\ &\Longleftrightarrow \vec{n} * \vec{x} = \vec{n} * \vec{u} \end{aligned}$$

Die *Hessesche Normalform* einer Geraden $g \in \mathbb{R}^2$:

$$\vec{n} * \vec{x} = d$$

für $d := \vec{n} * \vec{u}$ (hängt nicht von der Wahl von $\vec{u} \in g$ ab!).

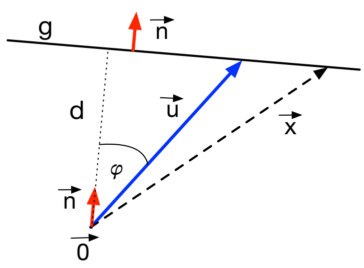
Für $\vec{n} = \binom{n_1}{n_2}$ ergibt sich:

$$n_1 x_1 + n_2 x_2 = d$$

mit $n_1^2 + n_2^2 = 1$ (wegen $||\vec{n}||= 1$; 'Normierung').
Dabei ist d der *vorzeichenbehaftete Abstand* der Geraden g zum Nullpunkt:

$$d = \vec{n} * \vec{u} = ||\vec{n}|| \cdot ||\vec{u}|| \cdot \cos\varphi = ||\vec{u}|| \cdot \cos\varphi$$

Positiv falls $-90^\circ < \varphi < 90^\circ$, negativ falls $90^\circ < \varphi < 270^\circ$.



Rechtfertigung des Begriffes *Abstand*:

- für alle $\vec{x} \in g$ gilt

  $$d \leq ||\vec{x}||$$

  denn $d = \vec{n} * \vec{x} \leq ||\vec{n}|| \cdot ||\vec{x}||= ||\vec{x}||$ nach Cauchy-Schwarz (Satz 6.1.3).

- Es gibt ein $\vec{x}_0 \in g$ mit $|d| = ||\vec{x}_0||$: siehe Abschnitt 6.2.5.

6.2.3 Koordinatendarstellung

Gerade g im $\mathbb{R}^2$: für $(a_1, a_2) \neq (0,0)$

$$g = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = a_0\} = \text{Lös}((a_1, a_2), a_0)$$

Lösungsmenge eines linearen Gleichungssystems der Form $a_1x_1 + a_2x_2 = a_0$.

Zugehörige Hessesche Normalform:

$$\underbrace{\frac{a_1}{\sqrt{a_1^2 + a_2^2}}}_{=:n_1} x_1 + \underbrace{\frac{a_2}{\sqrt{a_1^2 + a_2^2}}}_{=:n_2} x_2 = \underbrace{\frac{a_0}{\sqrt{a_1^2 + a_2^2}}}_{=:d}$$

es gilt $n_1^2 + n_2^2 = 1$.

Parameterdarstellung im Fall $a_2 \neq 0$ ist z.B.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{a_0}{a_2} \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -\frac{a_1}{a_2} \end{pmatrix}$$

Setze $x_1 = \lambda$, dann $x_2 = (a_0 - a_1\lambda)/a_2$. Analog für $a_1 \neq 0$.

6.2.4 (Orthogonale) Projektionen

Sei g eine Gerade in $\mathbb{R}^2$ mit Richtungsvektor $\vec{v}$. Die *(orthogonale) Projektion* eines Punktes $\vec{q} \in \mathbb{R}^2$ auf g ist $\vec{p} \in g$ mit $(\vec{q} - \vec{p}) \perp \vec{v}$, das heißt $(\vec{q} - \vec{p}) * \vec{v} = 0$.
Existiert stets, und ist eindeutig. Bild!

Berechnung aus Hessescher Normalform.
Gegeben: $g = \{\vec{x} \in \mathbb{R}^2 \mid \vec{n} * \vec{x} = d\}$ und $\vec{q} \in \mathbb{R}^2$.
Gesucht: Projektion $\vec{p}$ von $\vec{q}$ auf Gerade g.
Antwort: $\vec{p} = \vec{q} + \lambda_0 \cdot \vec{n}$ für $\lambda_0 = d - \vec{n} * \vec{q}$.
Beweis:

$$\begin{aligned} (\vec{q} - \vec{p}) * \vec{v} &= (\vec{q} - \vec{q} - \lambda_0\vec{n}) * \vec{v} = 0 \\ \vec{n} * \vec{p} &= \vec{n} * (\vec{q} + \lambda_0\vec{n}) \\ &= \vec{n} * \vec{q} + \lambda_0\|\vec{n}\|^2 = \vec{n} * \vec{q} + d - \vec{n} * \vec{q} = d \end{aligned}$$

(Existenz gezeigt.)

Berechnung aus Parameterform.
Gegeben: $g = \vec{u} + \mathbb{R}\vec{v}$ und $\vec{q} \in \mathbb{R}^n$.
Gesucht: Projektion $\vec{p}$ von $\vec{q}$ auf Gerade g.

Antwort:

$$\vec{p} = \vec{u} + \lambda_1 \vec{v} \text{ mit } \lambda_1 = \frac{(\vec{q} - \vec{u}) * \vec{v}}{\|\vec{v}\|^2} \tag{6.2}$$

Beweis: $(\vec{q} - \vec{p}) \perp \vec{v}$ und $\vec{p} \in g$. Das heißt, $(\vec{q} - \vec{p}) * \vec{v} = 0$ und $\exists \lambda_1 : \vec{p} = \vec{u} + \lambda_1 \vec{v}$. Einsetzen ergibt

$$\begin{aligned} &(\vec{q} - \vec{u} - \lambda_1 \vec{v}) * \vec{v} = 0 \\ \Rightarrow\ &(\vec{q} - \vec{u}) * \vec{v} - \lambda_1 (\vec{v} * \vec{v}) = 0 \\ \Rightarrow\ &\lambda_1 = \frac{(\vec{q} - \vec{u}) * \vec{v}}{\|\vec{v}\|^2} . \end{aligned}$$

(Eindeutigkeit gezeigt.)

6.2.5 Zusammenhang Projektionen mit Hessescher Normalform

Es sei g eine Gerade im $\mathbb{R}^2$ und $\vec{p}_0$ die Projektion von $\mathbf{0}$ auf g. Dann ist $\vec{p}_0 \perp \vec{v}$, also ist $\vec{n} := \frac{\vec{p}_0}{\|\vec{p}_0\|}$ Normaleneinheitsvektor von g. Das bedeutet, dass

$$\vec{n} * \vec{x} = d$$

mit $d := \vec{n} * \vec{p}_0$ die Hessesche Normalform von g ist.

Es gilt $|d| = |\vec{n} * \vec{p}_0| = |\frac{\vec{p}_0}{\|\vec{p}_0\|} * \vec{p}_0| = \|p_0\|$ und damit ist p_0 wirklich der Punkt auf g, der nächstmöglich an $\mathbf{0}$ liegt – daher also die Bezeichnung von $|d|$ als der *Abstand* von g zum Ursprung (siehe Abschnitt 6.2.2).

Wenn g in Parameterdarstellung gegeben ist durch $g = \vec{u} + \mathbb{R}\vec{v}$, so können wir mit (6.2) die Hessesche Normalform von g bestimmen, indem wir $\vec{p}_0$ ausrechnen wie folgt:

$$\vec{p}_0 = \vec{u} + \frac{-\vec{u} * \vec{v}}{\|\vec{v}\|^2} \vec{v}$$

6.2.6 Abstand Punkt-Gerade

Es sei $g \subseteq \mathbb{R}^2$ eine Gerade, gegeben über die Hessesche Normalform $\vec{n} * \vec{x} = d$. Der (vorzeichenbehaftete) *Abstand* $d_{\vec{q}} \in \mathbb{R}$ zwischen $\vec{q} \in \mathbb{R}^2$ und der Geraden g ist

$$d_{\vec{q}} := \vec{n} * \vec{q} - d$$

und es gilt

- $\|\vec{q} - \vec{p}\| = |d_{\vec{q}}|$ für $\vec{p} := \vec{q} - d_{\vec{q}} \cdot \vec{n}$ (der Projektion von $\vec{q}$ auf g)

  $$\|\vec{q} - \vec{p}\| = \|\vec{q} - \vec{q} + d_{\vec{q}} \cdot \vec{n}\| = |d_{\vec{q}}|$$

- für alle $\vec{x} \in g$ gilt $d_{\vec{q}} \leq \|\vec{q} - \vec{x}\|$, denn wegen Cauchy-Schwarz gilt

  $$d_{\vec{q}} = d - \vec{n} * \vec{q} = \vec{n} * \vec{x} - \vec{n} * \vec{q} = \vec{n} * (\vec{x} - \vec{q}) \leq \|\vec{n}\| \cdot \|\vec{x} - \vec{q}\| = \|\vec{x} - \vec{q}\| .$$

- $d_{\vec{p}} = 0$ falls $\vec{q} \in g$.
  $d_{\vec{p}} > 0$ falls $\vec{n}$ und $(\vec{q} - \vec{u})$ spitzen Winkel bilden, für ein (äquivalent: für alle) $\vec{u} \in g$
  $d_{\vec{p}} < 0$ falls $\vec{n}$ und $(\vec{q} - \vec{u})$ stumpfen Winkel bilden. Bild!

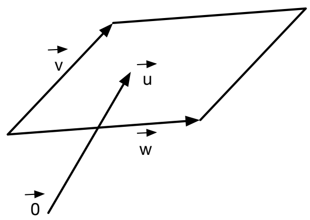
Übung 30. Zeigen Sie: drei Punkte $u, v, w \in \mathbb{R}^n$ liegen genau dann auf einer Geraden, wenn es $a, b, c \in \mathbb{R}$ gibt, die nicht alle gleich Null sind, so dass $au + bv + cw = 0$ und $a + b + c = 0$.

6.3 Ebenendarstellungen

6.3.1 Parameterdarstellung

Darstellung durch Punkt $\vec{u}$ und zwei linear unabhängige Richtungsvektoren $\vec{v}$, $\vec{w}$:

$$\begin{aligned} E &:= \vec{u} + \langle \vec{v}, \vec{w} \rangle & \text{(Abschnitt 2.4.1)} \\ &= \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w} \end{aligned}$$



$\vec{x} \in E$ falls es $\lambda, \mu \in \mathbb{R}$ gibt mit $\vec{u} + \lambda\vec{v} + \mu\vec{w}$.

Koordinatenweise:

$$\begin{aligned} x_1 &= u_1 + \lambda v_1 + \mu w_1 \\ &\vdots \\ x_n &= u_n + \lambda v_n + \mu w_n \end{aligned}$$

6.3.2 Hessesche Normalform einer Ebene im $\mathbb{R}^3$

Analog zu Abschnitt 6.2.2 (Geraden im $\mathbb{R}^2$).

$$E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}$$

Sei $\vec{n}$ *normierter Normalenvektor von* E, d.h., $||\vec{n}|| = 1$, $\vec{n} \perp \vec{v}$, $\vec{n} \perp \vec{w}$.

$$\begin{aligned} \vec{x} \in E &\Leftrightarrow (\vec{x} - \vec{u}) \perp \vec{n} \\ &\Leftrightarrow \vec{n} * (\vec{x} - \vec{u}) = 0 & (\textit{Hessesche Normalform} \text{ von } E) \end{aligned}$$

Anders geschrieben:

$$E = \{\vec{x} \mid \vec{n} * \vec{x} = d\}$$

mit $d := \vec{n} * \vec{u}$ (der vorzeichenbehaftete Abstand zwischen Nullpunkt und E).

Allgemein (analog zu Abschnitt 6.2.6): Für Punkt Q, mit $\vec{q} = \overrightarrow{\mathbf{0}Q}$, ist

$$\vec{n} * (\vec{q} - \vec{u})$$

der vorzeichenbehaftete Abstand von Q zur Ebene E.
Also: wenn dieser Ausdruck Null wird, liegt $\vec{q}$ auf der Ebene.

6.3.3 Koordinatendarstellung

Koordinatendarstellung einer Ebene E im $\mathbb{R}^3$:

$$E = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid a_1x_1 + a_2x_2 + a_3x_3 = d\} = \text{Lös}((a_1, a_2, a_3), d)$$

Von Koordinatendarstellung in Hessesche Normalform: definiere

$$\vec{a} := \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{und} \quad \vec{n} := \frac{\vec{a}}{||\vec{a}||}.$$

Es gilt

$$E = \{\vec{x} \mid \vec{n} * \vec{x} = d\}.$$

6.3.4 Orthogonalprojektion von Punkt auf Ebene

Gegeben: Punkt $\vec{q} \in \mathbb{R}^3$, Ebene

$$E = \vec{u} + \mathbb{R}\vec{v} + \mathbb{R}\vec{w}.$$

Gesucht: Orthogonalprojektion $\vec{p}$ von $\vec{q}$ auf E, d.h.,

1. $\vec{p} \in E$, d.h., es gibt λ_0 und μ_0 mit $\vec{p} = \vec{u} + \lambda_0\vec{v} + \mu_0\vec{w}$
2. $\vec{v} * (\vec{q} - \vec{p}) = 0$ und $\vec{w} * (\vec{q} - \vec{p}) = 0$.

Einsetzen von 1. in 2. ergibt Gleichungssystem für λ_0 und μ_0:

$$\begin{aligned} \lambda_0(\vec{v} * \vec{v}) + \mu_0(\vec{v} * \vec{w}) &= \vec{v} * (\vec{q} - \vec{u}) \\ \lambda_0(\vec{w} * \vec{v}) + \mu_0(\vec{w} * \vec{w}) &= \vec{w} * (\vec{q} - \vec{u}) \end{aligned}$$

Lösung mit Cramerscher Regel (Abschnitt 4.1.7):

$$\lambda_0 = \frac{\begin{vmatrix} \vec{v} * (\vec{q} - \vec{u}) & \vec{v} * \vec{w} \\ \vec{w} * (\vec{q} - \vec{u}) & \vec{w} * \vec{w} \end{vmatrix}}{\begin{vmatrix} \vec{v} * \vec{v} & \vec{v} * \vec{w} \\ \vec{w} * \vec{v} & \vec{w} * \vec{w} \end{vmatrix}}$$

$$\mu_0 = \frac{\begin{vmatrix} \vec{v} * \vec{v} & \vec{v} * (\vec{q} - \vec{u}) \\ \vec{w} * \vec{v} & \vec{w} * (\vec{q} - \vec{u}) \end{vmatrix}}{\begin{vmatrix} \vec{v} * \vec{v} & \vec{v} * \vec{w} \\ \vec{w} * \vec{v} & \vec{w} * \vec{w} \end{vmatrix}}$$

Dann ist der gesuchte Vektor $\vec{p}$:

$$\vec{p} = \vec{u} + \lambda_0 \vec{v} + \mu_0 \vec{w}$$

Ist E in Hessescher Normalform gegeben,

$$E = \{\vec{x} \in \mathbb{R}^3 \mid \vec{n} * (\vec{x} - \vec{u}) = 0\}$$

so lässt sich die Projektion $\vec{p}$ von $\vec{q}$ auf E wie folgt berechnen:
$\vec{q} - \vec{p}$ hat gleiche Richtung wie $\vec{n}$.

$$\begin{aligned} \vec{p} &= \vec{q} + (\vec{p} - \vec{q}) \\ &= \vec{q} + \underbrace{(d - \vec{n} * \vec{q})}_{\text{vorzeichenbehaftete Länge von } \vec{p}-\vec{q}} \cdot \vec{n} \end{aligned}$$

6.4 Das äußere Produkt (Vektorprodukt)

Anwendungen in Mathematik, Physik und Informatik, z.B.:

- Berechnung des Drehmoments, oder der Lorenzkraft
  (bewegte Ladung im magnetischen Feld)
- Abstandsformel windschiefer Geraden
- Algorithmische Geometrie

Ausgangsidee: Wollen von zwei Richtungsvektoren, die eine Ebene im $\mathbb{R}^3$ definieren, möglichst bequem an einen Normelenvektor der Ebene kommen.

Das *äußere Produkt* oder *Vektorprodukt*

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 : (\vec{a}, \vec{b}) \mapsto \vec{a} \times \vec{b}$$

wird durch folgende Eigenschaften (eindeutig!) definiert:

1. $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$
2. $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$
3. $\vec{e}_3 \times \vec{e}_1 = \vec{e}_2$
4. $\times$ ist bilinear (siehe Abschnitt 6.1.3)
5. $\times$ ist *schiefsymmetrisch*, d.h., $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$.

Bemerkung: $\vec{e}_i \times \vec{e}_i = \mathbf{0}$ folgt aus Schiefsymmetrie für $\vec{a} = \vec{b}$.

Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$,

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3 \qquad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3$$

dann gilt

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ -a_1b_3 + a_3b_1 \\ a_1b_2 - a_2b_1 \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{pmatrix} \tag{6.3}$$

Nachrechnen: Ausdruck in (6.3) erfüllt alle Bedingungen der Definition des Vektorproduktes; wir haben also inbesondere die Existenz eines Vektorproduktes bewiesen. Zur Eindeutigkeit: Aus der Bilinearität von $\times$ erhählt man

$$\begin{aligned}
\vec{a} \times \vec{b} &= (a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3) \times \vec{b} \\
&= a_1(\vec{e}_1 \times \vec{b}) + a_2(\vec{e}_2 \times \vec{b}) + a_3(\vec{e}_3 \times \vec{b}) && \text{(Bilinearität)} \\
&= \quad a_1(b_1(\vec{e}_1 \times \vec{e}_1) + b_2(\vec{e}_1 \times \vec{e}_2) + b_3(\vec{e}_1 \times \vec{e}_3)) \\
&\quad + a_2(b_1(\vec{e}_2 \times \vec{e}_1) + b_2(\vec{e}_2 \times \vec{e}_2) + b_3(\vec{e}_2 \times \vec{e}_3)) \\
&\quad + a_3(b_1(\vec{e}_3 \times \vec{e}_1) + b_2(\vec{e}_3 \times \vec{e}_2) + b_3(\vec{e}_3 \times \vec{e}_3)) && \text{(Bilinearität)} \\
&= a_1(b_2\vec{e}_3 - b_3\vec{e}_2) + a_2(-b_1\vec{e}_3 + b_3\vec{e}_1) + a_3(b_1\vec{e}_2 - b_2\vec{e}_1) && \text{(1., 2., 3., und 5.)} \\
&= (a_2b_3 - a_3b_2)\vec{e}_1 + (a_1b_3 - a_3b_1)\vec{e}_2 + (-a_1b_2 + a_2b_1)\vec{e}_3 && \text{(Zusammenfassen)}
\end{aligned}$$

6.4.1 Beziehungen zwischen Vektorprodukt und Skalarprodukt

Vektorprodukt weder kommutativ noch assoziativ.

1. Der *Grassmannsche Entwicklungssatz*:

$$(\vec{a} \times \vec{b}) \times \vec{c} = (\vec{a} * \vec{c}) \cdot \vec{b} - (\vec{b} * \vec{c}) \cdot \vec{a}$$

2. Das *Spatprodukt* $(\vec{a} \times \vec{b}) * \vec{c}$ erfüllt

$$(\vec{a} \times \vec{b}) * \vec{c} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Aus unserem Wissen über Determinanten lassen sich nun viele Eigenschaften für das Spatprodukt herleiten, z.B. $(\vec{a} \times \vec{b}) * \vec{c} = (\vec{b} \times \vec{c}) * \vec{a} = (\vec{c} \times \vec{a}) * \vec{b}$

3. Die *Lagrangesche Identität*:

$$(\vec{a} \times \vec{b}) * (\vec{c} \times \vec{d}) = (\vec{a} * \vec{c}) \cdot (\vec{b} * \vec{d}) - (\vec{a} * \vec{d}) \cdot (\vec{b} * \vec{c})$$

Beweis: ausrechnen mit Hilfe von (6.3).

Beweis. Für Grassmann:

$$\begin{aligned}
(\vec{a} \times \vec{b}) \times \vec{c} &= \begin{pmatrix} a_2b_3 - a_3b_2 \\ -a_1b_3 + a_3b_1 \\ a_1b_2 - a_2b_1 \end{pmatrix} \times \vec{c} \\
&= \begin{pmatrix} (-a_1b_3 + a_3b_1)c_3 - (a_1b_2 - a_2b_1)c_2 \\ -(a_2b_3 - a_3b_2)c_3 + (a_1b_2 - a_2b_1)c_1 \\ (a_2b_3 - a_3b_2)c_2 - (-a_1b_3 + a_3b_1)c_1 \end{pmatrix} \\
&= \begin{pmatrix} -a_1b_3c_3 + a_3b_1c_3 - a_1b_2c_2 + a_2b_1c_2 \\ -a_2b_3c_3 + a_3b_2c_3 + a_1b_2c_1 - a_2b_1c_1 \\ (a_2b_3c_2 - a_3b_2c_2 + a_1b_3c_1 - a_3b_1c_1 \end{pmatrix} \\
&= \begin{pmatrix} a_1b_1c_1 + a_2b_2c_1 + a_3b_3c_1 \\ a_1b_1c_2 + a_2b_2c_2 + a_3b_3c_2 \\ a_1b_1c_3 + a_2b_2c_3 + a_3b_3c_3 \end{pmatrix} - \begin{pmatrix} a_1b_1c_1 + a_1b_2c_2 + a_1b_3c_3 \\ a_2b_1c_1 + a_2b_2c_2 + a_2b_3c_3 \\ a_3b_1c_1 + a_3b_2c_2 + a_3b_3c_3 \end{pmatrix} \\
&= (\vec{a} * \vec{c}) \cdot \vec{b} - (\vec{b} * \vec{c}) \cdot \vec{a}
\end{aligned}$$

Für das Spatprodukt:

$$\begin{aligned}
(\vec{a} \times \vec{b}) * \vec{c} &= \begin{pmatrix} a_2b_3 - a_3b_2 \\ -a_1b_3 + a_3b_1 \\ a_1b_2 - a_2b_1 \end{pmatrix} * \vec{c} \\
&= a_2b_3c_1 - a_3b_2c_1 - a_1b_3c_2 + a_3b_1c_2 + a_1b_2c_3 - a_2b_1c_3 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}
\end{aligned}$$

Für Lagrange:

$$\begin{aligned}
(\vec{a} \times \vec{b}) * (\vec{c} \times \vec{d}) &= ((\vec{c} \times \vec{d}) \times a) * \vec{b} && \text{(Gleichheit fürs Spatprodukt)} \\
&= ((\vec{c} * \vec{a})\vec{d} - (\vec{d} * \vec{a})\vec{c}) * \vec{b} && \text{(Grassmann)} \\
&= (\vec{a} * \vec{c}) \cdot (\vec{b} * \vec{d}) - (\vec{a} * \vec{d}) \cdot (\vec{b} * \vec{c}) && \text{(Rechnen mit Skalarprodukt)} \ \square
\end{aligned}$$

6.4.2 Geometrische Interpretation des Vektorproduktes

Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$. Dann gilt

- $(\vec{a} \times \vec{b}) \perp \vec{a}$ und $(\vec{a} \times \vec{b}) \perp \vec{b}$
- $\|\vec{a} \times \vec{b}\| = \|\vec{a}\| \cdot \|\vec{b}\| \cdot |\sin\varphi|$ ist der Flächeninhalt des von den Vektoren $\vec{a}$ und $\vec{b}$ aufgespannten Parallelogramms.

Beweis. Rechnen zunächst nach, dass $(\vec{a} \times \vec{b}) * \vec{a} = 0$.
Setze $\vec{c} := \vec{a}$ im Spatprodukt aus Punkt 3 in Abschnitt 6.4.1: erhalten

$$(\vec{a} \times \vec{b}) * \vec{a} = \begin{vmatrix} a_1 & b_1 & a_1 \\ a_2 & b_2 & a_2 \\ a_3 & b_3 & a_3 \end{vmatrix} = 0$$

Der Flächeninhalt des von $\vec{a}$ und $\vec{b}$ aufgespannten Parallelograms mit Höhe h ist $\|\vec{a}\|h$, wobei $h = \|\vec{b}\||\sin\varphi|$ mit $\varphi := \measuredangle(\vec{a}, \vec{b})$. Nach der Lagrangeschen Identität gilt

$$\begin{aligned}(\vec{a} \times \vec{b}) * (\vec{a} \times \vec{b}) &= \|\vec{a}\|^2\|\vec{b}\|^2 - (\vec{a} * \vec{b})^2 \\ &= \|\vec{a}\|^2\|\vec{b}\|^2 - \|\vec{a}\|^2\|\vec{b}\|^2(\cos\varphi)^2 \\ &= \|\vec{a}\|^2\|\vec{b}\|^2\,(1 - \cos^2\varphi) \\ &= \|\vec{a}\|^2\|\vec{b}\|^2\,\sin^2\varphi. \end{aligned}$$

□

Das Spatprodukt

$$(\vec{a} \times \vec{b}) * \vec{c} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

ist das *vorzeichenbehaftete Volumen* des von $\vec{a}, \vec{b}, \vec{c}$ aufgespannten *Spats* (Form der Kristalle im Kalkspat; auch *Parallelepiped*).

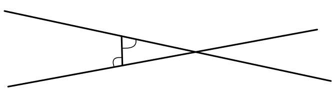
- Vorzeichen positiv, falls $\vec{a}, \vec{b}, \vec{c}$ "Rechtssystem"
(sonst "Linkssystem", linke Hand Regel).

- $(\vec{a} \times \vec{b}) * \vec{c} = \|\vec{a} \times \vec{b}\| \cdot \|\vec{c}\| \cdot |\cos\varphi|$; hier ist $\|\vec{c}\| \cdot |\cos\varphi|$ die *Höhe* des Spats und $\|\vec{a} \times \vec{b}\|$ der Flächeninhalt der von $\vec{a}$ und $\vec{b}$ aufgespannten *Grundfläche* des Spats.

6.4.3 Anwendung: Abstand zweier Geraden

Seien g_1, g_2 Geraden in $\mathbb{R}^3$, gegeben als

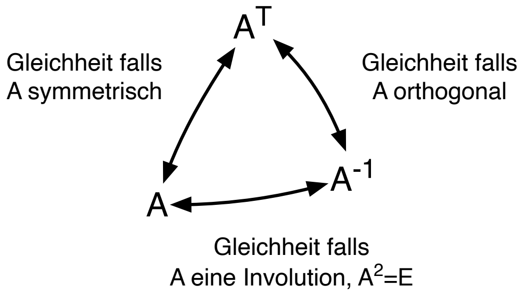
$$\begin{aligned} g_1 &= \vec{u}_1 + \mathbb{R}\vec{v}_1 \\ g_2 &= \vec{u}_2 + \mathbb{R}\vec{v}_2 \end{aligned}$$



Das Vektorprodukt $\vec{v}_3 := \vec{v}_1 \times \vec{v}_2$ ist genau dann $\mathbf{0}$, wenn $\vec{v}_1$ und $\vec{v}_2$ Vielfache voneinander sind, d.h., wenn g_1 und g_2 parallel sind. Ansonsten steht $\vec{v}_3$ senkrecht auf $\vec{v}_1$ und $\vec{v}_2$. Sei E_1 die Ebene durch $\vec{u}_1$ mit Normalenvektor $\vec{v}_3$. Alle Punkte von g_2 haben den gleichen Abstand zu E_1. Es genügt also, den Abstand von $\vec{u}_2$ zu E_1 zu berechnen – und das geht wie in Abschnitt 6.3.4.

6.5 Orthogonale lineare Abbildungen

Eine lineare Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^n$ heißt *orthogonal* falls für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt: $\vec{x} * \vec{y} = f(\vec{x}) * f(\vec{y})$. Die Abbildung f *erhält* das Skalarprodukt. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal* falls $A^\top = A^{-1}$ (insbesondere: A ist invertierbar).[1]



Proposition 6.5.1. *Falls A eine orthogonale Matrix ist, dann ist $x \mapsto Ax$ eine orthogonale lineare Abbildung.*

Beweis. Falls $A^\top A = E_n$, dann gilt für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$

$$\vec{x} * \vec{y} = \vec{x}^\top \vec{y} = \vec{x}^\top (A^\top A)\vec{y} = (A\vec{x})^\top (A\vec{y}) = f(\vec{x}) * f(\vec{y}). \qquad \square$$

Bemerkung 6.5.2. Es gilt sogar die Umkehrung: falls eine lineare Abbildung f orthogonal ist, und $A = M_B^B(f)$ für die Standardbasis $B = (e_1, \ldots, e_n)$ von $\mathbb{R}^n$, so ist A orthogonal (Satz 8.3.5).

Eigenschaften von orthogonalen Abbildungen:

Proposition 6.5.3. *Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ orthogonal. Dann gilt für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$:*

1. $\|\vec{x}\| = \|f(\vec{x})\|$.
2. $\vec{x} \perp \vec{y} \Leftrightarrow f(\vec{x}) \perp f(\vec{y})$.
3. $|\det f| = 1$.

Beweis. 1 und 2 folgen direkt aus der Definition von Orthogonalität (f erhält das Skalarprodukt, also auch Rechtwinkligkeit und Norm). Zu 3:

$$\begin{aligned} \det(f) = \det(A) &= \det(A^\top) && \text{(Proposition 4.1.14)} \\ &= \det(A^{-1}) && \text{(Proposition 6.5.1)} \\ &= \det(A)^{-1} = \det(f)^{-1} && \text{(Satz 4.1.15)} \end{aligned}$$

Also muss gelten $|\det A| = 1$. $\square$

[1] In Abschnitt 8.3.2 werden wir diese Begriffe noch etwas allgemeiner definieren.

6.5.1 Die Gruppen $O(n)$, $SO(n)$

- Das Produkt orthogonaler Matrizen ist orthogonal.
- Die inverse Matrix einer orthogonalen Matrix ist wieder orthogonal.

Also bildet $O(n) := \{M \in \mathbb{R}^{n\times n} \mid M \text{ orthogonal}\}$ bezüglich der Matrizenmultiplikation eine Gruppe (eine Untergruppe von $\mathrm{GL}(n, \mathbb{R})$, Abschnitt 3.2.1).

$$SO(n) := \{M \in O(n) \mid \det M = 1\}$$

ist eine Untergruppe von $O(n)$, die *spezielle orthogonale Gruppe*.
Verwenden $\det(A \cdot B) = \det(A) \cdot \det(B)$

6.5.2 Die orthogonale Gruppe $O(2)$

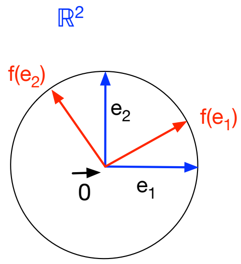
$f\colon \mathbb{R}^2 \to \mathbb{R}^2 : \vec{x} \mapsto A\vec{x}$ durch Bilder $f(e_1), f(e_2)$ (Spalten von A) eindeutig festgelegt.

$$f \text{orthogonal} \Leftrightarrow \|f(e_1)\| = 1 = \|f(e_2)\| \text{ und } f(e_1) \perp f(e_2)$$

Wählt man $f(e_1)$ beliebig mit $\|f(e_1)\| = 1$, so gibt es nur 2 Möglichkeiten für $f(e_2)$:

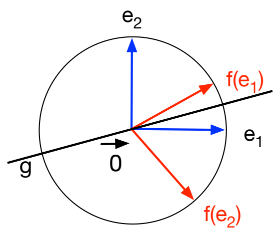
- Drehung um Winkel $\alpha = \measuredangle(\vec{e}_1, f(\vec{e}_1)) = \measuredangle(\vec{e}_2, f(\vec{e}_2))$.

$$M_B^B(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} =: D(\alpha)$$



- Spiegelung an Gerade g.

$$M_B^B(f) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} =: S(\alpha)$$



Also:

$$O(2) = \underbrace{\{D(\alpha) \mid 0 \le \alpha \le 2\pi\}}_{\text{Drehungen}} \cup \underbrace{\{S(\alpha) \mid 0 \le \alpha \le 2\pi\}}_{\text{Spieglungen}}$$

Wegen $\det D(\alpha) = 1$ und $\det S(\alpha)) = -1$ folgt

$$SO(2) = \{D(\alpha) \mid 0 \le \alpha \le 2\pi\} \qquad \text{(Menge aller Drehungen).}$$

Berechnung des Drehwinkels α. Sei $M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in SO(2)$. Dann ist

$$M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

und es gilt

$$\cos\alpha = a_{11} = a.$$

Berechnung der Spiegelungsachse g. Für $b = 0$ ist

- $\alpha \in \{0, \pi\}$, da $\sin\alpha = b$,
- $a = 1$, da $\cos 0 = \cos\pi = 1$, und
- $g = \mathbb{R} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbb{R} \cdot \vec{e}_1$.

Also nehmen wir im folgenden an, dass $b \neq 0$.

Ansatz ohne Winkel:

$$\vec{x} \in g \Leftrightarrow M\vec{x} = \vec{x}$$

Sprich: $\vec{x}$ ist Eigenvektor von M zum Eigenwert 1.
Lineares Gleichungssystem:

$$\begin{aligned} ax_1 + bx_2 &= x_1 \\ bx_1 - ax_2 &= x_2 \end{aligned}$$

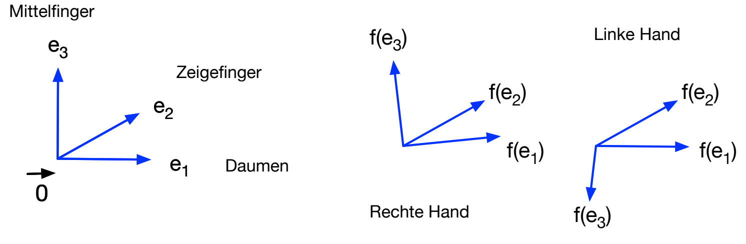
gesucht ist nichttriviale Lösung. Probieren zunächst $x_1 = 1$.
Wir erhalten $a + bx_2 = 1$ und damit $x_2 = (1-a)/b$ (wie bereits erwähnt ist $b \neq 0$). Also:

$$g = \mathbb{R} \cdot \begin{pmatrix} 1 \\ (1-a)/b \end{pmatrix}$$

Falls g keinen Punkt der Form $\binom{1}{x_2}$ enthält, so gilt $g = \mathbb{R} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbb{R} \cdot \vec{e}_2$.

6.5.3 Die orthogonale Gruppe $O(3)$

$f\colon \mathbb{R}^2 \to \mathbb{R}^3 : \vec{x} \mapsto A\vec{x}$ durch Bilder $f(e_1), f(e_2), f(e_3)$ (Spalten von A) eindeutig festgelegt. Zwei Möglichkeiten: sind im Rechtssystem ($\det A = 1$) oder im Linkssystem ($\det A = -1$).



Bemerkung: Jede Drehung $A \in SO(3)$ lässt sich als Hintereinanderausführungen von Drehungen um die Koordinatenachsen eindeutig beschreiben.

$$A = D_1(\alpha) \circ D_2(\beta) \circ D_3(\gamma)$$

(α, β, γ: *Eulersche Winkel*.)

Anwendung: Satellitenjustierung.

Kapitel 7

Normalformen von Matrizen

7.1 Klassifikation und Normalformen

7.1.1 Was heißt 'klassifizieren'?

Nicht exakt definierbar – es hängt davon ab, was man erreichen will.

Ausgangssituation:

- Menge M.
  Z.B. $\mathbb{K}^{n\times n}$, $\mathrm{Hom}(V,W)$, ...
- Äquivalenzrelation $E \subseteq M \times M$ (Abschnitt 1.2.1).
  Z.B. Ähnlichkeit von Matrizen, Äquivalenz von Matrizen (im engeren Sinne), ...

Klassifikation heißt

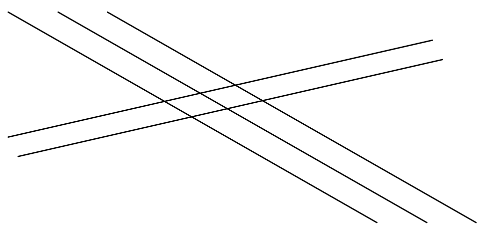
- Festlegen einer Äquivalenzrelation
- Gutes Verständnis der Faktormenge M/E und der Zuordnung $M \mapsto M/E$.
  Insbesondere: wann sind zwei Element äquivalent.

A: Klassifikation durch charakteristische Daten

Beispiel. M: Menge aller Geraden in $\mathbb{R}^2$.

$$(g_1, g_2) \in E :\Longleftrightarrow g_1 || g_2 \qquad \text{(Parallelität)}$$

Charakteristisches Datum: Anstiegswinkel α.



$$\begin{aligned} f\colon M &\to \overbrace{D}^{\text{"Daten"}} \\ g &\mapsto \alpha \qquad\qquad\qquad\qquad \alpha \text{ Anstiegswinkel von } g \end{aligned}$$

Es gilt $E(g_1, g_2) \iff f(g_1) = f(g_2) \iff [g_1]_E = [g_2]_E$ d.h.,

$$E = \operatorname{Kern} f := \{(x,y) \in M \mid f(x) = f(y)\} \tag{7.1}$$

Kern einer Abbildung im Sinne von (1.1). Ist formal verschieden vom Kern einer linearen Abbildung. Aber es gibt natürlich einen Zusammenhang, siehe Übung 31.

Übung 31. Es sei $f\colon A \to B$ eine lineare Abbildung, und $E := \{(x,y) \in M \mid f(x) = f(y)\}$ der Kern von f im Sinne von (7.1), und sei $\operatorname{Kern}(f)$ der Kern von f im Sinne der linearen Algebra (Definition 3.8). Zeigen Sie:

- für alle $a \in A$ gilt genau dann $a \in \operatorname{Kern}(f)$, wenn $E(a, 0)$.
- für alle $a, b \in A$ gilt $E(a,b)$ genau dann, wenn $a - b \in \operatorname{Kern}(f)$.

B: Klassifikation durch Repräsentanten

Auswahl eines Repräsentanten aus jeder Äquivalenzklasse aus M/E:
Gesucht: $N \subseteq M$ so dass jede Äquivalenzklasse genau einen Repräsentanten in N hat.

D.h.,

$$N \to M/E : m \mapsto [m]_E$$

ist bijektiv.
Elemente aus N heißen auch *Normalformen*.

Beispiel. M ist Menge aller Geraden in $\mathbb{R}^2$ und E ist Parallelität.
N: Menge der Geraden durch $\mathbf{0}$.

Typische Anforderungen:

- Für gegebene $m_1, m_2 \in M$, entscheide ob $(m_1, m_2) \in E$.
- Zu jedem $m \in M$ finde Normalform, d.h., finde $n \in N$ mit $(m, n) \in E$.

Erstes Problem lässt sich auf zweites zurückführen!

7.1.2 Äquivalenz

Wiederholung: Seien $A, B \in \mathbb{K}^{n \times m}$.

$$A \sim B :\iff A \text{ und } B \text{ sind } \textit{äquivalent} \text{ (im engeren Sinne)}$$

d.h., es gibt invertierbare Matrizen $S \in \mathbb{K}^{n \times n}$ und $T \in \mathbb{K}^{m \times m}$ so dass

$$B = TAS\,.$$

Eine Äquivalenzrelation.

Klassifikation durch charakteristische Daten: $A \sim B \iff \mathrm{rg}(A) = \mathrm{rg}(B)$ (Satz 3.4.27, Charakterisierung von Äquivalenz, Abschnitt 3.4.8).

Klassifikation durch Repräsentanten: als Normalform für die Äquivalenzklasse aller Matrizen in $\mathbb{K}^{m\times n}$ vom Rang r kann man folgende Matrix wählen (siehe Beweis von Satz 3.4.27):

$$\begin{pmatrix} E_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \in \mathbb{K}^{m\times n}$$

7.1.3 Zeilenäquivalenz

Wiederholung Definition 3.4.24: $A, B \in \mathbb{K}^{n\times m}$ heissen *zeilenäquivalent* (oder *linksäquivalent*) falls es eine invertierbare Matrix S gibt so dass $B = SA$.

- Zeilenäquivalenz definiert eine Äquivalenzrelation auf $\mathbb{K}^{n\times m}$.
- Falls A und B zeilenäquivalent sind, dann auch äquivalent. (Zeilenäquivalenz liefert *feinere* Unterscheidung als Äquivalenz.)

Motivation: Wenn die erweiterten Koeffizientenmatrizen von zwei Gleichungssystemen zeilenäquivalent sind, dann haben sie den gleichen Lösungsraum (Lemma 3.3.12).

Jede Matrix ist zeilenäquivalent zu einer Matrix in Stufenform (Definition 3.2.21); aber offensichtlich gibt es zeilenäquivalente Matrizen mit derselben Stufenform: beispielsweise

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \xrightarrow[(2)]{2z_1 \rightsquigarrow z_1} \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix} \xrightarrow[(2)]{\frac{1}{2}z_1 \rightsquigarrow z_1} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Definition 7.1.1. Eine Matrix $A \in \mathbb{K}^{m\times n}$ ist in *reduzierter Stufenform*, falls

- A in Stufenform ist,
- der führende (linkeste) Eintrag jeder Zeile, der nicht 0 ist, ist 1, und
- jede Spalte, die eine 1 enthält, in allen anderen Einträgen 0 ist.

$$A = \begin{pmatrix}
0 & \cdots & 0 & \mathbf{1} & a_{1j_1} & 0 & \cdots & 0 & \cdots & 0 & \cdots & a_{1n} \\
0 & \cdots & 0 & \cdots & 0 & \mathbf{1} & \ldots & a_{2j_3} & \cdots & & & \vdots \\
\vdots & & & & \cdots & 0 & \cdots & 0 & \ddots & & & \vdots \\
\vdots & & & & & & & & \ddots & \mathbf{1} & \cdots & a_{rn} \\
0 & \cdots & & & & & & & \cdots & 0 & \cdots & 0 \\
\vdots & & & & & & & & & & & \vdots \\
0 & \cdots & & & & & & & \cdots & 0 & \cdots & 0
\end{pmatrix}$$

Satz 7.1.2. *Jede Matrix $A \in \mathbb{K}^{m\times n}$ lässt sich durch elementare Zeilenumformungen in eine* eindeutige *Matrix N in reduzierter Stufenform überführen; N ist zeilenäquivalent zu A.*

Beweis. Wir wissen bereits, dass sich A durch elementare Zeilenumformungen in eine Matrix in Stufenform (Abschnitt 3.2.4) überführen lässt. Es ist leicht zu sehen, dass sich jede Matrix in Stufenform durch elementare Zeilenumformungen weiter in eine Matrix N in reduzierter Stufenform umformen lässt: zum Beispiel

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

Elementare Zeilenumformungen lässen sich durch Multiplikation von links mit invertierbaren Elementarmatrizen beschreiben. Da das Produkt von invertierbaren Matrizen ebenfalls invertierbar ist (3.3), gibt es also eine invertierbare Matrix S mit $SA = N$.

Zur Eindeutigkeit: seien A und B zeilenäquivalente Matrizen in reduzierter Stufenform. Angenommen $A \neq B$; wir wollen dies zum Widerspruch führen. Sei $i \in \{1, \dots, n\}$ minimal, so dass sich A und B in der i-ten Spalte unterscheiden. Sei A' (bzw. B') die Matrix, die aus der i-ten Spalte von A (bzw. B) und allen Spalten mit kleinerem Index besteht, deren Einträge nur einmal 1 und sonst nur 0 sind. Dann ist A' notwendigerweise von der Gestalt

$$A' = \begin{pmatrix} E_k & r \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \text{ oder } A' = \begin{pmatrix} E_k & \mathbf{0} \\ \mathbf{0} & 1 \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Analog ist B' von der Gestalt

$$B' = \begin{pmatrix} E_k & s \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \text{ oder } B' = \begin{pmatrix} E_k & \mathbf{0} \\ \mathbf{0} & 1 \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

für ein $k \in \{1, \dots, n\}$ und $r, s \in K^k$. Die Matrizen A' und B' sind zeilenäquivalent, da A' aus A und B' aus B durch Wegstreichen von Spalten entsteht und sich Zeilenäquivalenz dadurch nicht ändert. Wenn also A' von der zweiten Gestalt ist, so ist auch B' von der zweiten Gestalt, was nicht sein kann, da sich A und B per Annahme in der i-ten Spalte unterscheiden.

Beide Matrizen können aufgefasst werden als erweiterte Matrizen eines linearen Gleichungssystems (wie in Abschnitt 3.3.4). Das Gleichungssystem für A' ist

$$\begin{pmatrix} E_k \\ \mathbf{0} \end{pmatrix} \bar{x} = \begin{pmatrix} r \\ \mathbf{0} \end{pmatrix},$$

also von der Gestalt

$$x_1 = r_1, \dots, x_k = r_k, 0 = 0, \dots, 0 = 0,$$

hat also die *eindeutige* Lösung r. Analog hat das System für B' die eindeutige Lösung s. Da beide Systeme zeilenäquivalent sind, gilt $r = s$, wieder ein Widerspruch. □

Äquivalenzbegriff	Normalform
Äquivalenz	Abschnitt 7.1.2
Zeilenäquivalenz	Reduzierte Stufenform, Abschnitt 7.1.3
Ähnlichkeit	Frobenius-Normalform (Satz 7.2.31), Abschnitt 7.1.4

Abbildung 7.1: Übersicht zu Normalformen von Matrizen.

Korollar 7.1.3. *$A \in \mathbb{K}^{n\times n}$ ist genau dann invertierbar, wenn sich A schreiben lässt als Produkt von Elementarmatrizen.*

Beweis. Offensichtlich ist jede Elementarmatrix invertierbar (Bemerkung 3.2.19), und das Produkt von invertierbaren Matrizen ist ebenfalls invertierbar (3.3).

Sei umgekehrt A invertierbar. Dann lässt sich A nach Satz 7.1.2 durch elementare Zeilenumformungen in eine Matrix N in reduzierter Stufenform überführen. Diese hat den gleichen Rang wie A (Lemma 3.2.17), nämlich n; also ist N von der Gestalt E_n. Die Zeilenumformungen lassen sich beschreiben durch ein Produkt $S = S_1 S_2 \cdots S_k$ von Elementarmatrizen. Die Matrix S ist invertierbar und es gilt $SA = N = E_n$, also ist $A = S^{-1} = S_k^{-1} \cdots S_2^{-1} S_1^{-1}$ ein Produkt von Elementarmatrizen. □

7.1.4 Ähnlichkeit

Wiederholung: $A \approx A' :\Leftrightarrow A$ und A' sind *ähnlich*, d.h., es gibt invertierbare Matrix S mit

$$A' = S^{-1} A S.$$

Eine Äquivalenzrelation. Ist feiner als Zeilenäquivalenz (und Äquivalenz).

Motivation: Ähnliche Matrizen beschreiben die gleiche lineare Abbildung! (Satz 3.4.28)

Fragen:

- wie entscheiden wir, ob zwei Matrizen ähnlich sind?
- was ist möglichst einfache/schöne/praktische Normalform?

Bemerkung 7.1.4. Falls A diagonalisierbar: die Diagonalmatrix als NF (Satz 4.3.19; eindeutig bis auf Reihenfolge der Eigenwerte). Aber nicht jede Matrix ist diagonalisierbar.

Bemerkung 7.1.5. Falls $\mathbb{K} = \mathbb{C}$: jede Matrix trigonalisierbar (Abschnitt 4.3.6). Allerdings: wenig Kontrolle über die Einträge der Dreiecksmatrix oberhalb der Diagonalen.

7.2 Die Frobenius-Normalform

Müssen ein wenig ausholen …

7.2.1 Das Charakteristische Polynom II

$\mathbb{K}[X]$: der Polynomring, siehe Abschnitt 4.2.

Wiederholung: das charakteristische Polynom $\chi_A(X)$ von $A = (a_{ij})_{i,j\leq n} \in \mathbb{K}^{n\times n}$:

$$\chi_A(X) := \begin{vmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & & -a_{2n} \\ \vdots & & \ddots & \\ -a_{n1} & -a_{n2} & & X - a_{nn} \end{vmatrix}$$

aus Abschnitt 4.3.2 (Kapitel zur Berechnung von Eigenwerten). Das folgende Beispiel zeigt, dass jedes Polynom das charakteristische Polynom einer Matrix ist.

Beispiel 7.2.1. Sei $\varphi(X)$ das Polynom $X^n + \alpha_{n-1}X^{n-1} + \cdots + \alpha_1 X + \alpha_0 \in \mathbb{K}[X]$, und sei

$$Z_\varphi := \begin{pmatrix} 0 & & \cdots & & 0 & -\alpha_0 \\ 1 & 0 & & & 0 & -\alpha_1 \\ 0 & 1 & 0 & & & -\alpha_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & & \ddots & 1 & 0 & -\alpha_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -\alpha_{n-1} \end{pmatrix}.$$

Dann ist φ das charakteristische Polynom von Z_φ. Berechnen dazu

$$\chi_{Z_\varphi}(X) = \begin{vmatrix} X & 0 & \cdots & 0 & \alpha_0 \\ -1 & X & & 0 & \alpha_1 \\ 0 & -1 & \ddots & 0 & \alpha_2 \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & -1 & X + \alpha_{n-1} \end{vmatrix}$$

wie folgt: Addition des X-fachen der zweiten Zeile zur ersten und anschließendes Vertauschen der ersten beiden Zeilen liefert

$$\chi_{Z_\varphi}(X) = -\begin{vmatrix} -1 & X & 0 & \dots & 0 & \alpha_1 \\ 0 & X^2 & 0 & \dots & 0 & \alpha_0 + X\alpha_1 \\ 0 & -1 & X & & 0 & \alpha_2 \\ & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & -1 & X & \alpha_{n-2} \\ 0 & \cdots & & 0 & -1 & X + \alpha_{n-1} \end{vmatrix}.$$

Addition des X^2-fachen der dritten Zeile zur zweiten und anschließendes Vertauschen der zweiten und dritten Zeile liefert

$$\chi_{Z_\varphi}(X) = \begin{vmatrix} -1 & X & 0 & 0 & & \alpha_1 \\ 0 & -1 & X & 0 & & \alpha_2 \\ 0 & 0 & X^3 & 0 & & \alpha_0 + X\alpha_1 + X^2\alpha_2 \\ 0 & 0 & -1 & X & & \\ \vdots & & \ddots & \ddots & & \alpha_{n-2} \\ 0 & \cdots & & 0 & -1 & X + \alpha_{n-1} \end{vmatrix}.$$

Durch Wiederholung erhalten wir schließlich

$$\chi_{Z_\varphi}(X) = (-1)^{n-1} \begin{vmatrix} -1 & X & 0 & 0 & & \alpha_1 \\ 0 & -1 & X & 0 & & \alpha_2 \\ & \ddots & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & X & \alpha_{n-2} \\ \vdots & & & \ddots & -1 & X + \alpha_{n-1} \\ 0 & \cdots & & & 0 & \alpha_0 + X\alpha_1 + X^2\alpha_2 + \cdots + \alpha_{n-1}X^{n-1} + X^n \end{vmatrix}$$
$$= \varphi(X). \qquad \triangle$$

Übung 32. Geben sie einen alternativen Beweis von $\chi_{Z_\varphi}(X) = \varphi(X)$ durch Entwicklung nach der letzten Spalte (unter Verwendung von (4.3) und Übung 19).

Die Matrix Z_φ aus Beispiel 7.2.1 wird in diesem Abschnitt und auch später im Abschnitt 7.3.5 eine besondere Rolle spielen (siehe Beweis von Satz 7.2.12, Lemma 7.2.28, Theorem 7.2.29, Theorem 7.2.31), und hat einen Namen verdient.

Definition 7.2.2 (Begleitmatrix)**.** Für $\varphi \in \mathbb{K}[X]$ heißt die Matrix Z_φ aus Beispiel 7.2.1 auch *Begleitmatrix* von φ.

Beispiel 7.2.3. Die Begleitmatrix von $\varphi = (\alpha + X)$ ist $(-\alpha)$. $\triangle$

Beispiel 7.2.4. Die Begleitmatrix von $X^3 - 1$ ist

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \qquad \triangle$$

Das folgende Beispiel zeigt, dass das charakteristische Polynom χ_A im allgemeinen *nicht* verrät, ob A diagonalisierbar ist.

Beispiel 7.2.5.

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ und } B := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

haben das gleiche charakteristische Polynom

$$\chi_A(X) = \chi_B(X) = (X - 1)^2 ,$$

aber nur die erste Matrix ist diagonalisierbar (Übung 26). $\triangle$

7.2.2 Das Minimalpolynom

Ziel: Polynom für A, welches verrät, ob A diagonalisierbar.

Definition 7.2.6. Eine Teilmenge $\mathcal{I}$ eines kommutatives Ringes R heißt *Ideal*, wenn gilt

$$\varphi, \psi \in \mathcal{I} \Rightarrow \varphi + \psi \in \mathcal{I}$$
$$\varphi \in \mathcal{I}, \psi \in R \Rightarrow \varphi \cdot \psi \in \mathcal{I}$$

Beispiel 7.2.7. Sei V ein $\mathbb{K}$-Vektorraum. Wir können Polynome $\varphi \in \mathbb{K}[X]$ auswerten in $\mathrm{End}(V)$ (siehe Beispiel 4.2.14). Sei $f \in \mathrm{End}(V)$. Betrachten

$$\mathcal{I}_f := \{\varphi \in \mathbb{K}[X] \mid \underbrace{\varphi(f)}_{\text{Auswerten von } f} = \underbrace{\underline{0}}_{\in \mathrm{End}(V)}\}$$

Dann ist $\mathcal{I}_f$ ein Ideal, und heißt das *Ideal von* f. △

Bemerkung 7.2.8. Für jedes $f \in \mathrm{End}(V)$ gilt $\mathcal{I}_f \neq \{\underline{0}\}$. Denn: für $\dim(V) = n$ ist $\dim(\mathrm{End}(V)) = n^2$. Daher sind $1, f, f^2, f^3, \ldots, f^{n^2}$ linear abhängig. Es gibt also eine nicht-triviale Linearkombination von $\underline{0}$, d.h., es gibt $\alpha_0, \ldots, \alpha_{n^2} \in K$, die nicht alle 0 sind, so dass

$$\alpha_0 + \alpha_1 f + \alpha_2 f^2 + \cdots + \alpha_{n^2} f^{n^2} = \underline{0}.$$

Dann ist $\alpha_{n^2} X^{n^2} + \cdots + \alpha_1 X + \alpha_0$ ein Polynom in $\mathcal{I}_f \setminus \{\underline{0}\}$.

Der folgende Satz ist eine besondere Eigenschaft des Polynomrings $\mathbb{K}[X]$.

Satz 7.2.9. *Jedes Ideal $\mathcal{I} \subseteq \mathbb{K}[X]$ mit $\mathcal{I} \neq \{0\}$ enthält ein eindeutiges Polynom φ mit folgenden Eigenschaften:*

- *φ ist* normiert, *d.h., $\varphi = X^d + \cdots$ wobei $d = \mathrm{grad}(\varphi)$;*
- *Für jedes $\psi \in \mathcal{I}$ existiert $\psi_q \in \mathbb{K}[X]$ so dass $\psi = \varphi \cdot \psi_q$.*

φ heißt Minimalpolynom von $\mathcal{I}$, *im Falle von $\mathcal{I}_f$ auch* Minimalpolynom μ_f von f.

Beweis. Existenz: Sei d minimaler Grad eines Polynoms aus $\mathcal{I}$, und $\varphi \in \mathcal{I}$ vom Grad d und normiert. Für beliebiges $\psi \in \mathcal{I}$ dividieren wir durch φ mit Rest (Abschnitt 4.2.7):

$$\psi = \varphi \cdot \psi_q + \psi_r$$

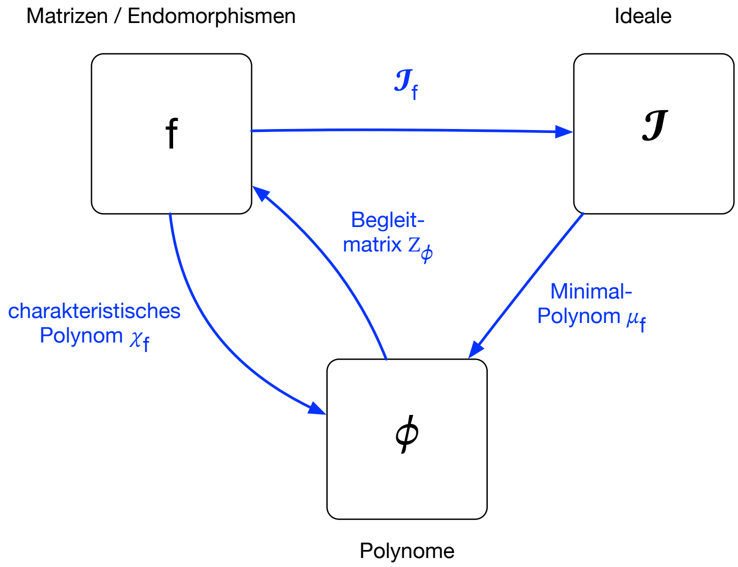
wobei $\psi_r = 0$ oder $\mathrm{grad}(\psi_r) < d$. Falls $\psi_r \neq 0$, dann wäre

$$\psi_r = \underbrace{\psi}_{\in \mathcal{I}} - \underbrace{\underbrace{\varphi}_{\in \mathcal{I}} \cdot \underbrace{\psi_q}_{\in \mathbb{K}[X]}}_{\in \mathcal{I}} \in \mathcal{I}$$

ein Widerspruch zur Minimalität von d.

Eindeutigkeit: Falls φ' ein anderes Polynom ist mit diesen Eigenschaften, dann teilen sich φ und φ' gegenseitig, was impliziert dass $\mathrm{grad}(\varphi) = \mathrm{grad}(\varphi')$. Sei $\psi \in \mathbb{K}[X]$ mit $\varphi = \varphi' \cdot \psi$. Dann ist ψ vom Grad 0. Da φ und φ' normiert sind, gilt $\psi = 1$ und $\varphi = \varphi'$. □

Bemerkung 7.2.10. Der Grad des Minimalpolynoms ist höchstens n^2: dies folgt unmittelbar aus Bemerkung 7.2.8. Wir werden später sehen, dass $\mathcal{I}_f \setminus \{0\}$ sogar ein Polynom vom Grad höchstens n enthält (dies folgt aus dem Satz von Cayley-Hamilton, Satz 7.2.12).



Definition 7.2.11. Sei $A \in \mathbb{K}^{n\times n}$ und $f := f_A$ die lineare Abbildung $x \mapsto Ax$. Schreiben

- $\mathcal{I}_A$ für $\mathcal{I}_f$.
- μ_A für μ_f.

Was ist der Zusammenhang zwischen Minimalpolynom von f und dem charakteristischen Polynom χ_f von f? Zunächst beweisen wir den folgenden wichtigen Satz. Wir erinnern uns: Polynome aus $\mathbb{K}[X]$ können ausgewertet werden im Matrizenring $\mathbb{K}^{n\times n}$ (Beispiel 4.2.13).

Satz 7.2.12 (Cayley-Hamilton)**.** *Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt*

$$\chi_A^{\mathbb{K}^{n\times n}}(A) = \mathbf{0} \in \mathbb{K}^{n\times n}$$

Also gilt für $f := f_A$

$$\chi_f \in \mathcal{I}_f \textit{ und } \mu_f | \chi_f.$$

'Jede Matrix erfüllt ihr eigenes charakteristisches Polynom.'

Bemerkung 7.2.13. Es folgt insbesondere, dass der Grad des Minimalpolynoms höchstens n ist, da der Grad des charakteristischen Polynoms offensichtlich höchstens n ist. Das verbessert die quadratische Schranke aus Bemerkung 7.2.8.

Bemerkung 7.2.14. Satz 7.2.12 hat die folgende Variante für lineare Abbildungen: sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, und $f \in \operatorname{End}(V)$. Dann gilt

$$\chi_f^{\operatorname{End} V}(f) = \underline{0}$$

wobei $\underline{0}\colon V \to V : v \mapsto \mathbf{0}$ die Nullabbildung.

Was ist faul an folgender Rechnung?

$$?? \qquad \chi_A^{\mathbb{K}^{n\times n}}(A) = \det(AE_n - A) = \det(\mathbf{0}) = 0 \qquad ¿¿$$

Beweis von Satz 7.2.12. Zu zeigen: für alle $v \in V$ gilt $\chi_f(f)(v) = \mathbf{0}$. Falls $v = \mathbf{0}$ ist die Aussage klar, da $\chi_f(f)$ eine lineare Abbildung ist.

Zwischenschritt: zeigen, dass es ein $U \leq V$ gibt mit

(a) $v \in U$

(b) U ist *(f-) invariant*, d.h., $\forall u \in U\colon f(u) \in U$.

(c) Für die Einschränkung $g := f|_U\colon U \to U$ (ergibt Sinn wegen (b)) gilt

$$\left(\chi_g^{\operatorname{End} V}(f)\right)(v) = \mathbf{0}.$$

Seien dazu

$$\begin{aligned} u_1 &:= v \\ u_2 &:= f(u_1) \\ u_3 &:= f(u_2) = f^2(v) \\ u_{i+1} &:= f(u_i) = f^i(v) \\ &\dots \end{aligned}$$

Es gibt maximal $n = \dim V$ viele linear unabhängige Vektoren, also gibt es ein $m \in \{1, \dots, n\}$ so dass

$$\begin{aligned} &u_1, \dots, u_m \text{ linear unabhängig} \\ &u_1, \dots, u_m, u_{m+1} \text{ linear abhängig} \end{aligned}$$

(hier verwenden wir, dass $u_1 = v \neq \mathbf{0}$). Das bedeutet, es gibt $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ mit

$$u_{m+1} = \alpha_1 u_1 + \cdots + \alpha_m u_m \tag{7.2}$$

Sei $U := \langle u_1, \dots, u_m \rangle$.
Randbemerkung: falls $m = n$ dann ist $U = V$ und $g = f$ und (c) impliziert die Aussage.

- U erfüllt (a). $v = u_1 \in U$.
- U erfüllt (b). Sei $u \in U$, $u = \sum_{i=1}^m \beta_i u_i$. Dann

$$f(u) \overset{f \text{ linear}}{=} \sum_{i=1}^m \beta_i f(u_i) = \sum_{i=1}^m \beta_i u_{i+1} \in U$$

weil $u_{m+1} \in U$ wegen (7.2).

- Bestimmung von χ_g. Definieren $B' := (u_1, \dots, u_m)$ (Basis von U).

$$M' := M_{B'}^{B'}(g) = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_1 \\ 1 & 0 & \cdots & 0 & \alpha_2 \\ 0 & 1 & \cdots & 0 & \alpha_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_m \end{pmatrix} \qquad \text{Merkregel! (3.14)}$$

denn: $g = f|_U$, d.h.,

$$\begin{aligned} g(u_i) &= f(u_i) = u_{i+1} && \text{für } i < m, \\ \text{und } g(u_m) &= u_{m+1} = \alpha_1 u_1 + \cdots + \alpha_m u_m. \end{aligned}$$

Die Matrix M' ist exakt die Begleitmatrix Z_φ des Polynoms

$$\varphi = X^m - \alpha_m X^{m-1} - \cdots - \alpha_2 X - \alpha_1.$$

Also (siehe Beispiel 7.2.1)

$$\chi_g(X) = \chi_{M'}(X) = \det(XE - M') = \varphi. \tag{7.3}$$

- g erfüllt (c):

$$\begin{aligned} &(\chi_g(f))(v) \\ &= f^m(v) - \alpha_m f^{m-1}(v) - \cdots - \alpha_2 f(v) - \alpha_1 v && (7.3) \\ &= u_{m+1} - \alpha_m u_m - \cdots - \alpha_2 u_2 - \alpha_1 u_1 && (\text{Def. von } u_1, \dots, u_{m+1}) \\ &= \mathbf{0} && (7.2) \end{aligned}$$

Wir zeigen nun $\chi_f(f)(v) = \mathbf{0}$.
Sei $B = (u_1, \dots, u_m, w_{m+1}, \dots, w_n)$ Ergänzung von B' zu Basis von V (existiert nach Austauschsatz von Steinitz, Satz 2.4.13). Dann hat $M_B^B(f)$ die Form

$$M := M_B^B(f) = \begin{pmatrix} M' & * \\ \mathbf{0} & M'' \end{pmatrix}$$

wobei $M' \in \mathbb{K}^{n \times n} = M_{B'}^{B'}(g)$ wie eben. Denn für $i \leq m$ ist $f(u_i) \in U$ und daher

$$f(u_i) = \sum_{j=1}^{m} a_j u_j + 0 \cdot w_{m+1} + \cdots + 0 \cdot w_n\,.$$

Also

$$\begin{aligned} \chi_f = \chi_M &= \chi_{M'} \cdot \chi_{M''} && \text{(Übung 24)} \\ &= \chi_g \cdot \chi_{M''} = \chi_{M''} \cdot \chi_g && \text{(Polynommultiplikation ist kommutativ).} \end{aligned}$$

Einsetzen von f (Auswerten in $\operatorname{End} V$):

$$\begin{aligned}
\chi_f^{\operatorname{End} V}(f)(v) &= \big((\chi_{M''}\chi_g)(f)\big)(v) && \text{(siehe oben)}\\
&= \big(\chi_{M''}(f) \circ \chi_g(f)\big)(v) && \text{(Satz 4.2.12)}\\
&= \chi_{M''}(f)\big(\underbrace{\chi_g(f)(v)}_{=\mathbf{0}}\big) && \text{(Def. Multiplikation in End } V\text{)}\\
&= \chi_{M''}(f)(\mathbf{0}) && \text{(wegen (c))}\\
&= \mathbf{0} && \text{(denn } \chi_{M''}(f) \text{ ist lineare Abbildung).} \quad \square
\end{aligned}$$

Beispiel 7.2.15. Das Minimalpolynom von $f = \mathrm{id}_V$ ist $\mu_f(X) = (X-1)$, also verschieden vom charakteristischen Polynom $\chi_f = (X-1)^n$ für $n = \dim V$. △

Beispiel 7.2.16. Eine Abbildung $f \in \operatorname{End}(V)$ heißt *Involution* falls $f^2 = \mathrm{id}_V$. Den Fall $f \in \{\mathrm{id}_V, -\mathrm{id}_V\}$ haben wir bereits im vorigen Beispiel behandelt. Ansonsten ist das Minimalpolynom von f

$$\mu_f(X) = X^2 - 1 .$$

Für Eigenwerte λ einer Involution gilt $\lambda^2 = 1$: Nullstellen von $\mu_f(X)$!
Behauptung: f ist diagonalisierbar, d.h., es gibt eine Basis von V aus Eigenvektoren von f (Lemma 4.3.15 und Definition 4.3.17). Es gilt sogar

$$V = \operatorname{Kern}(f - \mathrm{id}_V) \oplus \operatorname{Kern}(f + \mathrm{id}_V).$$

Denn:

- $\operatorname{Kern}(f - \mathrm{id}_V)$ und $\operatorname{Kern}(f + \mathrm{id}_V)$ sind Eigenräume der Eigenwerte 1 und -1.
- $\operatorname{Kern}(f - \mathrm{id}_V) \cap \operatorname{Kern}(f + \mathrm{id}_V) = \{\mathbf{0}\}$: falls $v \in \operatorname{Kern}(f - \mathrm{id}_V) \cap \operatorname{Kern}(f + \mathrm{id}_V)$, dann gilt $-v = f(v) = v$, also $v = \mathbf{0}$.
- $\operatorname{Kern}(f - \mathrm{id}_V) + \operatorname{Kern}(f + \mathrm{id}_V) = V$: jedes $v \in V$ kann geschrieben werden als

$$v = \underbrace{\frac{v - f(v)}{2}}_{\in \operatorname{Kern}(f + \mathrm{id}_V)} + \underbrace{\frac{v + f(v)}{2}}_{\in \operatorname{Kern}(f - \mathrm{id}_V)}$$

Denn: $(f - \mathrm{id}_V)(v + f(v)) = f(v + f(v)) - v - f(v) = f(v) + v - f(v) = 0$, also $v + f(v) \in \operatorname{Kern}(f - \mathrm{id}_V)$.
Analog: $v - f(v) \in \operatorname{Kern}(f + \mathrm{id}_V)$. △

Das Minimalpolynom gibt uns das meiste von dem, was wir typischerweise vom charakteristischen Polynom bekommen.

Proposition 7.2.17. *Sei V endlichdimensional und $f \in \operatorname{End}(V)$. Dann gelten:*

- *$\lambda \in \mathbb{K}$ ist genau dann Eigenwert von f, wenn $\mu_f(\lambda) = 0$.*
- *f ist genau dann invertierbar wenn $\mu_f(0) \neq 0$.*

Beweis. Es sei
$$\mu_f = X^d + \alpha_{d-1}X^{d-1} + \cdots + \alpha_1 X + \alpha_0$$

Teil 1: Sei $\lambda \in \mathbb{K}$ ein EW von f, und $v \in V \setminus \{\mathbf{0}\}$ ein zugehöriger Eigenvektor, $f(v) = \lambda v$. Dann gilt wegen $f^i(v) = \lambda^i v$
$$\begin{aligned} 0 = \mu_f(f)(v) &= (f^d + \alpha_{d-1}f^{d-1} + \cdots + \alpha_1 f + \alpha_0 \operatorname{id}_V)(v) \\ &= \lambda^d v + \alpha_{d-1}\lambda^{d-1}v + \cdots + \alpha_1\lambda v + \alpha_0 v = \mu_f(\lambda)v \end{aligned}$$
also $\mu_f(\lambda) = 0$ da $v \neq \mathbf{0}$. Umgekehrt, falls $\mu_f(\lambda) = 0$, dann ist $\chi_f(\lambda) = 0$ (Lemma 4.2.16), da $\mu_f | \chi_f$ nach dem Satz von Cayley-Hamilton (Satz 7.2.12). Also ist λ ein Eigenwert von f (Satz 4.3.10).

Teil 2: Wenn $\mu_f(0) = \alpha_0 \neq 0$, dann können wir $\mu_f(f) = \underline{0}$ umschreiben zu
$$1 = f(f^{d-1} + \alpha_{d-1}f^{d-2} + \cdots + \alpha_1)/-\alpha_0$$
also gilt
$$f^{-1} = \frac{-1}{\alpha_0}(f^{d-1} + \alpha_{d-1}f^{d-2} + \cdots + \alpha_1 \operatorname{id}_V)\,.$$

Umgekehrt, wenn $\mu_f(0) = 0$, dann ist 0 ein Eigenwert nach Teil 1. Also gibt es ein $v \in V \setminus \{\mathbf{0}\}$ mit $f(v) = 0$, und f ist nicht injektiv, damit nicht invertierbar. □

Teil 1 von Proposition 7.2.17 in Kombination mit Satz 4.3.10 ergibt direkt die folgende Aussage.

Korollar 7.2.18. *Für V endlichdimensional und $f \in \operatorname{End}(V)$ haben χ_f und μ_f dieselben Nullstellen.*

In anderen Worten: die Polynome χ_f und μ_f haben dieselben Faktoren der Gestalt $(X - \lambda)$, für $\lambda \in \mathbb{K}$ (Lemma 4.2.16). Diese Aussage werden wir später auf Faktoren allgemeinerer Gestalt erweitern (Lemma 7.2.37).

Im Gegensatz zum charakteristischen Polynom kann die Berechnung des Minimalpolynoms von A aufwendig sein. Allerdings hilft der Satz von Cayley-Hamilton (Satz 7.2.12), da man nicht mehr alle Polynome testen muss, sondern nur noch die Teiler von χ_A. Wir demonstrieren das in den Beweisen der folgenden Propositionen.

Proposition 7.2.19. *Die Begleitmatrix eines normierten Polynoms $\varphi \in \mathbb{K}[X]$ hat das Minimalpolynom φ, d.h., $\mu_{Z_\varphi} = \varphi$.*

Beweis. Sei $\varphi = X^n + \alpha_{n-1}X^{n-1} + \cdots + \alpha_1 X + \alpha_0$. Wir kennen bereits das charakteristische Polynom von Z_φ, es gilt nämlich $\chi_{Z_\varphi} = \varphi$ (Beispiel 7.2.1). Da μ_{Z_φ} normiert ist und χ_{Z_φ} teilt (Satz 7.2.12), genügt es zu zeigen, dass der Grad von μ_{Z_φ} gleich n ist. Offenbar gilt für
$$Z_\varphi = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{pmatrix}$$

dass

$$\begin{aligned} Z_\varphi e_1 &= e_2 \\ Z_\varphi e_2 &= e_3 = Z_\varphi^3 e_1 \\ \vdots\ &\quad \vdots \\ Z_\varphi e_{n-1} &= e_n = Z_\varphi^{n-1} e_1 . \end{aligned}$$

Angenommen, es gäbe in $\mathcal{I}_{Z_\varphi}$ ein normiertes Polynom

$$\psi = X^m + \beta_{m-1}X^{m-1} + \dots + \beta_1 X + \beta_0$$

vom Grad $m < n$, dann wäre also

$$\begin{aligned} \mathbf{0} = \underbrace{\psi(Z_\varphi)}_{\in \mathbb{K}^{n\times n}} e_1 &= Z_\varphi^m e_1 + \beta_{m-1} Z_\varphi^{m-1} e_1 + \dots + \beta_1 Z_\varphi e_1 + \beta_0 e_1 \\ &= e_{m+1} + \beta_{m-1} e_m + \dots + \beta_1 e_2 + \beta_0 e_1 . \end{aligned}$$

und damit wären die Basisvektoren $e_1, \dots, e_{m+1}$ linear abhängig, ein Widerspruch. □

Proposition 7.2.20. *Das Minimalpolynom von*

$$A := \begin{pmatrix} \lambda_1 & 1 & & 0 \\ 0 & \lambda_2 & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_n \end{pmatrix}$$

ist

$$\mu_A(X) = (X - \lambda_1)\cdots(X - \lambda_n) = \chi_A(X) .$$

Beweis. Es ist klar (siehe (4.3)), dass

$$\chi_A(X) = (X - \lambda_1)\cdots(X - \lambda_n).$$

Da μ_A ein Teiler ist von χ_A, genügt es zu zeigen, dass für jedes $i \in \{1, \dots, n\}$ das Polynom $\varphi_i(X) := \chi_A(X)/(X - \lambda_i)$ nicht in $\mathcal{I}_f$ liegt. Für $i = n$ berechnen wir

$$(A - \lambda_1 E_n)(A - \lambda_2 E_n)\cdots(A - \lambda_{n-1}E_n)$$

$$= \begin{pmatrix} 0 & 1 & 0 & \\ 0 & \lambda_2 - \lambda_1 & 1 & \\ 0 & 0 & \lambda_3 - \lambda_1 & \\ & & & \ddots \end{pmatrix} \begin{pmatrix} \lambda_1 - \lambda_2 & 1 & 0 & \\ 0 & 0 & 1 & \\ 0 & 0 & \lambda_3 - \lambda_2 & \\ & & & \ddots \end{pmatrix} \cdots \begin{pmatrix} \ddots & & & \\ & \lambda_{n-2} - \lambda_{n-1} & 1 & 0 \\ & 0 & 0 & 1 \\ & 0 & 0 & \lambda_n - \lambda_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 1 & \\ 0 & 0 & * & \\ 0 & 0 & * & \\ & & & \ddots \end{pmatrix} \cdots \begin{pmatrix} \ddots & & & \\ & \lambda_{n-2} - \lambda_{n-1} & 1 & 0 \\ & 0 & 0 & 1 \\ & 0 & 0 & \lambda_n - \lambda_{n-1} \end{pmatrix}$$

$$= \dots = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & * \\ \vdots & \vdots & & \vdots & * \\ 0 & 0 & \cdots & 0 & * \end{pmatrix} \neq \mathbf{0}$$

Die Aussage folgt für alle $i \in \{1, \dots, n\}$ durch Umsortierung der Faktoren von $\varphi_i(X)$. $\square$

Übung 33. Sei $A = \begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}$. Zeigen Sie, dass $\mu_A = \mathrm{kgV}(\mu_B, \mu_C)$.

Übung 34. Sei $A \in \mathbb{K}^{n\times n}$. Zeigen Sie, dass $\mu_A = \mu_{A^\top}$.

Übung 35. Was ist das Minimalpolynom der Nullabbildung $\underline{0}$? Was ist das Minimalpolynom der identischen Abbildung?

Übung 36. Sei $f \in \mathrm{End}(V)$ so, dass μ_f ein konstantes Polynom ist. Zeigen Sie, dass dann $V = \{\mathbf{0}\}$.

7.2.3 Minimalpolynom und Diagonalisierbarkeit

Wiederholung (Abschnitt 4.3.4): Wie entscheiden wir, ob eine Matrix $A \in \mathbb{K}^{n\times n}$ diagonalisierbar ist?

1. Berechne die Eigenwerte von A.
2. Berechne die zugehörigen Eigenräume.
3. Entscheide, ob man eine Basis aus Eigenvektoren finden kann.

Insbesondere: wenn $\chi_A = (X - \lambda_1)\cdots(X - \lambda_n)$ für paarweise verschiedene $\lambda_1, \dots, \lambda_n$, dann ist A diagonalisierbar (Bemerkung 4.3.20). Wenn wir hier das charakteristische Polynom durch das Minimalpolynom ersetzen, erhalten wir sogar ein hinreichendes *und* notwendiges Kriterium für Diagonalisierbarkeit!

Satz 7.2.21 (Minimalpolynom and Diagonalisierbarkeit). *Sei $A \in \mathbb{K}^{n\times n}$ und seien $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von A. Dann sind äquivalent:*

(1) *A ist diagonalisierbar;*

(2) *$(A - \lambda_1 E_n)\cdots(A - \lambda_k E_n) = \mathbf{0}$;*

(3) *$\mu_A(X) = (X - \lambda_1)\cdots(X - \lambda_k)$.*

Beweis. (1) $\Rightarrow$ (2). Definiere $\varphi(X) := (X - \lambda_1)\cdots(X - \lambda_k) \in \mathbb{K}[X]$ und $\varphi_i = \varphi/(X - \lambda_i)$. Es ist zu zeigen, dass $\varphi(A) = \mathbf{0}$. Sei $v \in \mathbb{K}^n$ beliebig. Wir zeigen, dass $\varphi(A)v = \mathbf{0}$. Da A per Annahme diagonalisierbar ist, gibt es nach Satz 4.3.19 (3c) eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A. Also kann man v schreiben als Summe $u_1 + \cdots + u_k$ (falls $v = \mathbf{0}$ ist $k = 0$) wobei für alle $i \in \{1, \dots, k\}$ der Vektor u_i ein Eigenvektor zum Eigenwert λ_i ist.

$$\begin{aligned}
\varphi(A)v &= \varphi(A)(u_1 + \cdots + u_k) \\
&= \varphi(A)u_1 + \cdots + \varphi(A)u_k \\
&= \varphi_1(X - \lambda_1)(A)u_1 + \cdots + \varphi_k(X - \lambda_k)(A)u_k \\
&= \varphi_1(A - \lambda_1 E_n)u_1 + \cdots + \varphi_k(A - \lambda_k E_n)u_k \\
&= \varphi_1(A)(Au_1 - \lambda_k u_1) + \cdots + \varphi_k(A)(Au_k - \lambda_k u_k) = \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}.
\end{aligned}$$

(2) $\Rightarrow$ (3): Nach (2) ist $\varphi \in \mathcal{I}_A$ und damit $\mu_A | \varphi$. Umgekehrt gilt $\varphi | \mu_A$: zeigen dazu, dass jeder Faktor $(X - \lambda_i)$ von φ ein Teiler von μ_A ist. Dafür genügt es zu zeigen, dass λ_i eine Nullstelle ist von μ_A (Lemma 4.2.16). Satz 4.3.10 impliziert, dass λ_i eine Nullstelle ist von χ_A, also auch eine von μ_A nach Korollar 7.2.18. Da sowohl μ_A also auch φ normiert sind, muss also gelten $\mu_A = \varphi$.

(3) $\Rightarrow$ (1): Beweis per Induktion nach n. Für $n = 1 = k$ ist A sicher diagonalisierbar. Sei nun $n > 1$, und sei die Aussage richtig für alle $m < n$.

Behauptung: für $f := f_A$ und $V := \mathbb{K}^n$ sind $\mathrm{Kern}(f - \lambda_1 \,\mathrm{id}_V)$ und $\mathrm{Bild}(f - \lambda_1 \,\mathrm{id}_V)$ komplementär, d.h. (Definition 2.4.17)

$$\mathrm{Kern}(f - \lambda_1 \,\mathrm{id}_V) \oplus \mathrm{Bild}(f - \lambda_1 \,\mathrm{id}_V) = V.$$

Wegen der Dimensionsformel $\dim \mathrm{Bild} + \dim \mathrm{Kern} = n$ (3.3.6) genügt es zu zeigen, dass

$$\mathrm{Kern}(f - \lambda_1 \,\mathrm{id}_V) + \mathrm{Bild}(f - \lambda_1 \,\mathrm{id}_V) = V\,.$$

Dividieren $(X - \lambda_2)\cdots(X - \lambda_k)$ mit Rest durch $(X - \lambda_1)$, und erhalten $\varphi, \psi \in \mathbb{K}[X]$, $\mathrm{grad}(\psi) < 1$, so dass

$$(X - \lambda_2)\cdots(X - \lambda_k) = \varphi(X - \lambda_1) + \psi.$$

Da $\mathrm{grad}(\psi) < 1$, ist ψ ein Körperelement; weiterhin ist $\psi \neq 0$, da $(X - \lambda_1)$ für $i \in \{2, \dots, k\}$ kein Teiler von $(X - \lambda_i)$ ist. Setzen f ein, stellen um, und erhalten

$$(f - \lambda_2 \,\mathrm{id}_V)\cdots(f - \lambda_m \,\mathrm{id}_V) - (f - \lambda_1 \,\mathrm{id}_V)\varphi(f) = \psi \,\mathrm{id}_V\,. \tag{7.4}$$

Sei nun $v \in V$ beliebig. Dann folgt aus (7.4), dass

$$\psi \,\mathrm{id}_V(v) = \underbrace{(f - \lambda_2 \,\mathrm{id}_V)\cdots(f - \lambda_m \,\mathrm{id}_V)(v)}_{=:v_1} - \underbrace{(f - \lambda_1 \,\mathrm{id}_V)\varphi(f)(v)}_{=:v_2}.$$

Also $\psi \cdot v = v_1 + v_2$. Wir haben $(f - \lambda_1 \,\mathrm{id}_V)(v_1) = \mathbf{0}$, da $\mu_f(f) = \underline{0}$ nach Satz 7.2.12. Also $v_1 \in \mathrm{Kern}(f - \lambda_1 \,\mathrm{id}_V)$. Ausserdem haben wir

$$v_2 = (f - \lambda_1 \,\mathrm{id}_V)\varphi(f)(v) \in \mathrm{Bild}(f - \lambda_1 \,\mathrm{id}_V).$$

Weil $\frac{1}{\psi}v_1 + \frac{1}{\psi}v_2 = v$, folgt die Behauptung.

Da $\dim(\mathrm{Kern}(f - \lambda_1 \,\mathrm{id}_V)) > 0$, gilt $\dim(\mathrm{Bild}(f - \lambda_1 \,\mathrm{id})) < n$. Wenden Induktionsvoraussetzung an auf $U := \mathrm{Bild}(f - \lambda_1 \,\mathrm{id}) \leq V$. Da $f(U) \subseteq U$, ist die Einschränkung f_U von f auf U aus $\mathrm{End}(U)$. Ferner ist χ_{f_U} ein Teiler von μ_f. Also zerfällt χ_{f_U} in paarweise verschiedene Linearfaktoren, und f_U ist diagonalisierbar nach Induktionsvoraussetzung.

Sei $v_1, \dots, v_m$ eine Basis von U aus Eigenwerten von f_U, und sei $v_{m+1}, \dots, v_n$ eine Basis von $\mathrm{Kern}(f - \lambda_1 \,\mathrm{id}) = \mathrm{Eig}_{\lambda_1}(f)$. Da $U \cap \mathrm{Kern}(f - \lambda_1 \,\mathrm{id}) = \{\mathbf{0}\}$, ist $v_1, \dots, v_m, v_{m+1}, \dots, v_n$ eine Basis von V aus Eigenvektoren von f. Die Aussage folgt nun aus dem ersten Diagonalisierbarkeitskriterium (Satz 4.3.19). □

Und was, wenn f nicht diagonalisierbar ist?

7.2.4 Zyklische Unterräume

Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Die folgende Definition extrahiert wichtige Ideen aus dem Beweis des Satzes von Cayley-Hamilton (Satz 7.2.12).

Definition 7.2.22. Sei $f \in \mathrm{End}(V)$. Dann heißt $S \leq V$ *invariant unter* f (oder f-invariant) falls

$$f(S) \subseteq S\,.$$

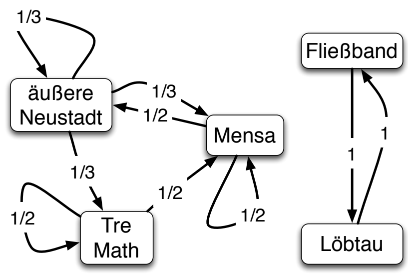
Bemerkung 7.2.23. $f|_S \in \mathrm{End}(S)$.

Bemerkung 7.2.24. Für alle $v \in V$ gilt

$$\begin{aligned} \langle v \rangle \text{ ist } f\text{-invariant} &\iff (v \text{ ist Eigenvektor von } f \text{ oder } v = \mathbf{0}) \\ &\iff f(v) \in \langle v \rangle. \end{aligned}$$

Bemerkung 7.2.25. Die Eigenräume von f (Definition 4.3.3) sind invariant.

Beispiel 7.2.26. Wir betrachten wieder das folgende System (Beispiel 4.3.27).



Beschreibung durch Matrix:

$$\begin{array}{cc} \begin{pmatrix} 1/3 & 1/2 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/2 & 0 & 0 \\ 1/3 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} & \begin{array}{c} \text{äußere Neustadt} \\ \text{Mensa} \\ \text{Tre Math} \\ \text{Fließband} \\ \text{Löbtau} \end{array} \end{array}$$

Dann ist $\langle (0.3, 0.4, 0.3, 0, 0) \rangle$ invariant. Weiterhin ist $\langle (0, 0, 0, 0.5, 0.5) \rangle$ invariant. △

Ziel dieses Abschnittes: Dekomposition

$$V = S_1 \oplus \cdots \oplus S_k$$

für invariante $S_i \leq V$, so dass die Matrixdarstellung von $f_i := f|_{S_i}$ durch $\mu_{f_i}(X)$ eindeutig bestimmt.

Definition 7.2.27. Für $v \in S \setminus \{\mathbf{0}\}$, definiere

$$Z_v := \langle v, f(v), f^2(v), \dots \rangle$$

der *(von v erzeugte) zyklische Unterraum* von V.

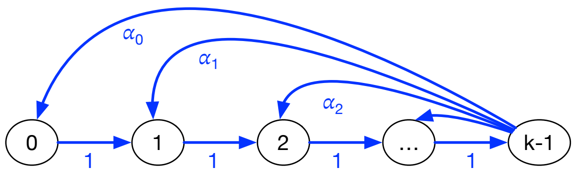


Abbildung 7.2: Illustration der Matrix in (7.5). Zur Erklärung der Bedeutung der Illustrationen siehe Beispiel 4.3.27.

Da $\dim(V) = n$ existiert $k \in \{1, \dots, n\}$ mit

$$f^k(v) \in \langle v, f(v), f^2, \dots, f^{k-1}(v)\rangle \,.$$

Lemma 7.2.28. *Sei $v \in V \setminus \{\mathbf{0}\}$, und sei $k \in \mathbb{N}$ kleinstmöglich, so dass $f^k(v) \in \langle v, f(v), f^2, \dots, f^{k-1}(v)\rangle$. Dann ist Z_v invariant unter f, und*

$$B = (v, f(v), \dots, f^{k-1}(v))$$

ist Basis für Z_v. Ausserdem gilt (siehe Bild 7.2)

$$M_B^B(f|_{Z_v}) = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_0 \\ 1 & 0 & \cdots & 0 & \alpha_1 \\ 0 & 1 & \cdots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_{k-1} \end{pmatrix} = Z_\varphi \tag{7.5}$$

wobei

$$\begin{aligned} & f^k(v) = \alpha_0 v + \alpha_1 f(v) + \dots + \alpha_{k-1} f^{k-1}(v) \\ \textit{und} \quad & \varphi(X) = X^k - \alpha_{k-1} X^{k-1} - \dots - \alpha_1 X - \alpha_0. \end{aligned}$$

Beweis. Die Vektoren $v, f(v), \dots, f^{k-1}(v)$ sind linear unabhängig, da k kleinstmöglich gewählt. Zu zeigen ist, dass

$$\langle B\rangle = Z_v = \langle v, f(v), f^2(v), \dots\rangle.$$

Sei $m \in \mathbb{N}$ beliebig. Zeigen per Induktion nach m, dass $f^m(v) \in \langle B\rangle$. Falls $m < k$, dann $v^m(v) \in B$. Angenommen $f^{m-1}(v) = \beta_0 v + \beta_1 f(v) + \dots + \beta_{k-1} f^{k-1}(v)$. Dann ist

$$f^m(v) = \beta_0 f(v) + \beta_1 f^2(v) + \dots + \beta_{k-1} \underbrace{f^k(v)}_{\in\langle B\rangle} \in \langle B\rangle \,.$$

Sei nun $u \in Z_v$ beliebig. Dann gilt

$$u = \gamma_0 \underbrace{v}_{\in\langle B\rangle} + \gamma_1 \underbrace{f(v)}_{\in\langle B\rangle} + \dots + \gamma_{k-1} \underbrace{f^{k-1}(v)}_{\in\langle B\rangle} \in \langle B\rangle$$

Die Matrixdarstellung von $f|_{Z_v}$ bezüglich B ist

$$\begin{aligned}&\begin{pmatrix} f(v) & f(f(v)) & \cdots & f(f^{k-2}(v)) & f(f^{k-1}(v)) \end{pmatrix}\\ &= \begin{pmatrix} v & f(v) & \cdots & f^{k-2}(v) & f^{k-1}(v) \end{pmatrix} \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_0 \\ 1 & 0 & \cdots & 0 & \alpha_1 \\ 0 & 1 & \cdots & 0 & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_{k-1} \end{pmatrix}\end{aligned}$$

(Merkregel, (3.14)! Spalten sind Koordinaten der Bilder der Einheitsvektoren.) □

Satz 7.2.29 (Zerlegung in zyklische Unterräume). *Sei V n-dimensional und $f \in \mathrm{End}(V)$. Dann gibt es $v_1, \ldots, v_k \in V$ so dass*

$$V = Z_{v_1} \oplus \cdots \oplus Z_{v_k}$$

und V hat eine Basis B so dass

$$M_B^B(f) = \begin{pmatrix} Z_{\varphi_1} & 0 & \cdots & 0 \\ 0 & Z_{\varphi_2} & & \\ \vdots & & \ddots & \\ 0 & & & Z_{\varphi_k} \end{pmatrix}.$$

für normierte $\varphi_1, \ldots, \varphi_k \in \mathbb{K}[X]$.

Beweis. Per Induktion nach n. Sei $v \in V$ so, dass $m = \dim Z_v$ größtmöglich. Falls $m = n$ ist nichts zu zeigen. Betrachte $g\colon V \to \mathbb{K}^m$ definiert durch

$$g(u) := \begin{pmatrix} h(u) \\ h(f(u)) \\ \vdots \\ h(f^{m-1}(u)) \end{pmatrix}$$

wobei $h\colon V \to \mathbb{K}$ Linearform mit $h(v) = \cdots = h(f^{m-2}(v)) = 0$ und $h(f^{m-1}(v)) = 1$. ($v, \ldots, f^{m-1}(v)$ sind linear unabhängig.)

Behauptung 1: $g|_{Z_v}\colon Z_v \to \mathbb{K}^m$ ist Isomorphismus. Die Darstellungsmatrix dieser Abbildung bezüglich der Basis $B = (v_1, f(v), \ldots, f^{m-1}(v))$ von Z_v und der Standardbasis von $\mathbb{K}^m$ ist

$$A = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & 1 & * \\ 0 & \iddots & \iddots & \vdots \\ 1 & * & \cdots & * \end{pmatrix}$$

also offensichtlich invertierbar.

Behauptung 2: $\mathrm{Kern}(g)$ ist f-invariant. Sei $u \in \mathrm{Kern}(g)$, also

$$g(u) = \begin{pmatrix} h(u) \\ h(f(u)) \\ \vdots \\ h(f^{m-1}(u)) \end{pmatrix} = \mathbf{0}\,.$$

Dann ist

$$g(f(u)) = \begin{pmatrix} h(f(u)) \\ h(f^2(u)) \\ \vdots \\ h(f^{m-1}(u)) \\ h(f^m(u)) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ h(f^m(u)) \end{pmatrix}$$

Da $f^m(u) \in \langle u, f(u), \ldots, f^{m-1}(u) \rangle$ für alle u ist $h(f^m(u)) = 0$. Also $f(u) \in \mathrm{Kern}(g)$ wie behauptet.

Behauptung 3: $V = Z_v \oplus \mathrm{Kern}(g)$. Es gilt $Z_v \cap \mathrm{Kern}(f) = \{\mathbf{0}\}$ da $g|_{Z_v}$ injektiv. Ausserdem

$$\dim Z_v + \dim \mathrm{Kern}(g) = \dim \mathrm{Bild}(g) + \dim \mathrm{Kern}(g) = \dim(V)$$

also $V = Z_v + \mathrm{Kern}(g)$.

Wenden nun die Induktionsvoraussetzung an auf die Einschränkung von f auf $\mathrm{Kern}(g)$, und der Satz folgt aus Lemma 7.2.28. □

Beispiel 7.2.30. Die Darstellung von f aus Satz 7.2.29 ist nicht eindeutig: die Matrix

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = Z_{X^2-1}$$

ist bereits von der Form in Satz 7.2.29. Auf der anderen Seite gilt für die Basis $B = (e_2, e_1)$ von $\mathbb{R}^2$, dass

$$M_B^B(f_A) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} Z_{X-1} & 0 \\ 0 & Z_{X+1} \end{pmatrix}.$$ △

7.2.5 Die Frobenius-Normalform

Die Frobenius-Normalform (bisweilen auch *rationale Normalform*) liefert eine Klassifikation von quadratischen Matrizen bis auf Ähnlichkeit.

Satz 7.2.31 (Frobenius-Normalform)**.** *Sei V n-dimensional und $f \in \mathrm{End}(V)$. Dann hat V eine Basis B so dass*

$$M_B^B(f) = \begin{pmatrix} Z_{\varphi_1} & 0 & \cdots & 0 \\ 0 & Z_{\varphi_2} & & \\ \vdots & & \ddots & \\ 0 & & & Z_{\varphi_k} \end{pmatrix}.$$

mit der Eigenschaft dass $\varphi_i | \varphi_{i-1}$ für alle $i \in \{2, \dots, k\}$; die Polynome $\varphi_1, \dots, \varphi_k$ sind hierbei eindeutig, und $\varphi_1 = \mu(f)$.

Die Polynome $\varphi_1, \dots, \varphi_k$ heißen auch *Ähnlichkeitsinvarianten* (denn ähnliche Matrizen haben die gleichen Ähnlichkeitsinvarianten), und die Untermatrizen $Z_{\varphi_1}, \dots, Z_{\varphi_k}$ die *Kästchen* der Frobenius-Normalform.

Beweis. Wie im Beweis der Zyklendekomposition (Satz 7.2.29) wählen wir $v_1 \in V$ so, dass $m = \dim Z_{v_1}$ größtmöglich.[1] Mit dieser Wahl gilt

$$f^m(v_1) = \alpha_0 v_1 + \alpha_1 f(v_1) + \dots + \alpha_{m-1} f^{m-1}(v_1) \tag{7.6}$$

für $\alpha_0, \dots, \alpha_{k-1} \in \mathbb{K}$. Sei

$$\varphi_1(X) := X^m - \alpha_{m-1} X^{m-1} - \dots - \alpha_1 X - \alpha_0$$

Behauptung 1: $\varphi_1(f) = \underline{0}$, d.h., $\varphi_1(f)(u) = 0$ für alle $u \in V$. Für $u = v_1$ stimmt das wegen (7.6). Für $u = f(v_1)$ haben wir

$$\begin{aligned} & f^m(f(v_1)) - \alpha_{m-1} f^{m-1}(f(v_1)) - \dots - \alpha_1 f(f(v)) - \alpha_0 f(v_1) \\ = {} & f(f^m(v_1) - \alpha_{m-1} f^{m-1}(v_1) - \dots - \alpha_1 f(v_1) - \alpha_0 v_1) \\ = {} & f(0) = 0 \end{aligned}$$

Analog für $u = f^2(v_1), \dots, u = f^{m-1}(v_1)$, und die Aussage folgt für $u \in Z_{v_1}$.
Aus der Zyklendekomposition (Satz 7.2.29) folgt, dass $V = Z_{v_1} \oplus W$. Sei nun $u \in W$.
Da m größtmöglich, gilt

$$\begin{aligned} f^m(v_1 + u) &= \gamma_{m-1} f^{m-1}(v_1 + u) + \dots + \gamma_0(v_1 + u) \\ &= \underbrace{\underbrace{\gamma_{m-1} f^{m-1}(v_1) + \dots + \gamma_0(v_1)}_{\in Z_{v_1}}}_{= f^m(v_1)} + \underbrace{\underbrace{\gamma_{m-1} f^{m-1}(w) + \dots + \gamma_0(u)}_{\in W}}_{= f^m(u)}. \end{aligned}$$

Also gilt insbesondere

$$\gamma_{m-1} f^{m-1}(v_1) + \dots + \gamma_0(v_1) = f^m(v_1) = \alpha_{m-1} f^{m-1}(v_1) + \dots + \alpha_0(v_1)$$

und da $v_1, f(v_1), \dots, f^{m-1}(v_1)$ linear unabhängig, gilt $\alpha_i = \gamma_i$ für $i \in \{0, \dots, m-1\}$.
Also

$$\begin{aligned} \varphi_1(f)(u) &= \varphi_1(f)(v_1) + \varphi_1(f)(u) \\ &= \varphi_1(f)(v_1 + u) && \text{(Linearität von } \varphi_1(f)) \\ &= f^m(v_1 + u) - \alpha_{m-1} f^{m-1}(v_1 + u) - \alpha_0(v_1 + u) && \text{(Definition von } \varphi_1) \\ &= \gamma_{m-1} f^{m-1}(v_1 + u) + \dots + \gamma_0(v_1 + u) && \text{(Definition von } \gamma_0, \dots, \gamma_{m-1}) \\ &\quad - \alpha_{m-1} f^{m-1}(v_1 + u) - \alpha_0(v_1 + u) \\ &= 0 && \text{(da } \alpha_i = \gamma_i). \end{aligned}$$

[1] Es ist nicht unmittelbar klar, wie dieser Schritt algorithmisch durchgeführt werden kann. Ein effizientes Verfahren zur Berechnung der Frobenius-Normalform wird in Abschnitt 7.3.6 vorgestellt.

Behauptung 2: $\varphi_1 = \mu_f$ (und ist damit eindeutig). Wir wissen bereits, dass $\varphi_1(f) = 0$, also $\mu_f | \varphi_1$. Auf der anderen Seite sind $v_1, f(v_1), \dots, f^{m-1}(v_1)$ linear unabhängig, also $\operatorname{grad}(\mu_f) \geq m$. Da φ_1 und μ_f normiert gilt $\varphi_1 = \mu_f$.

Wählen nun $v_2 \in W$ und $\varphi_2(X)$ auf die gleiche Art wie v_1 und $\varphi_1(X)$.

Behauptung 3: φ_2 teilt φ_1. Da $\operatorname{grad}\varphi_2 \leq \operatorname{grad}\varphi_1$ können wir schreiben

$$\varphi_1 = \psi_q \varphi_2 + \psi_r$$

für $\psi_q, \psi_r \in \mathbb{K}[X]$ mit $\operatorname{grad}\psi_r < l := \operatorname{grad}\varphi_2$ (Polynomdivision).

$$\begin{aligned}
0 &= \varphi_1(f)(v_2) && \text{(Nach Behauptung 1.)} \\
&= \psi_q(f)(\underbrace{\varphi_2(f)(v_2)}_{=0}) + \psi_r(f)(v_2) \\
&= \psi_r(f)(v_2) \\
&= \beta_0 v_2 + \beta_1 f(v_2) + \dots + \beta_{l-1} f^{l-1}(v_2) && \text{für } \beta_0, \dots, \beta_{l-1} \in \mathbb{K}.
\end{aligned}$$

Da $v_2, f(v_2), \dots, f^{l-1}(v_2)$ linear unabhängig, gilt $\beta_0 = \beta_1 = \dots = \beta_{l-1} = 0$.
Also $\psi_r = 0$ und ψ_2 teilt ψ_1.

Behauptung 4: φ_2 ist eindeutig. Obwohl Z_{v_1} das nicht ist!
Es sei v_1' so, dass

$$Z_{v_1'} \oplus W' = V = Z_{v_1} \oplus W\,.$$

Wir wissen bereits, dass es Basen B, B' von V gibt, so dass

$$M_B^B(f) = \begin{pmatrix} Z_{\varphi_1} & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} \text{ und } M_{B'}^{B'}(f) = \begin{pmatrix} Z_{\varphi_1} & \mathbf{0} \\ \mathbf{0} & A' \end{pmatrix}.$$

Es genügt daher zu zeigen, dass $\mu_A = \mu_{A'}$.
Sei $\psi \in \mathbb{K}[X]$. Es gibt invertierbare Matrix $T \in \mathbb{K}^{n \times n}$ mit

$$T^{-1} \begin{pmatrix} Z_{\varphi_1} & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} T = \begin{pmatrix} Z_{\varphi_1} & \mathbf{0} \\ \mathbf{0} & A' \end{pmatrix}.$$

Also

$$\begin{aligned}
\begin{pmatrix} \psi(Z_{\varphi_1}) & \mathbf{0} \\ \mathbf{0} & \psi(A') \end{pmatrix} &= \psi \begin{pmatrix} Z_{\varphi_1} & \mathbf{0} \\ \mathbf{0} & A' \end{pmatrix} \\
&= \psi(T^{-1} \begin{pmatrix} Z_{\varphi_1} & \mathbf{0} \\ \mathbf{0} & A \end{pmatrix} T) = T^{-1} \begin{pmatrix} \psi(Z_{\varphi_1}) & \mathbf{0} \\ \mathbf{0} & \psi(A) \end{pmatrix} T.
\end{aligned}$$

Insbesondere haben die Matrizen $\psi(A)$ und $\psi(A')$ den gleichen Rang. Es folgt aus dem zweiten Teil von Proposition 7.2.17, dass $\psi(A) = 0 \Leftrightarrow \psi(A') = 0$. Da ψ beliebig gewählt war, haben A und A' also das gleiche Minimalpolynom. □

Bemerkung 7.2.32. Falls $\mathbb{K} \subseteq \mathbb{K}'$ und $A \in \mathbb{K}^{n\times n} \subseteq (\mathbb{K}')^{n\times n}$, so spielt es für die Frobenius-Normalform keine Rolle, ob wir bezüglich $\mathbb{K}$ oder bezüglich $\mathbb{K}'$ rechnen. Wenn wir beispielsweise starten mit einer Matrix mit rationalen Einträgen, dann sind die Koeffizienten von $\varphi_1(X), \ldots, \varphi_k(X)$ und damit der Frobenius-Normalform (die wie bereits erwähnt auch *rationale Normalform* genannt wird) ebenfalls rational.

Beispiel 7.2.33. Das Minimalpolynom beschreibt eine lineare Abbildung nicht eindeutig. Betrachten dazu die Matrizen

$$A := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ und } B := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Es ist $\chi_A(X) = X^4 = \chi_B(X)$ und $\mu_A(X) = X = \mu_B(X)$ aber A und B haben nicht die gleiche Frobenius-Normalform: B hat die Ähnlichkeitsinvarianten $\varphi_1 = \varphi_2 = X^2$, während $p_1 = X^2$, $p_2 = X$, $p_3 = X$ die für A sind. △

Übung 37. Geben Sie eine Diagonalmatrix an, die nicht in Frobenius-Normalform ist.

Übung 38. Bestimmen Sie alle $\mathbb{K}^{n\times n}$-Matrizen, bei denen die Frobenius-Normalform aus n Kästchen besteht.

Übung 39. Zeigen Sie, dass die Frobenius Normalform einer Diagonalmatrix D mit paarweise verschiedenen Diagonaleinträgen gleich der Begleitmatrix von χ_A ist.

Übung 40. Was ist die Frobenius-Normalform von $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$?

Übung 41. Ist die Basis B aus der Frobenius-Normalform ebenfalls eindeutig, oder sind nur die Ähnlichkeitsinvarianten eindeutig?

Übung 42. Ist jede Matrix $A \in \mathbb{K}^{n\times n}$ ähnlich zu einer Matrix, die höchstens $2n - 1$ Einträge hat, die ungleich 0 sind? Beweisen Sie oder finden Sie ein Gegenbeispiel.

Korollar 7.2.34. *Für $M \in \mathbb{K}^{n\times n}$ sind die folgenden Aussagen äquivalent.*

1. *Die Frobenius-Normalform von M hat nur ein Kästchen, d.h., ist von der Gestalt $M = Z_{\mu_M}$;*
2. $\mu_M = \chi_M$.

Beweis. 1 ⇒ 2: Sei $\varphi := \mu_M$. Dann gilt

$$\begin{aligned} \chi_M &= \chi_{Z_\varphi} && \text{(Voraussetzung)} \\ &= \varphi && \text{(Proposition 7.2.1)} \\ &= \mu_M. \end{aligned}$$

2 ⇒ 1: Wenn die Frobenius-Normalform neben μ_M noch eine weitere Ähnlichkeitsinvariante ψ besitzt, denn ist $\varphi\psi$ ein Teiler von χ_M. Da $\mathrm{grad}(\psi) > 0$ folgt, dass $\mu_M \neq \chi_M$. □

Bemerkung 7.2.35. Das Produkt $\varphi_1 \cdots \varphi_k$ der Ähnlichkeitsinvarianten von $f \in \mathrm{End}(V)$ ist gleich dem charakteristischen Polynom χ_f: dies folgt direkt aus der Frobenius-Normalform und der Beobachtung dass $\chi(Z_\varphi) = \varphi$ (Beispiel 7.2.1).

Übung 43. Beweisen Sie den Satz von Cayley-Hamilton (Satz 7.2.12) aus der Frobenius-Normalform.

Zwei Polynome heißen *teilerfremd*, wenn sie nur konstante gemeinsame Teiler haben.

Beispiel 7.2.36. Seien $\varphi, \psi \in \mathbb{K}[X]$ teilerfremd, und

$$A := \begin{pmatrix} Z_\varphi & \mathbf{0} \\ \mathbf{0} & Z_\psi \end{pmatrix}.$$

Dann hat A die Frobenius-Normalform $Z_{\varphi\psi}$. Da $\chi_A = \varphi\psi$, genügt es nach Korollar 7.2.34 zu zeigen, dass $\mu_A = \chi_A$. Nach dem Satz von Cayley Hamilton ist μ_A ein Teiler von $\chi_A = \varphi\psi$. Umgekehrt wird μ_A von sowohl φ als auch ψ geteilt. Denn

$$\mathbf{0} = \mu_A(A) = \begin{pmatrix} \mu_A(Z_\varphi) & \mathbf{0} \\ \mathbf{0} & \mu_A(Z_\psi) \end{pmatrix}$$

bedeutet, dass $0 = \mu_A(Z_\varphi)$ und $0 = \mu_A(Z_\psi)$. Also wird μ_A von $\mu_{Z_\varphi} = \varphi$ und von $\mu_{Z_\psi} = \psi$ geteilt. Da φ und ψ teilerfremd sind, wird μ_A von $\varphi\psi$ geteilt. △

Ein Polynom $\varphi \in \mathbb{K}[X]$ heißt *irreduzibel* falls es nicht geschrieben werden kann als Produkt von Polynomen kleineren Grades.

Lemma 7.2.37. *Sei V endlichdimensional und $f \in \mathrm{End}(V)$. Dann haben χ_f und μ_f dieselben irreduziblen Faktoren.*

Beweis. Nach dem Satz von Cayley-Hamilton (Satz 7.2.12) gilt $\mu_f | \chi_f$. Sei also ψ ein irreduzibler Faktor von χ_f. Nach Bemerkung 7.2.35 ist $\chi_f = \varphi_1 \cdots \varphi_k$, wobei $\varphi_1, \dots, \varphi_i$ die Ähnlichkeitsinvarianten von f in der Frobenius-Normalform. Da ψ irreduzibel ist, gibt es ein $i \in \mathbb{N}$ so dass $\psi | \varphi_i$. In der Frobenius-Normalform gilt $\varphi_i | \varphi_1$ und $\varphi_1 = \mu_f$. Also $\psi | \mu_f$. □

Wir haben bereits die Existenz und Eindeutigkeit der Frobenius-Normalform bewiesen (Satz 7.2.31), aber bisher noch keinen Algorithmus kennengelernt, um diese Normalform zu berechnen. Tatsächlich gibt es sogar einen Algorithmus, der dies in polynomieller Zeit leistet, wie wir in Abschnitt 7.3.5 sehen werden.

7.2.6 Die Jordan-Weierstrass Normalform

Genau wie die Frobenius-Normalform klassizifiziert die *Jordan-Weierstrass Normalform*[2] Matrizen bis auf Ähnlichkeit. Sie teilt viele Vorzüge von Diagonalmatrizen, und tatsächlich ist die Jordan-Normalform einer diagonalisierbaren Matrix eine Diagonalmatrix (im Gegesatz zur Frobenius-Normalform, siehe Beispiel 7.2.30). Allerdings existiert die Jordan-Weierstrass Normalform nur für trigonalisierbare Matrizen, also wenn das charakteristische Polynom in Linearfaktoren zerfällt (siehe Abschnitt 4.3.6). Laut Gerd Fischer [5]

[2] Häufig auch nur *jordansche Normalform* genannt, aber etwa zeitgleich von Weierstrass entdeckt.

handelt es sich hier um das *'wohl schwierigste Theorem der elementaren linearen Algebra'*. Zunächst betrachten wir allerdings eine Vorstufe, die sogenannte *Jordan-Chevalley-Zerlegung*. Dafür benötigen wir den folgenden wichtigen Begriff.

Definition 7.2.38. Man nennt $f \in \mathrm{End}(V)$ *nilpotent* falls $f^k = \underline{0}$ für ein $k \in \mathbb{N}$. Analog dazu heißt $A \in \mathbb{K}^{n \times n}$ *nilpotent* falls $A^k = \mathbf{0}$ für ein $k \in \mathbb{N}$.

Intuitiv kann man nilpotente Matrizen als solche betrachten, die "besonders schlimm nicht diagonalisierbar" sind.

Satz 7.2.39. *Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}(V)$. Dann sind äquivalent:*

1. *f ist nilpotent.*
2. *$\mu_f = X^d$ für ein $d \in \{1, \dots, n\}$.*
3. *$\chi_f = X^n$.*
4. *Es gibt eine Basis B von V, so dass*

$$M_B^B(f) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}. \tag{7.7}$$

Beweis. Falls f nilpotent ist, dann gibt es ein $k \in \mathbb{N}$ mit $X^k \in \mathcal{I}_f$ (das Ideal von f, definiert in Beispiel 7.2.7). Also ist $\mu_f = X^d$ mit $1 \leq d \leq n$ (siehe Bemerkung 7.2.13). Die Implikation 2. $\Rightarrow$ 3. folgt aus dem Lemma, dass μ_f und χ_f dieselben irreduziblen Faktoren haben (Lemma 7.2.37), denn nach Annahme ist X ist der einzige irreduzible Faktor von μ_f, und χ_f ist vom Grad n. Die Implikation 3. $\Rightarrow$ 4. folgt aus unserem Trigonalisierungskriterium (Satz 4.3.24): wenn $\chi_f = X^n$, dann ist f trigonalisierbar. Da alle Eigenwerte 0 sind, hat die Dreiecksform auf der Diagonalen nur Einträge 0, und ist daher von der gewünschten Gestalt (7.7). Die Implikation 4. $\Rightarrow$ 1. folgt aus Übung 44. □

Bemerkung 7.2.40. Eine Matrix in der Gestalt von (7.7) zeichnet sich dadurch aus, dass der zugehörige gerichtete Graph (wie etwa in Beispiel 4.3.27) keine gerichteten Kreise hat.

Übung 44. Zeigen Sie, dass eine Matrix der Gestalt (7.7) nilpotent ist.

Satz 7.2.41 (Jordan-Chevalley-Zerlegung)**.** *Sei $A \in \mathbb{K}^{n \times n}$ so, dass χ_A in Linearfaktoren zerfällt. Dann ist $A = D + N$, wobei*

- *D diagonalisierbar,*
- *N nilpotent, und*
- *$DN = ND$;*

hierbei sind D und N eindeutig.

Dieser Satz wird bisweilen auch *Hauptraumzerlegung* genannt, nach einem wichtigen Begriff, der für den Beweis des Satzes benötigt wird: der Begriff eines *Hauptraumes* (Im Englischen: *generalised eigenspace*. Vergleiche mit der Definition eines Eigenraumes (Definition 4.3.3).

Definition 7.2.42. Sei V ein n-dimensionaler Vektorraum und $f \in \operatorname{End}(V)$. Sei λ ein Eigenwert von f und $m \in \mathbb{N}$. Dann heißt

$$\operatorname{Hau}_{\lambda,m}(f) := \operatorname{Kern}\big((f - \lambda \operatorname{id}_V)^m\big)$$

der *m-te Hauptraum vom f zum Eigenwert λ.*

Beispiel 7.2.43. Falls V ein n-dimensionaler Vektorraum und $f \in \operatorname{End}(V)$ nilpotent, dann gilt $\operatorname{Hau}_{\lambda,n}(f) = V$. △

Beweis des Satzes von der Hauptraumzerlegung, Satz 7.2.41. Es sei

$$\chi_A(X) = (X - \lambda_1)^{m_1} \cdots (X - \lambda_k)^{m_r}$$

für $\lambda_1, \ldots, \lambda_r$ paarweise verschieden. (Falls $m_1 = \cdots = m_r = 1$, so ist A diagonalisierbar (Bemerkung 4.3.20), und die Aussage stimmt mit $N = \mathbf{0}$.)

Sei $f := f_A$ und $V_i := \operatorname{Hau}_{\lambda_i,r_i}(f)$. Dann gilt (und daher rührt der Name 'Hauptraumzerlegung'):

$$V = V_1 \oplus \cdots \oplus V_r. \tag{7.8}$$

Um das zu sehen, schreiben wir

$$\begin{aligned} V &= \operatorname{Kern}\big(\chi_f(f)\big) && \text{(Satz 7.2.12)} \\ &= \operatorname{Kern}\big((f - \lambda_1 \operatorname{id}_V)^{m_1} \oplus \cdots \oplus (f - \lambda_k \operatorname{id}_V)^{m_r}\big). \end{aligned}$$

Es genügt also zu zeigen, dass falls $\varphi, \psi \in \mathbb{K}[X]$ teilerfremd sind, sich der Kern der linearen Abbildung $(\varphi\psi)(f)$ wie folgt schreiben läßt:

$$\operatorname{Kern}\big((\varphi\psi)(f)\big) = \operatorname{Kern}\big(\varphi(f)\big) \oplus \operatorname{Kern}\big(\psi(f)\big) \tag{7.9}$$

Denn $\varphi := (X - \lambda_1)^{m_1}$ ist teilerfremd zu $\psi := \prod_{i \in \{2,\ldots,k\}} (X - \lambda_i)^{m_i}$, also

$$\underbrace{\operatorname{Kern}(\varphi(f))}_{=V_1} \oplus \operatorname{Kern}(\psi(f)),$$

und die Aussage folgt per vollständiger Induktion.

Wir zeigen zunächst, dass

$$U := \operatorname{Kern}\big(\varphi(f)\big) \cap \operatorname{Kern}\big(\psi(f)\big) = \{\mathbf{0}\}.$$

Der Untervektorraum U ist f-invariant, denn falls $v \in U$, dann gilt $\varphi(f)(v) = \mathbf{0}$, und damit

$$\varphi(f)(f(v)) = f \underbrace{\varphi(f)(v)}_{\mathbf{0}} = \mathbf{0},$$

also $f(v) \in \operatorname{Kern}\big(\varphi(f)\big)$. Analog zeigt man, dass $f(U) \subseteq \operatorname{Kern}\big(\psi(f)\big)$. Also $f(U) \subseteq U$. Dann gilt $\varphi(f|_U) = \underline{0}$ und $\psi(f|_U) = \underline{0}$; das Minimalpolynom von $f|_U$ teilt also sowohl φ als auch ψ, ist also nach Annahme konstant. Das ist nur dann möglich, wenn $U = \{\mathbf{0}\}$ (Übung 36).

Es sei $W := \operatorname{Kern}\big((\varphi\psi)(f)\big)$. Offenbar gilt $\operatorname{Kern}\big(\varphi(f)\big) \subseteq W$ und $\operatorname{Kern}\big(\psi(f)\big) \subseteq W$. Um Aussage (7.9) zu zeigen, genügt es also zu zeigen, dass

$$\dim \underbrace{\operatorname{Kern}\big((\varphi\psi)(f)\big)}_{=W} = \dim \operatorname{Kern}\big(\varphi(f)\big) + \dim \operatorname{Kern}\big(\psi(f)\big). \tag{7.10}$$

Analog zur f-Invarianz von U zeigt man die f-Invarianz von W. Es gilt

$$\operatorname{Bild}(\psi(f|_W)) \subseteq \operatorname{Kern}(\varphi(f|_W)). \tag{7.11}$$

Denn falls $u \in \operatorname{Bild}(\psi(f|_W))$, so gibt es $v \in W$ mit $\psi(f|_W(v)) = u$. Da $v \in W$, gilt

$$\varphi(f|_W)(u) = \varphi\psi(f|_W)(v) = \mathbf{0},$$

also $u \in \operatorname{Kern}(\phi(f|_W))$.

Weiterhin gilt

$$\operatorname{Kern}(\varphi(f|_W)) = \operatorname{Kern}(\varphi(f)) \tag{7.12}$$

Trivialerweise ist $\operatorname{Kern}(\varphi(f|_W)) \subseteq \operatorname{Kern}(\varphi(f))$. Umgekehrt, falls $u \in \operatorname{Kern}(\varphi(f))$, dann ist $u \in W$, also in $\operatorname{Kern}(\varphi(f|_W))$.

$$\begin{aligned}
\dim W &= \dim \operatorname{Kern}(\varphi(f|_W)) + \dim \operatorname{Bild}(\varphi(f|_W)) && \text{(Dimensionsformel, Satz 3.3.6)}\\
&\leq \dim \operatorname{Kern}(\varphi(f|_W)) + \dim \operatorname{Kern}(\psi(f|_W)) && \text{(7.11)}\\
&= \dim \operatorname{Kern}(\varphi(f)) + \dim \operatorname{Kern}(\psi(f)) && \text{(7.12).}
\end{aligned}$$

Es gilt sogar Gleichheit, da $\operatorname{Kern}(\varphi(f)) \subseteq W$ und $\operatorname{Kern}(\varphi(f)) \subseteq W$, und $\operatorname{Kern}(\varphi(f)) \cap \operatorname{Kern}(\varphi(f)) = \{\mathbf{0}\}$. Damit haben wir (7.10), und in Folge (7.9) und (7.8) bewiesen.

Wir definieren nun $f_D, f_N \in \operatorname{End}(V)$ durch

$$\begin{aligned}
(f_D)|_{V_i} &:= \lambda_i \operatorname{id}_{V_i}\\
(f_N)|_{V_i} &:= (f - \lambda_i \operatorname{id})|_{V_i}
\end{aligned}$$

für alle $i \in \{1, \ldots, r\}$. Sei B eine Basis von V zusammengesetzt aus den Basen für $V_1, \ldots, V_r$; definieren $D := M_B^B(f_D)$ und $N := M_B^B(f_N)$. Offensichtlich ist $f = f_D + f_N$ und $A = D + N$. Ausserdem ist D in Diagonalform, und daher $DN = ND$.

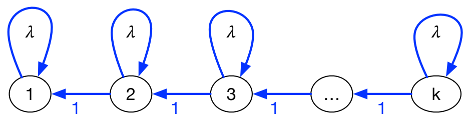


Abbildung 7.3: Illustration eines Jordankästchens der Größe k zum Eigenwert $\lambda \in \mathbb{K}$ aus Beispiel 7.2.44. Zur Erklärung der Bedeutung der Illustrationen siehe Beispiel 4.3.27.

Wir zeigen nun, dass N nilpotent ist. Sei $v \in V$; wir zeigen, dass $N^n(v) = \mathbf{0}$. Das ist trivial für $v = \mathbf{0}$. Ansonsten liegt v nach (7.8) in $V_i = \operatorname{Kern}\big((f - \lambda_i \operatorname{id}_V)^{m_i}\big)$ für genau ein $i \in \{1, \dots, r\}$. Es gilt also $\mathbf{0} = (f - \lambda_i \operatorname{id}_V)^{m_i}(v) = (f_N)^{m_i}(v)$, und damit auch $N^n(v) = \mathbf{0}$.
Eindeutigkeit: Siehe Peter Petersen, *Linear Algebra* [9], Übung 14 in Abschnitt 2.8, oder Gerd Fischer, *Lineare Algebra* [5], 'Zusatz' in Abschnitt 4.6.3 (Normalformen werden in Albrecht Beutelspacher, *Lineare Algebra* [1] nicht behandelt). □

Beispiel 7.2.44. Betrachte

$$J_n(\lambda) := \underbrace{\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & & \\ \vdots & & \ddots & \\ 0 & & & \lambda \end{pmatrix}}_{=D} + \underbrace{\begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}}_{=N} = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

Siehe Abbildung 7.3. Dann ist $\mu_{J_n(\lambda)}(X) = (X - \lambda)^n = \chi_{J_n(\lambda)}(X)$ (Proposition 7.2.20). Also ist $Z_{(X-\lambda)^n}$ ähnlich zu $J_n(\lambda)$. △

Die Matrix $J_n(\lambda)$ aus Beispiel 7.2.44 heißt *Jordankästchen* der Größe n zum Eigenwert $\lambda \in \mathbb{K}$. Diese Matrizen sind die Bausteine für die Jordan-Normalform. Spezialfall $n = 1$:

$$J_1(\lambda) = (\lambda) \in \mathbb{K}^{1\times 1}$$

Satz 7.2.45 (Jordan-Weierstrass Normalform). *Sei $A \in \mathbb{K}^{n\times n}$ und das charakteristische Polynom zerfalle in Linearfaktoren:*

$$\chi_A(X) = (\lambda_1 - X)^{m_1} \cdots (\lambda_r - X)^{m_r}$$

$\lambda_1, \dots, \lambda_r$*: Eigenwerte von A.*
$m_1, \dots, m_r$*: algebraische Vielfachheiten.*
$n_1, \dots, n_r$*: geometrische Vielfachheiten.*
Dann existiert eine zu A ähnliche Matrix J der Gestalt

$$\begin{pmatrix} J_{s(1,1)}(\lambda_1) & & & & & & \mathbf{0} \\ & \ddots & & & & & \\ & & J_{s(1,n_1)}(\lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & J_{s(r,1)}(\lambda_r) & & \\ & & & & & \ddots & \\ \mathbf{0} & & & & & & J_{s(r,n_r)}(\lambda_r) \end{pmatrix}.$$

Zu jedem Eigenwert λ_i*,* $i \in \{1, \dots, r\}$*, gibt es* n_i *Jordankästchen* $J_{s(i,1)}(\lambda_i), \dots, J_{s(i,n_i)}(\lambda_i)$*, deren Länge sich zu* m_i *aufsummiert, d.h.,* $\sum_{j \in \{1,\dots,n_i\}} s(i,j) = m_i$*. Die Gesamtanzahl der Jordankästchen ist demnach* $t := \sum_{i \in \{1,\dots,r\}} n_i$*.*

Die Matrix J heißt auch *jordansche Normalform* und ist bis auf die Reihenfolge der Jordankästchen eindeutig. Diese kann eindeutig gemacht werden durch Festlegung $s(i,1) \geq s(i,2) \geq \cdots \geq s(i,n_i)$ für jedes $i \in \{1, \dots, r\}$ und $\lambda_1 < \lambda_2 < \cdots < \lambda_r$ (im Falle von $\mathbb{K} = \mathbb{R}$, ansonsten fixiere beliebig Ordnung auf $\mathbb{K}$).

Bemerkung 7.2.46. Aus der jordanschen Normalform lassen sich sofort ablesen:

- die Eigenwerte: die Hauptdiagonalelemente.
- die algebraische Vielfachheit von EW λ_i:
  die Anzahl m_i der λ_i auf der Hauptdiagonale.
- die geometrische Vielfachheit von EW λ_i:
  die Anzahl n_i der Jordankästchen zum EW λ_i.
- das charakteristische Polynom in faktorisierter Form, nämlich $\prod_{i \in \{1,\dots,r\}} (X - \lambda_i)^{m_i}$.
- das Minimalpolynom in faktorisierter Form, nämlich $\prod_{i \in \{1,\dots,r\}} (X - \lambda_i)^{s(i,1)}$.

$\rightsquigarrow$ Klassifikation durch charakteristische Daten: brauchen blos die Werte

$$\lambda_1, \dots, \lambda_r, n_1, \dots, n_r, s(1,1), \dots, s(r, n_r)$$

Beweis von Satz 7.2.45. Sei $f := f_A \in \mathrm{End}(V)$. Wie im Beweis von Satz 7.2.41 schreiben wir $f = f_D + f_N$ für f_D diagonalisierbar und f_N nilpotent. Dann zerlegen wir V in Eigenräume für f_D:

$$V = \mathrm{Kern}(f_D - \lambda_1 \,\mathrm{id}) \oplus \cdots \oplus \mathrm{Kern}(f_D - \lambda_k \,\mathrm{id})$$

Die Eigenräume sind auch f_N-invariant: sei v so, dass $(f_D - \lambda_1 \,\mathrm{id})(v) = \mathbf{0}$. Dann ist

$$\begin{aligned} (f_D - \lambda_1 \,\mathrm{id})(f_N v) &= (f_D f_N - \lambda_1 \,\mathrm{id}\, f_N)(v) \\ &= f_N \underbrace{(f_D - \lambda_1 \,\mathrm{id})(v)}_{=\mathbf{0}} = \mathbf{0}. \end{aligned}$$

Nach Satz 7.2.31 existiert eine Basis B, so dass

$$M_B^B(f_N) = \begin{pmatrix} Z_{\varphi_1} & 0 & \cdots & 0 \\ 0 & Z_{\varphi_2} & & \\ \vdots & & \ddots & \\ 0 & & & Z_{\varphi_k} \end{pmatrix}$$

die Frobenius Normform von f_N ist. Die Ähnlichkeitsinvarianten $\varphi_1, \dots, \varphi_k$ sind alle von der Gestalt X^k, für ein $k \leq n$, da $f_N^n = \underline{0}$, also sehen die Blöcke der FNF so aus:

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

Sei B' die geordnete Basis, die man aus B durch Umdrehen der Reihenfolge erhält. Dann ist $J := M_{B'}^{B'}(f_A) = M_{B'}^{B'}(f_N) + M_{B'}^{B'}(f_D)$ die jordansche Normalform von A. Es bleibt zu zeigen, dass die Anzahl der Jordankästchen zum Eigenwert λ_i genau die geometrische Vielfachheit von λ_i ist, also $n_i = \dim(\operatorname{Kern}(A - \lambda_i E_n))$. Zum einen sind zwei Spalten von J linear unabhängig, wenn sie in unterschiedlichen Jordankästchen liegen. Zum anderen gilt für die jeweils erste Spalte s jedes Jordankästchens zum Eigenwert λ_i, dass $(A - \lambda_i E_n)s = 0$, und damit folgt die Behauptung. □

7.2.7 Beispiele

Beispiele zur Berechnung der Frobenius-Normalform, und, falls möglich, der jordanschen Normalform (letzteres kann abhängen vom Körper, in dem wir rechnen).

Beispiel 7.2.47.

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

Für Frobenius-Normalform: suchen $v_1 \in V$ so, dass $\dim Z_{v_1}$ größtmöglich.

Wähle $v_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, erhalten $Av_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ und $A^2 v_1 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} = 2Av_1$. Dann ist $\dim Z_{v_1} = 2$.

Geht noch besser: Wähle $v_1 := \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, erhalten

$$Av_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \text{ und } A^2 v_1 = \begin{pmatrix} 4 \\ 1 \\ 4 \end{pmatrix}$$

allesamt linear unabhängig, und damit ist $Z_{v_1} = V$ von größtmöglicher Dimension. Haben

$$A^3 v_1 = (8, 1, 8) = (12, 3, 12) - (4, 2, 4) = 3A^2 v_1 - 2 \cdot Av_1 + 0 \cdot v_1 \,.$$

Mit Basis $B = (v_1, Av_1, A^2 v_1)$ ist

$$M_B^B(f_A) = Z_{\varphi_1}$$

für $\varphi_1(X) = X^3 - 3X^2 + 2X = \mu_A(X)$. Die Frobenius-Normalform von A ist also

$$M_B^B(f_A) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

Zur Berechnung der jordanschen Normalform. Wir bestimmen zunächst das charakteristische Polynom.

$$\begin{aligned} \chi_A(X) &= \det(A - XE_3) = (1-X)^3 - (1-X) = -X^3 + 3X^2 - 2X \\ &= -X(X^2 - 3X + 2) = -X(X-1)(X+2) \end{aligned}$$

Wir haben also drei verschiedene Eigenwerte bei $\dim V = 3$, und damit eine Basis aus Eigenvektoren. Also ist A diagonalisierbar (siehe Bemerkung 4.3.20). Sei S eine Matrix, deren Spalten aus Eigenvektoren von A zu den verschiedenen Eigenwerten besteht, also etwa

$$S = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 3 \end{pmatrix}.$$

Dann gilt

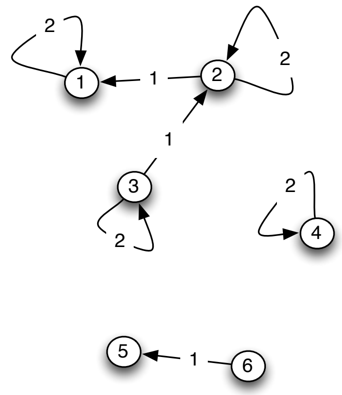
$$\underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 3 \end{pmatrix}}_{=S} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 3 \end{pmatrix}}_{=S} \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}}_{=:D}$$

also $S^{-1}AS = D$ und damit ist A ähnlich zu einer Diagonalmatrix. △

Beispiel 7.2.48. Die Matrix

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

ist nicht diagonalisierbar, aber bereits in Jordan-Normalform.



Zwei Eigenwerte:

- $\lambda_1 = 2$, algebraische Vielfachheit 4, geometrische Vielfachheit 2.
- $\lambda_2 = 0$, algebraische Vielfachheit 2, geometrische Vielfachheit 1.

Das charakteristisches Polynom ist $\chi_A(X) = (2 - X)^4 X^2$. △

Beispiel 7.2.49.

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & -1 & 3 & 1 & 0 & -1 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 1 & -4 & 2 & 2 & 4 & 0 \\ 0 & 2 & -1 & -1 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{6\times 6}$$

1. Eigenwerte bestimmen. Wir berechnen die Determinante für das charakteristische Polynom durch Entwicklung nach der ersten und 4ten Reihe, und dann nach der 4ten Spalte.

$$\begin{aligned}
\det(XE_6 - A) &= \begin{vmatrix} X-2 & 0 & 0 & 0 & 0 & 0 \\ -1 & X & -1 & -1 & -1 & 0 \\ -1 & 1 & X-3 & -1 & 0 & 1 \\ 0 & 0 & 0 & X-2 & 0 & 0 \\ -1 & 4 & -2 & -2 & X-4 & 0 \\ 0 & -2 & 1 & 1 & 1 & X-2 \end{vmatrix} \\
&= (X-2)^2 \begin{vmatrix} X & -1 & -1 & 0 \\ 1 & X-3 & 0 & 1 \\ 4 & -2 & X-4 & 0 \\ -2 & 1 & 1 & X-2 \end{vmatrix} \\
&= (X-2)^3 \begin{vmatrix} X & -1 & -1 \\ 1 & X-3 & 0 \\ 4 & -2 & X-4 \end{vmatrix} + (X-2)^2 \begin{vmatrix} X & -1 & -1 \\ 4 & -2 & X-4 \\ -2 & 1 & 1 \end{vmatrix} \\
&= \cdots = (X-2)^5 (X-3) = \chi_A(X)
\end{aligned}$$

Eigenwerte:

$$\begin{array}{ll} \lambda_1 = 2 & \text{algebraische Vielfachheit } m_1 = 5 \\ \lambda_2 = 3 & \text{algebraische Vielfachheit } m_2 = 1 \end{array}$$

Jordansche Normalform existiert also.

2. Basen der Eigenräume bestimmen. Idee: Basisvektoren bilden "Startpunkte der Jordankästchen"
   - Für $\lambda_1 = 2$: $\mathrm{Eig}_{\lambda_1}(A) = \mathrm{Kern}(A - \lambda_1 E_n)$:

     $$A_1 := A - \lambda_1 E_n = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 2 & 2 & 2 & 0 \\ 0 & 2 & -1 & -1 & -1 & 0 \end{pmatrix}$$

     $\mathrm{rg}(A_1) = 3$ (Zwei Nullzeilen und $z_6 = z_2 - z_5$ zeigt $\mathrm{rg}(A_2) \leq 3$, und z_1, z_2, z_3 sind linear unabhängig) Also:

     $$\dim \mathrm{Eig}_{\lambda_1} = \dim \mathrm{Kern}\, A_1 = n - \mathrm{rg}(A_1) = 6 - 3 = 3$$

     Finden drei linear unabhängige Eigenvektoren aus $\mathrm{Lös}(A_1, \mathbf{0})$, z.B. mit dem gaußschen Algorithmus:

     $$u_1 := \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, u_2 := \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, u_3 := \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

     $B = (u_1, u_2, u_3)$ ist Basis für $\mathrm{Eig}_{\lambda_1}(A)$.
   - Für $\lambda_2 = 3$: $\dim \mathrm{Eig}_{\lambda_2} \leq m_2 = 1$ ($m_2 = 1$ ist algebraische Vielfachheit),

     $$A_2 = A - \lambda_2 E_n.$$

     Lösung von $A_2 u = \mathbf{0}$:

     $$u_4 := \begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \\ -2 \\ 1 \end{pmatrix}$$

     Also $B_{\lambda_2} = \{u_4\}$.

3. Bestimmung der Jordankästchen.
   Allgemeine Idee: Wenn B eine Basis mit $J = M_B^B(f_A)$ in jordanscher NF, dann muss für die zu einem Jordanblock von J zu Eigenwert λ gehörigen Spalten $v^{(1)}, \dots, v^{(s)}$ gelten

$$\begin{aligned} Av^{(1)} &= \lambda v^{(1)} \\ Av^{(2)} &= v^{(1)} + \lambda v^{(2)} \\ &\vdots \\ Av^{(s)} &= v^{(s-1)} + \lambda v^{(s)} \end{aligned}$$

   - $u_1^{(1)} := u_1 = (0,1,2,0,0,1)^\top$.
     $u_1^{(2)}$: Lösung von $A_1 u_1^{(2)} = u_1^{(1)}$, etwa $u_1^{(2)} = (2,1,1,0,0,0)^\top$.
     Keine Lösung für $u_1^{(3)}$ in $A_1 u_1^{(3)} = u_1^{(2)}$.
     (Erste Zeile von A_1 Null, erste Komponente von $u_1^{(2)}$ nicht Null.)
     Jordankästchen $K_{u_1} := (u_1^{(1)}, u_1^{(2)})$.
   - $u_2^{(1)} := u_2 = (0,1,1,0,1,0)^\top$.
     $u_2^{(2)}$: Lösung von $A_1 u_2^{(2)} = u_2^{(1)}$, etwa $u_2^{(2)} = (1,0,0,0,0,0)^\top$.
     Keine Lösung für $u_1^{(3)}$ in $A_1 u_2^{(3)} = u_2^{(2)}$.
     (erste Komponente von $u_2^{(2)}$ nicht Null, erste Zeile von A_1 Null.)
     Jordankästchen $K_{u_2} = (u_2^{(1)}, u_2^{(2)})$.
   - $u_3^{(1)} := u_3 = (0,0,-1,1,0,0)^\top$,
     und keine Lösung für $u_3^{(2)}$ in $A_1 u_3^{(2)} = u_3^{(1)}$.
     Jordankästchen $K_{u_3} = (u_3^{(1)})$.
   - Jordankästchen $K_{u_4} = (u_4^{(1)}) = ((0,-1,-1,0,-2,1)^\top)$.

4. Zusammensetzen: $B = (u_1^{(1)}, u_1^{(2)}, u_2^{(1)}, u_2^{(2)}, u_3^{(1)}, u_4^{(1)})$ ist Basis von $\mathbb{R}^6$ und liefert die Spalten für

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

Die Jordansche Normalform von C ist

$$J = S^{-1}AS = \begin{pmatrix} 2 & 1 & & & & \\ 0 & 2 & & & & \\ & & 2 & 1 & & \\ & & 0 & 2 & & \\ & & & & 2 & \\ & & & & & 3 \end{pmatrix}. \qquad \triangle$$

Algorithmus zur Berechnung der Jordanschen Normalform:

Gegeben: $A \in \mathbb{K}^{n\times n}$.

1. Bestimme alle Eigenwerte $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ von A und deren algebraische Vielfachheiten $m_1, \ldots, m_r$ (faktorisiere das charakteristische Polynom). Falls $\sum_{i=1}^r m_i < n$, gebe aus 'A besitzt keine JNF'.

2. Für $i \in \{1, \ldots, r\}$, sei $A_i := A - \lambda_i E_n$. Berechne Basis $B_i = (u_1, \ldots, u_{n_i})$ des Eigenraums
$$\mathrm{Eig}_{\lambda_i}(A) = \mathrm{Kern}(A_i).$$
Größe n_i der Basis ist die Anzahl der Jordankästchen zum Eigenwert λ_i.

3. Bestimmung der Jordankästchen: Für jeden Basisvektor $u \in B_i$, suchen nach größtmöglichem s, so dass es Vektoren $u^{(1)}, \ldots, u^{(s)}$ gibt mit $u^{(1)} := u$ und
$$\begin{aligned} Au^{(1)} &= \lambda_i u^{(1)} && \text{i.e., } (A - \lambda_i E_n)u^{(1)} = \mathbf{0} \\ Au^{(2)} &= 1 \cdot u^{(1)} + \lambda_i u^{(2)} && \text{i.e., } (A - \lambda_i E_n)u^{(2)} = u^{(1)} \\ &\dots \\ Au^{(s)} &= 1 \cdot u^{(s-1)} + \lambda_i u^{(s)} && \text{i.e., } (A - \lambda_i E_n)u^{(s)} = u^{(s-1)}. \end{aligned} \tag{7.13}$$
Siehe Abbildung 7.4.

4. Aneinanderreihung aller Vektoren $u^{(j)}$ aus Schritt 3. ergibt Basis B von $\mathbb{K}^n$.
Sei S die Matrix mit Vektoren von B als Spalten. Dann gilt:
$$J := S^{-1}AS = M_B^B(f_A)$$
ist die jordansche Normalform von A (J gewinnt man direkt aus den charakteristischen Daten, die wir zuvor allesamt berechnet haben).

Wichtig: Probe machen. Müssen wir dazu die Inverse von S berechnen? Nein!
Machen die Probe mit
$$SJ = SA.$$

$$J = \begin{array}{c} \begin{array}{ccccccccc} & & & & u^{(1)} & u^{(2)} & u^{(3)} & & \end{array} \\ \begin{pmatrix} \lambda_1 & & & & 0 & 0 & 0 & & \mathbf{0} \\ & \ddots & & & 0 & 0 & 0 & & \\ & & \lambda_1 & & 0 & 0 & 0 & & \\ & & & \ddots & 0 & 0 & 0 & & \\ & & & & \lambda_i & 1 & 0 & & \\ & & & & 0 & \lambda_i & 1 & & \\ & & & & 0 & 0 & \lambda_i & & \\ & & & & 0 & 0 & 0 & \ddots & \\ \mathbf{0} & & & & 0 & 0 & 0 & & \lambda_r \end{pmatrix} \end{array}$$

Abbildung 7.4: Illustration zu Schritt 3 im Algorithmus zur Berechnung der jordanschen Normalform.

Bemerkung 7.2.50. Es gibt zwei kritische Stellen in diesem Algorithmus. Der erste ist Schritt 1, denn wir haben noch kein effizientes Verfahren kennengelernt, um das charakteristische Polynom in faktorisierter Form zu berechnen. Wir werden später auf dieses Problem zurückkommen (siehe 7.3.22). Die zweite Stelle ist die Suche nach $u^{(1)}, \dots, u^{(s)}$ in Schritt 3., so dass s größtmöglich. Mit dem gaußschen Algorithmus können wir herausfinden, ob es in (7.13) eine Lösung gibt für $u^{(s)}$. Das liefert aber noch keine Gewissheit, dass s größtmöglich ist (denn dafür könnte es notwendig sein, bereits $u^{2}, \dots, u^{(s-1)}$ anders zu wählen). Auch dieses Problem werden wir in Abschnitt 7.3.5 lösen.

Bemerkung 7.2.51. Effiziente Algorithmen zur Berechnung der jordanschen Normalform und deren exakte Komplexität sind Gegenstand aktueller Forschung; es sei verwiesen auf [4].

Übung 45. Sei $J \in \mathbb{K}^{n \times n}$ in jordanscher Normalform. Zeigen Sie: die Anzahl der Jordankästchen zum Eigenwert 0 ist gleich $n - \mathrm{rg}(A)$.

Übung 46. Beschreiben Sie, wie man mit Hilfe der jordanschen Normalform die Frobenius-Normalform berechnen kann.[3]

Übung 47. Sei V ein beliebiger, nicht notwendigerweise endlichdimensionaler Vektorraum. Dann heißt $f \in \mathrm{End}(V)$ *diagonalisierbar*, wenn V eine Basis aus Eigenvektoren von f besitzt (wie im endlichdimensionalen Fall, siehe Satz 4.3.19). Finden Sie ein Beispiel für ein $f \in \mathrm{End}(V)$ mit V unendlichdimensional, welches sich *nicht* schreiben läßt als $f_N + f_D$ mit f_N nilpotent und f_D diagonalisierbar.[4] Tipp: Übung 28.

[3] Dank an die Teilnehmer:innen der VL im SS 2024 für die Idee zu dieser Übung.

[4] Inspiriert durch die Teilnehmer:innen der VL im SS 2024.

7.3 Die Hermite-Normalform

Für $A \in \mathbb{Q}^{m \times n}$ kennen wir bereits einen effizienten Algorithmus, um zu entscheiden, ob das Gleichungssystem $Ax = b$ eine Lösung in $\mathbb{Q}^n$ besitzt (Abschnitt 4.1.8). In diesem Abschnitt wollen wir einen effizienten Algorithmus für Lösbarkeit in $\mathbb{Z}^n$ vorstellen. Wenn man alle Zeilen mit dem kleinsten gemeinsamen Vielfachen der Nenner aller Koeffizienten multipliziert, erhält man ein System mit der gleichen Lösungsmenge, aber ganzzahligen Koeffizienten. Wir werden daher annehmen, dass $A \in \mathbb{Z}^{m \times n}$. Solche Systeme $Ax = b$ werden auch *lineare diophantische Gleichungssysteme* genannt. Ganzzahlige Lösungen sind in sehr vielen Anwendungen relevant, z.B. in der diskreten Optimierung.

Um ganzzahlige Lösungen zu finden, behandeln wir eine neue Normalform, die *Hermite-Normalform*[5] (Abschnitt 7.3.2). Unser Beweis für die Existenz der Hermite-Normalform liefert noch kein effizientes Verfahren, um diese auch wirklich auszurechnen; mit einem cleveren Trick aber gelingt uns das in Abschnitt 7.3.3. Die Anwendung auf lineare diophantische Gleichungssysteme folgt dann in Abschnitt 7.3.4. In diesen drei Abschnitten folgen wir im wesentlichen Kapitel 4 und 5 des Lehrbuchs von Schrijver [10].

Viele Aussagen in diesem Abschnitt lassen sich für alle, oder doch zumindest für gewisse Ringe verallgemeinern (z.B. für den Polynomring $\mathbb{Q}[X]$, und allgemeiner für *Hauptidealringe*; Gegenstand der Vorlesung AL10). Ähnliche Ideen wie bei der Hermite-Normalform helfen uns dann, auf überraschende Art und Weise (wie ich finde) auch die Frobenius-Normalform (und damit auch die Jordan-Normalform, siehe Beweis von Satz 7.2.45) effizient zu berechnen! Dazu benötigen wir eine weitere Normalform, nämlich die Smith-Normalform (Abschnitt 7.3.5), die wir ebenfalls für Matrizen aus dem Polynomring $\mathbb{Q}[X]$ berechnen können. Die Berechnung der Frobenius-Normalform mit Hilfe der Smith-Normalform ist dann Gegenstand von Abschnitt 7.3.6.

7.3.1 Unimodulare Spaltenäquivalenz

In diesem Abschnitt betrachten wir eine neue Äquivalenzrelation auf Matrizen. Wir benötigen dafür den folgenden Begriff.

Definition 7.3.1. Es sei R ein Ring mit Eins. Dann heißt $A \in R^{n \times n}$ *unimodular* falls $\det(A)$ eine Einheit in R ist (Definition 4.2.2). Zwei Matrizen $A, B \in R^{m \times n}$ heissen *unimodular spaltenäquivalent* falls $A = BU$ für eine unimodulare Matrix $U \in R^{n \times n}$.

Unimodulare Spaltenäquivalenz ist tatsächlich eine Äquivalenzrelation. *Unimodulare Zeilenäquivalenz* ist analog definiert.

Beispiel 7.3.2. Eine *Permutationsmatrix* (auch *Vertauschungsmatrix*) ist eine quadratische Matrix, in der in jeder Zeile und in jeder Spalte genau ein Eintrag eins ist und alle anderen Einträge null sind. Jede Permutationsmatrix $P \in \mathbb{K}^{n \times n}$ entspricht genau einer Permutation $\pi \in \mathrm{Sym}(\{1, 2, \ldots, n\})$: die zu π gehörige Permutationsmatrix hat die Einträge $p_{ij} = 1$ falls $\pi(i) = j$ und $p_{ij} = 0$ sonst. Alle Permutationsmatrizen sind

[5] Benannt nach Charles Hermite, geboren am 24.12.1822 in Dieuze, Lothringen; gestorben am 14.1.1901 in Paris.

unimodular (Proposition 4.1.3). Das gilt insbesondere für die Elementarmatrix für das Vertauschen zweier Spalten (Abschnitt 3.2.3). △

Beispiel 7.3.3. Die Elementarmatrix für die Multiplikation einer Spalte mit einem Skalar $r \in R$ (Abschnitt 3.2.3) ist genau dann unimodular, wenn r eine Einheit ist (Proposition 4.1.3). △

Beispiel 7.3.4. Die Elementarmatrix für die Addition einer Spalte mit einem Vielfachen einer anderen Spalte (Abschnitt 3.2.3) ist unimodular (Proposition 4.1.3). △

Im Folgenden sei $R = \mathbb{Z}$.

Bemerkung 7.3.5. Eine Matrix $A \in \mathbb{Z}^{n\times n}$ ist unimodular, falls $\det A \in \{-1, 1\}$.

Definition 7.3.6. Vertauschen von Spalten, Multiplikation einer Spalte mit -1, und Addition eines ganzzahligen Vielfachen einer Spalte zu einer anderen heißen *elementare unimodulare Spaltentransformationen*. Entsprechend heißen Permutationsmatrizen, Elementarmatrizen für die Multiplikation mit -1, und Elementarmatrizen für die Addition eines ganzzahligen Vielfachen einer Spalte zu einer anderen (nicht derselben!) *unimodulare Elementarmatrizen*.

Wir werden eine Normalform für Matrizen $A \in \mathbb{Z}^{m\times n}$ bis auf unimodulare Spaltenäquivalenz in Abschnitt 7.3.2 kennenlernen, die Hermite-Normalform.

7.3.2 Die Hermite-Normalform

Die Hermite-Normalform ist eine Normalform für Matrizen bis auf unimodulare Spaltenäquivalenz (Abschnitt 7.1.3). Analog erhält man auch eine Normalform für unimodulare Zeilenäquivalenz. Die Formulierung der Normalform für Spaltenäquivalenz (anstatt Zeilenäquivalenz) wird praktisch sein bei unserer Anwendung für Lösbarkeit linearer diophantischer Gleichungssysteme (Abschnitt 7.3.4).

Definition 7.3.7. Eine Matrix $M \in \mathbb{Z}^{m\times n}$ vom Rang m ist in *Hermite-Normalform* falls sie von der Gestalt $[B\ \mathbf{0}]$ ist, wobei $B \in \mathbb{N}^{m\times m}$ eine in $\mathbb{Q}^{m\times m}$ invertierbare Matrix in unterer Dreiecksform ist, in der jeder Diagonaleintrag strikt größer ist als alle anderen Einträge in der gleichen Zeile.

Satz 7.3.8. *Jede Matrix $A \in \mathbb{Z}^{m\times n}$ vom Rang m kann durch unimodulare Spaltenoperationen in eine Matrix in Hermite-Normalform überführt werden.*

Beweis. **Algorithmus, erster Teil.** Angenommen, A ist von der Gestalt

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \tag{7.14}$$

wobei B in unterer Dreiecksform und positiven Einträgen auf der Diagonalen. (Anfänglich ist $B \in \mathbb{Z}^{0\times 0}$.) Es sei $(d_{1,1}, \dots, d_{1,k})$ die erste Zeile von D.

1. Falls $d_{1,i} < 0$, multipliziere Spalte i mit -1. Wir können also ohne Beschränkung der Allgemeinheit annehmen, dass $d_{1,1}, \dots, d_{1,k} \in \mathbb{N}$.

2. Wende Spaltenoperationen an, so dass $d_{1,1} + \cdots + d_{1,k} \in \mathbb{N}$ so klein wie möglich ist.

3. Vertausche Spalten, so dass $d_{1,1} \geq \cdots \geq d_{1,k}$.

Beobachtung 1. $d_{1,1} > 0$ da sonst $d_{1,1} = \cdots = d_{1,k} = 0$, und damit $\mathrm{rg}(A) < m$, im Widerspruch zu unseren Annahmen.
Beobachtung 2. $d_{1,2} = \cdots = d_{1,k} = 0$. Ansonsten, falls $d_{12} > 0$: subtrahiere 2-te Spalte in D von der 1-ten, im Widerspruch zur Minimalität von $d_{1,1} + \cdots + d_{1,k} \in \mathbb{N}$.
Beobachtung 3. Die resultierende Matrix hat die Gestalt in (7.14) für eine Matrix B in Dreiecksform, die um eine Zeile und eine Spalte größer ist als zuvor.

Wenn wir dieses Verfahren endlich oft wiederholen, erhalten wir schließlich eine Matrix der Gestalt $(B|0)$ wobei B Dreiecksmatrix mit positiver Diagonale.

Algorithmus, zweiter Teil. Wir schreiben die Matrix $B = (b_{i,j})_{i \in \{1,\ldots,m\}, j \in \{1,\ldots,m\}}$ weiter um, damit alle Einträge nicht-negativ und jeder Diagonaleintrag $b_{i,i}$ strikt größer ist als alle anderen Einträge $b_{i,j}$ in der gleichen Zeile. Wir gehen Zeile für Zeile in aufsteigender Ordnung vor. Für Zeile i addieren wir zur Spalte $j < i$ ein ganzzahliges Vielfaches der Spalte i, so dass $0 \leq b_{i,j} < b_{i,i}$. Dabei ändern sich für $i' < i$ die Eintrage $b_{i',j}$ nicht, da $b_{i',i} = 0$. Die resultierende Matrix ist in Hermite-Normalform. □

Es folgt also, dass jede rationale Matrix $A \in \mathbb{Z}^{m \times n}$ vom Rang m unimodular ähnlich ist zu einer Matrix H in Hermite-Normalform. Wir nennen H dann *die Hermite-Normalform von A*. Ähnlich wie im Beweis von Satz 7.1.2 kann man zeigen, dass die Hermite-Normalform von A eindeutig ist.

Proposition 7.3.9. *Für $A \in \mathbb{Z}^{n \times n}$ sind die folgenden Aussagen äquivalent.*

1. *A ist unimodular.*

2. *A lässt sich schreiben als Produkt unimodularer Elementarmatrizen.*

3. *A hat eine inverse Matrix in $\mathbb{Z}^{n \times n}$.*

Beweis. 2. $\Rightarrow$ 3. Jede unimodulare Elementarmatrix hat ein Inverses. Das Inverse von A ergibt sich aus den Inversen der unimodularen Elementarmatrizen (3.3).

3. $\Rightarrow$ 1. Wenn A ein Inverses $B \in \mathbb{Z}^{n \times n}$ hat, dann ist $\det(A)$ eine Einheit in $\mathbb{Z}$, denn (mit Satz 4.1.15)

$$1 = \det(E_n) = \det(AB) = \det(A)\det(B).$$

Um die Implikation 1. $\Rightarrow$ 2. zu zeigen, transformieren wir A mit unimodularen Zeilenumformungen in Stufenform, wie bei der Berechnung der Hermite-Normalform in Satz 7.3.8. Die Stufenform muss sogar schon in Dreiecksform sein, denn sonst wäre $\det(A) = 0$ und damit A nicht unimodular. Alle Diagonaleinträge der Dreiecksmatrix müssen aus $\{-1, 1\}$ sein, denn sonst wäre $\det(A)$ keine Einheit in $\mathbb{Z}$ (siehe (4.3)). Wir können also durch weitere unimodulare Zeilenumformungen alle Diagonaleinträge zu 1 machen. Durch unimodulare Spaltentransformationen lassen sich dann alle Einträge ausserhalb der Diagonalen eliminieren, wir erhalten also die Matrix E_n. Also läßt sich A schreiben als Produkt unimodularer Elementarmatrizen. □

Beispiel 7.3.10. Wir betrachten die Matrix

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

Durch unimodulare Spaltenumformungen erhalten wir im ersten Teil des Algorithmus:

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

Im zweiten Teil des Algorithmus wird dies weiter vereinfacht:

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

Die Hermite-Normalform von A ist also $H = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ und für die Matrix $U = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$ gilt $\det(U) = -1$ und $AU = H$. △

Beispiel 7.3.11. Wir betrachten die (invertierbare) Matrix

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

Durch unimodulare Spaltenumformungen erhalten wir:

$$A \xrightarrow[(1)]{s_1 \leftrightsquigarrow s_3} \begin{pmatrix} 1 & 4 & 3 \\ 5 & 1 & 2 \\ 0 & -1 & 4 \end{pmatrix} \xrightarrow[(3)]{s_2 - 4s_1 \rightsquigarrow s_2} \begin{pmatrix} 1 & 0 & 3 \\ 5 & -19 & 2 \\ 0 & -1 & 4 \end{pmatrix} \xrightarrow[(3)]{s_3 - 3s_1 \rightsquigarrow s_3} \begin{pmatrix} 1 & 0 & 0 \\ 5 & -19 & -13 \\ 0 & -1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ C & D \end{pmatrix}.$$

Wir wiederholen das Verfahren nun mit der kleineren Matrix D.

$$D = \begin{pmatrix} -19 & -13 \\ -1 & 4 \end{pmatrix} \xrightarrow[(2)]{-s_1 \rightsquigarrow s_1, -s_2 \rightsquigarrow s_2} \begin{pmatrix} 19 & 13 \\ 1 & -4 \end{pmatrix} \xrightarrow[(1)]{s_1 - s_2 \rightsquigarrow s_1} \begin{pmatrix} 6 & 13 \\ 5 & -4 \end{pmatrix} \xrightarrow[(3)]{s_2 - s_1 \rightsquigarrow s_2} \begin{pmatrix} 6 & 1 \\ 5 & -9 \end{pmatrix}$$
$$\xrightarrow[(1)]{s_1 \leftrightsquigarrow s_2} \begin{pmatrix} 1 & 6 \\ -9 & 5 \end{pmatrix} \xrightarrow[(3)]{s_2 - 6s_1 \rightsquigarrow s_2} \begin{pmatrix} 1 & 0 \\ -9 & -46 \end{pmatrix}$$

Also lässt sich A mit unimodularen Spaltenumformungen in folgende Gestalt bringen:

$$\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & -9 & 46 \end{pmatrix}$$

Im zweiten Teil des Algorithmus wird die Matrix weiter wie folgt reduziert.

$$\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & -9 & 46 \end{pmatrix} \xrightarrow[(3)]{s_1 - 5s_2 \rightsquigarrow s_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -45 & -9 & 46 \end{pmatrix}$$
$$\xrightarrow[(3)]{s_1 + s_3 \rightsquigarrow s_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -45 & 35 & 46 \end{pmatrix} \xrightarrow[(3)]{s_1 + s_3 \rightsquigarrow s_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 35 & 46 \end{pmatrix}. \qquad \triangle$$

7.3.3 Ein polynomieller Algorithmus

Es lässt sich leicht zeigen, dass das Verfahren im Beweis von Satz 7.3.8 nach polynomiell vielen Rechenschritten terminiert.[6] Es kann aber das Problem auftreten, dass die Einträge der Matrizen während der Berechnung sehr groß werden; so groß, dass man sie nicht mehr mit polynomiell vielen Bits abspeichern kann (Beispiel 4.1.26 läßt sich entsprechend anpassen). Um das Problem zu beheben, verwenden wir einen Trick (der laut Schrijver [10] von Domich 1983 in seiner Masterarbeit gefunden wurde), und zeigen damit den folgenden Satz, der 1979 von Kannan und Bachem [7] gezeigt wurde (für Verbesserungen siehe Chou und Collins [3]).

Satz 7.3.12. *Zu einer gegebenen Matrix $A \in \mathbb{Z}^{m\times n}$ vom Rang m lässt sich in polynomieller Rechenzeit eine unimodulare Matrix $U \in \mathbb{Z}^{n\times n}$ und eine Matrix $H \in \mathbb{Z}^{m\times n}$ in Hermite-Normalform berechnen, so dass $AU = H$.*

Beweis. Es sei C eine beliebige quadratische Untermatrix von A vom Rang m; es ist klar, dass sich so ein C und $s := |\det(C)|$ in polynomieller Zeit berechnen lässt (z.B. mit dem gaußschen Algorithmus). Wir betrachten nun die Matrix

$$A' := [A \mid sE_m].$$

Behauptung. Die zusätzlichen Spalten sind ganzahlige Linearkombinationen der Spalten von C, und damit auch der Spalten von A.

Die Inverse von C berechnet sich nach Satz 4.1.33 durch $C^{-1} = \frac{C^{\#}}{\det C}$; da C ganzzahlig ist, ist auch $C^{\#}$ ganzzahlig. Also ist $\det(C)C^{-1}$ ganzzahlig. Es gilt

$$C(\det(C)C^{-1}) = \det(C)E_m \in \{sE_m, -sE_m\},$$

und die Behauptung ist bewiesen.

Die Behauptung impliziert, dass sich die Hermite-Normalform von A aus der für A' ergibt durch Entfernen von überschüssigen Spalten, die nur 0 enthalten. Wir folgen nun dem Algorithmus aus dem Beweis von Satz 7.3.8 mit der folgenden Modifikation. Falls beim Algorithmus eine Spalte erzeugt wird, deren i-ter Eintrag den Wert s überschreitet,

[6] Der interessanteste Teil der Laufzeitanalyse ist Schritt 2 im ersten Teil. Die Analyse hier ist ähnlich zur Analyse des euklidischen Algorithmus, der in der Fortsetzungsvorlesung AL10 ausführlich behandelt wird.

dann ziehen wir die i-te Spalte der Matrix sE_m ab (wir rechnen bei den Koeffizienten also *'modulo s'*).

Es ist klar, dass das Produkt der Elementarmatrizen für die unimodularen Zeilenumformungen die gesuchte unimodulare Matrix U liefert, und dass auch die Einträge dieser Matrix nicht zu groß werden. □

7.3.4 Ganzzahlige Lösungen für lineare Gleichungssysteme

Der folgende Satz hat viele Anwendungen in der theoretischen Informatik.

Satz 7.3.13. *Es gibt einen Algorithmus mit polynomieller Laufzeit, der für gegebenes $A \in \mathbb{Q}^{m \times n}$ und $b \in \mathbb{Q}^m$ entscheidet, ob $Ax = b$ eine ganzzahlige Lösung besitzt.*

Beweis. Zunächst stellen wir fest, dass wir ohne Beschränkung der Allgemeinheit annehmen können, dass A und b ganzzahlig ist, da die Multiplikation jeder Zeile der erweiterten Koeffizientenmatrix $(A|b)$ mit dem kleinsten gemeinsamen Vielfachen aller Koeffizienten der Zeile die Anzahl der Bits der Zahlen nur linear vergößert, und den Lösungsraum nicht verändert.

Als nächstes entscheiden wir mit dem gaußschen Algorithmus aus Abschnitt 3.3.4, ob $Ax = b$ eine *rationale* Lösung besitzt. Falls nein, dann sicherlich auch keine ganzzahlige. Falls ja, wählen wir eine maximale Menge linear unabhängiger Zeilen von A aus (das geht ebenfalls mit Hilfe des gaußschen Algorithmus). Das resultierende Untersystem hat die gleiche Lösungsmenge, und wir arbeiten daher im folgenden mit diesem Untersystem anstatt mit A. Wir nehmen also an, dass A vom Rang m ist.

Als nächstes berechnen wir mit dem Verfahren von Satz 7.3.12 in polynomieller Zeit eine unimodulare Matrix $U \in \mathbb{Z}^{n \times n}$, so dass $AU = [B\ \mathbf{0}]$, für $B \in \mathbb{N}^{m \times m}$ vom Rang m, die Hermite-Normalform von A ist. Dann hat B eine Inverse $B^{-1} \in \mathbb{Q}^{m \times m}$, die sich effizient berechnen lässt (siehe Abschnitt 3.2.7).

Behauptung: $Ax = b$ hat genau dann ganzzahlige Lösung, wenn $B^{-1}b$ ganzzahlig ist.

Falls $B^{-1}b$ ganzzahlig ist, dann ist $s := U(B^{-1}b, 0, \dots, 0)^\top \in \mathbb{Z}^n$ eine ganzzahlige Lösung von $Ax = b$, denn

$$As = AU(B^{-1}b, 0, \dots, 0)^\top = [B\ \mathbf{0}](B^{-1}b, 0, \dots, 0)^\top = b.$$

Umgekehrt, sei s ganzzahlig ist mit $As = b$. Nach Proposition 7.3.9 hat U eine Inverse $U^{-1} \in \mathbb{Z}^{n \times n}$. Dann ist $U^{-1}s$ eine ganzzahlige Lösung von $[B\ \mathbf{0}]\,x = b$, da

$$[B\ \mathbf{0}]U^{-1}s = As = b.$$

Insbesondere ist dann $B^{-1}(b)$ ganzzahlig. □

Beispiel 7.3.14. Wir betrachten das lineare diophantische Gleichungssystem $Ax = b$ für

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Wir wir in Beispiel 7.3.10 gesehen haben, gilt $AU = H$ für

$$H = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}, \quad U = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}.$$

Da

$$H^{-1}b = \underbrace{\begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}}_{\text{Beispiel 4.1.34}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$$

nicht ganzzahlig ist, hat das System

$$\begin{aligned} x + 2y &= 1 \\ 3x + 4y &= 2 \end{aligned}$$

keine ganzzahlige Lösung (aber die fraktionale Lösung $x = 0$, $y = \frac{1}{2}$).

Für $b = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ dagegen ist

$$H^{-1}b = \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

ganzzahlig, und tatsächlich ist

$$UH^{-1}b = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

eine ganzzahlige Lösung für

$$\begin{aligned} x + 2y &= -1 \\ 3x + 4y &= -1. \end{aligned} \qquad \triangle$$

Übung 48. Charakterisieren Sie die Teilmenge von $\mathbb{Q}^3$, die aus allen Tripeln (a_1, a_2, b) besteht, so dass ganze Zahlen $x_1, x_2 \in \mathbb{Z}$ existieren mit $a_1x_1 + a_2x_n = b$.

Übung 49. Zeigen Sie, dass sich lineare Gleichungssysteme über dem Ring $\mathbb{Z}_n$ (siehe Abschnitt 4.2.1) für beliebiges $n \in \{2, 3, 4, \dots\}$ in polynomieller Zeit lösen lassen (die Aufgabe ist besonders interessant, wenn n nicht prim ist!).

Hinweis. Ein möglicher Lösungsansatz besteht darin, die Aufgabe auf Lösbarkeit in $\mathbb{Z}$ zu reduzieren. Für $a \in \mathbb{Z}$ schreiben wir $[a]$ für die Restklasse von a modulo n. Es seien $a_0, a_1, \dots, a_n \in \{0, 1, \dots, n-1\}$. Zeigen Sie zunächst, dass es genau dann Elemente $x_1, \dots, x_n \in \mathbb{Z}_n$ gibt mit $[a_1]x_1 + \dots + [a_n]x_n = [a_0]$, wenn es Elemente $y, z_1, \dots, z_n \in \mathbb{Z}$ gibt mit

$$a_1z_1 + \dots + a_nz_n = a_0 + \underbrace{y + \dots + y}_{n \text{ mal}}.$$

Übung 50. Finden Sie einen Algorithmus mit polynomieller Laufzeit, der für gegebene lineare Gleichungssysteme $Ax = b$ und $Cx = d$ mit $A, B \in \mathbb{Q}^{m \times n}$ und $b, d \in \mathbb{Q}^m$ entscheidet, ob jede ganzzahlige Lösung von $Ax = b$ auch eine Lösung von $Cx = d$ ist.

7.3.5 Die Smith-Normalform

In diesem Kapitel betrachten wir eine Normalform von Matrizen bis auf unimodulare Äquivalenz (Definition 7.3.1), nämlich die *Smith-Normalform* (Satz 7.3.20). Viele der Ideen zur Berechnung der Hermite-Normalform sind auch für die Berechnung der Smith-Normalform nützlich. Die Smith-Normalform kann dazu verwendet werden, um die Frobenius-Normalform zu berechnen, und hat weitere Anwendungen in der Algebra. Unsere Anwendungen der Smith-Normalform verwenden Matrizen über dem Polynomring $\mathbb{K}[X]$, und wir beschränken uns ab jetzt auf diesen Fall. Hier sind die Einheiten gerade die Elemente von $\mathbb{K} \setminus \{0\}$.

Bemerkung 7.3.15. Die Smith-Normalform existiert auch für den Ring $\mathbb{Z}$, und allgemein für *Dedekind Ringe*, also insbesondere also für *Hauptidealringe* und damit auch für *euklidische Ringe*; diese Begriffe werden allerdings erst in der Vorlesung Algebra – grundlegende Konzepte (AL10) behandelt.

Die *unimodularen Elementarmatrizen* sind hier die Permutationsmatrizen, die Elementarmatrizen für die Multiplikation einer Spalte mit einer Einheit, und die Elementarmatrizen für die Addition von rs zu s', wobei s und s' verschiedene Spalten und $r \in R$ ist.

Lemma 7.3.16. *Sei $A \in \mathbb{K}[X]^{n\times n}$. Dann sind folgende Aussagen äquivalent.*

1. *A ist unimodular.*
2. *A kann geschrieben werden als Produkt von unimodularen Elementarmatrizen (Definition 7.3.6).*
3. *A hat ein Inverses in $\mathbb{K}[X]^{n\times n}$, d.h., es gibt ein $B \in \mathbb{K}[X]^{n\times n}$, so dass*

$$AB = BA = E_n.$$

Beweis. Der Beweis von 2. $\Rightarrow$ 3. und 3. $\Rightarrow$ 1. geht genau wie der Beweis von 7.3.9. Auch die Implikation 1. $\Rightarrow$ 2. geht wie dort: wir transformieren A mit unimodularen Zeilenumformungen in Stufenform, analog zum Algorithmus bei der Berechnung der Hermite-Normalform in Satz 7.3.8. Die Stufenform muss sogar schon in Dreiecksform sein, denn sonst wäre $\det(A) = 0$ und damit A nicht unimodular. Alle Diagonaleinträge der Dreiecksmatrix müssen aus $\mathbb{K} \setminus \{0\}$ sein, denn sonst wäre $\det(A)$ keine Einheit in $\mathbb{K}[X]$ (siehe (4.3)). Wir können also durch weitere unimodulare Zeilenumformungen alle Diagonaleinträge zu 1 machen. Durch unimodulare Spaltentransformationen lassen sich dann alle Einträge ausserhalb der Diagonalen eliminieren, wir erhalten also die Matrix E_n; also läßt sich A schreiben als Produkt unimodularer Elementarmatrizen. $\square$

Definition 7.3.17. Zwei Matrizen $A, B \in R^{m\times n}$ heißen *unimodular äquivalent* falls es unimodulare Matrizen $P \in R^{m\times m}$ und $Q \in R^{n\times n}$ gibt, so dass $PAQ = B$.

Um auch gleich die Eindeutigkeit der Smith-Normalform nachzuweisen, benötigen wir einige weitere, recht natürliche Definitionen. Falls $\varphi_1, \dots, \varphi_k \in \mathbb{K}[X]$, so existiert ein eindeutiges normiertes Polynom von maximalem Grad, welches alle $\varphi_1, \dots, \varphi_k$ teilt, welches wir den *größten gemeinsamen Teiler* $\mathrm{ggT}(\varphi_1, \dots, \varphi_k)$ von $\varphi_1, \dots, \varphi_k$ nennen.

Definition 7.3.18. Sei $A \in \mathbb{K}[X]^{m \times n}$ und $k \in \{1, \dots, \min(m,n)\}$. Ein *$k$-Minor von A* ist die Determinante einer $k \times k$ Untermatrix von A. Der größte gemeinsame Teiler aller k-Minoren von A wird *Determinantenteiler von A* genannt, und mit $d_k(A)$ bezeichnet; zudem definieren wir $d_0(A) := 1$. Der *Determinantenrang von A* ist die größte natürliche Zahl k, so dass A einen k-Minor ungleich 0 besitzt, und wird mit $r(A)$ bezeichnet.

Lemma 7.3.19. *Es seien $A, B \in \mathbb{K}[X]^{m \times n}$ unimodular äquivalent. Dann gilt $r(A) = r(B)$ und für alle $i \in \{1, \dots, r\}$ gilt $d_i(A) = d_i(B)$.*

Beweis. Seien $P, Q \in \mathbb{K}[X]^{n \times n}$ unimodular so dass $PAQ = B$. Die Zeilen von PA sind Linearkombinationen der Zeilen von A. Also ist für $k \leq \min(m,n)$ jeder k-Minor von PA eine Linearkombination der k-Minoren von A. Analog ist jede Spalte von $(PA)Q$ eine Linearkombination der Spalten von PA, also ist jeder k-Minor von $(PA)Q$ eine Linearkombination der k-Minoren von PA, und folglich der k-Minoren von A. Daher

- sind für $k > r(A)$ alle k-Minoren von B gleich 0, also $r(B) \leq r(A)$,
- und für $k \leq r(A)$ gilt $d_k(A) | d_k(B)$.

Mit dem gleichen Argument angewandt auf $P^{-1}BQ^{-1}A$ zeigt man $r(A) \leq r(B)$ und $d_k(B) | d_k(A)$. Da der ggT normiert ist, folgt $d_k(A) = d_k(B)$. □

Satz 7.3.20 (Smith-Normalform)**.** *Jede Matrix $A \in \mathbb{K}[X]^{n \times n}$ ist unimodular äquivalent zu einer eindeutigen Matrix der Gestalt*

$$\begin{pmatrix} \varphi_1 & & & & & 0 \\ & \ddots & & & & \\ & & \varphi_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

wobei $\varphi_1, \dots, \varphi_m \in \mathbb{K}[X]$ normiert, so dass $\varphi_i | \varphi_j$ für alle $i, j \in \{1, \dots, m\}$ mit $i < j$. Die Zahl m und die Polynome $\varphi_1, \dots, \varphi_m \in \mathbb{K}[X]$ berechnen sich wie folgt:

- *$m = r(A)$ ist der Determinantenrang von A, und*
- *$\varphi_i = \frac{d_i(A)}{d_{i-1}(A)}$ für $i \in \{1, \dots, m\}$.*

Das folgende Beispiel wird im nächsten Abschnitt eine wichtige Rolle spielen.

Beispiel 7.3.21. Es sei $\varphi \in \mathbb{K}[X]$ ein normiertes Polynom vom Grad n. Dann ist die Smith-Normalform von $XE_n - Z_\varphi$ von der Gestalt

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & \varphi \end{pmatrix}. \tag{7.15}$$

Wie wir in Beispiel 7.2.1 gesehen haben, lässt sich $XE_n - Z_\varphi$ durch unimodulare Zeilenumformungen in folgende Gestalt bringen:

$$\begin{pmatrix} -1 & X & 0 & 0 & & \alpha_1 \\ 0 & -1 & X & 0 & & \alpha_2 \\ & \ddots & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & X & \alpha_{n-2} \\ \vdots & & & \ddots & -1 & X + \alpha_{n-1} \\ 0 & \cdots & & & 0 & \alpha_0 + X\alpha_1 + X^2\alpha_2 + \cdots + \alpha_{n-1}X^{n-1} + X^n \end{pmatrix}$$

Durch unimodulare Spaltenumformungen können dann

- die Einträge -1 auf der Diagonalen zu 1 geändert werden, und
- die Einträge mit X oberhalb der Diagonalen zu 0 geändert werden,
- die ersten $n-1$ Einträge der letzten Spalte zu 0 geändert werden.

Die resultierende Matrix hat die Form (7.15). △

Beweis von Satz 7.3.20. Wir geben ein Verfahren an, welches nach einer endlichen Anzahl von Schritten terminiert, und unimodulare Matrizen $P \in \mathbb{K}[X]^{n\times n}$ und $Q \in \mathbb{K}[X]^{m\times m}$ und ein normiertes Polynom $\varphi_1 \in \mathbb{K}[X]$ liefert, so dass

$$PAQ = \begin{pmatrix} \varphi_1 & 0 \\ 0 & B \end{pmatrix} \tag{7.16}$$

wobei $B \in \mathbb{K}[X]^{(n-1)\times(m-1)}$ und φ_1 teilt alle Einträge von B.

1. Falls $A = \mathbf{0}$ dann sind wir fertig; wir nehmen also im Folgenden an, dass $A = (a_{ij}) \neq \mathbf{0}$.
2. Wende Zeilen- und Spaltenvertauschungen an, so dass $a_{1,1} \neq 0$, und so dass $a_{1,1} \in \mathbb{K}[X]$ unter allen Einträgen von A, die nicht 0 sind, den *kleinsten* Grad besitzt.
3. Schreibe jeden Eintrag $a_{1,j}$ in der ersten Zeile als $a_{1,j} = q_{1,j}a_{1,1} + r_{1,j}$ für $q_{1,j}, r_{1,j} \in \mathbb{K}[X]$ mit $\mathrm{grad}(r_{1,j}) < \mathrm{grad}(a_{1,1})$ (Polynomdivision), und führe folgende unimodulare Spaltenumformung durch: ziehe $q_{1,j}a_{*1}$ von der j-ten Spalte a_{*j} von A ab, so dass danach $a_{1,j} = r_{1,j}$.

4. Analog zum vorigen Schritt verfahren wir mit den Spalten anstatt der Zeilen.

5. Falls ein Eintrag von A strikt kleineren Grad hat als $a_{1,1}$, gehe zu Schritt 1.

6. Ansonsten: $a_{1,1} \neq 0$; alle anderen Einträge der ersten Zeile und ersten Spalte sind 0; und alle anderen Einträge von A, die nicht 0 sind, haben größeren Grad als $a_{1,1}$. Falls $a_{1,1}$ alle anderen Einträge von A teilt, so können wir durch Multiplikation der ersten Zeile mit einer Einheit in $\mathbb{K}[X]$ erreichen, dass $a_{1,1}$ normiert ist. Also ist A von der gewünschten Gestalt und das Verfahren bricht ab.

7. Ansonsten, falls $a_{1,1}$ den Eintrag $a_{i,j}$ nicht teilt, schreibe a_{ij} als $a_{i,j} = qa_{1,1} + r$ für $q, r \in \mathbb{K}[X]$ mit $0 \neq \mathrm{grad}(r) < \mathrm{grad}(a_{1,1})$. Addiere dann die erste Zeile von A zur i-ten Zeile. Subtrahiere das q-fache der ersten Spalte von der j-ten Spalte. Wir erhalten eine Matrix mit Eintrag r an der Stelle i, j. Fahre fort mit Schritt 1.

Wir wenden dieses Verfahren nun induktiv auf B anstatt auf A an. Bei Abbruch des Verfahrens ist die resultierende Matrix in Smith-Normalform. Das Verfahren terminiert nach endlich (polynomiell) vielen Schritten: zunächst ist die Tiefe der Induktion höchstens n. Im Induktionsschritt wird höchstens n mal zu Schritt 1 zurückgesprungen, da jedesmal der Grad von $a_{1,1}$ strikt kleiner wird.

Da alle auftretenden Umformungen im Verfahren durch Multiplikation mit unimodularen Elementarmatrizen von links oder von rechts beschrieben werden können, folgt, dass die resultierende Matrix S unimodular äquivalent ist zu A.

Die Eindeutigkeit folgt aus der zweiten Aussage. Nach Lemma 7.3.19 gilt

$$r(A) = r(S) = m$$

und für $i \in \{1, \ldots, m\}$ gilt

$$d_i(A) = d_i(S) = \varphi_1 \cdots \varphi_i,$$

und damit ist $\varphi_i = \frac{d_i(A)}{d_{i-1}(A)}$. □

Bemerkung 7.3.22. Wie bei der Berechnung der Hermite-Normalform besteht bei der algorithmischen Berechnung der Smith-Normalform die Gefahr, dass Einträge der Matrizen zu groß werden. Mit ähnlichen Methoden wie in Abschnitt 7.3.3 lässt sich das vermeiden, so dass auch die Smith-Normalform einer Matrix $A \in \mathbb{Q}[X]^{n \times n}$ in polynomieller Zeit berechnet werden kann [3, 7].

7.3.6 Zusammenhang Smith-Normalform und Frobenius-Normalform

Die Smith-Normalform kann dazu verwendet werden, um die Frobenius-Normalform (aus Abschnitt 7.2.5) von $A \in \mathbb{K}^{n \times n}$ zu berechnen! Wir berechnen dazu die Smith-Normalform

S der Matrix $XE_n - A$ mit Einträgen aus $\mathbb{K}[X]$. Diese ist von der Gestalt

$$\begin{pmatrix} \varphi_1 & & & & & 0 \\ & \ddots & & & & \\ & & \varphi_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \tag{7.17}$$

wobei $\varphi_1, \ldots, \varphi_m \in \mathbb{K}[X]$ normierte Polynome sind, so dass $\varphi_i | \varphi_j$ für alle $i, j \in \{1, \ldots, m\}$ mit $i < j$.

Bemerkung 7.3.23. Es gilt $m = n$ und $\prod_{i \in \{1,\ldots,n\}} \varphi_i = \chi_A$. Das folgt direkt aus der Definition von $\chi_A = \det(XE_n - A)$ und der Beobachtung, dass

- sowohl χ_A als auch $\prod_{i \in \{1,\ldots,n\}} \varphi_i$ normiert sind, und
- unimodulare Zeilen- und Spaltenumformungen die Determinante bis auf Multiplikation mit einer Einheit nicht ändern (siehe Proposition 4.1.3 für Zeilenumformungen, und kombiniere mit Proposition 4.1.14 für Spaltenumformungen).

Satz 7.3.24. *Falls $A \in \mathbb{K}^{n \times n}$, $\varphi_1, \ldots, \varphi_n \in \mathbb{K}^{n \times n}$ die Polynome aus (7.17) der Smith-Normalform. Dann gilt $\varphi_1 = \cdots = \varphi_\ell = 1$ für $\ell = \sum_{i=1}^n \operatorname{grad}(\varphi_i)$, und $\varphi_{\ell+1}, \ldots, \varphi_n$ sind die Ähnlichkeitsinvarianten von A. Die Frobenius-Normalform F von A hat also die Gestalt*

$$F = \begin{pmatrix} Z_{\varphi_n} & & 0 \\ & \ddots & \\ 0 & & Z_{\varphi_{\ell+1}} \end{pmatrix}.$$

Beweis. Es seien $\psi_1, \ldots, \psi_k$ die Ähnlichkeitsinvarianten von A, so dass $\varphi_i | \varphi_{i-1}$ für alle $i \in \{2, \ldots, k\}$. Wenn ψ_i für $i \in \{1, \ldots, k\}$ vom Grad d_i ist, dann haben wir in Beispiel 7.3.21 gesehen, dass $XE_{d_i} - Z_{\psi_i}$ unimodular äquivalent ist zu

$$C_i := \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \psi_i \end{pmatrix}.$$

Man sieht daher leicht, dass sich $XE_n - A$ mit unimodularen Zeilen- und Spaltenumformungen in die Gestalt

$$\begin{pmatrix} C_k & & 0 \\ & \ddots & \\ 0 & & C_1 \end{pmatrix}$$

bringen läßt. Durch Vertauschen von Zeilen und Spalten erhalten wir dann die Matrix

$$\begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \psi_k & & \\ & & & & \ddots & \\ 0 & & & & & \psi_1 \end{pmatrix}.$$

Da $1|\phi_i$ und $\psi_i|\psi_j$ für $i,j \in \{1,\dots,k\}$ mit $i < j$, ist diese Matrix in SNF, und aus der Eindeutigkeit der Smith-Normalform (Satz 7.3.20) folgt, dass $k = n - \ell$ und $\psi_1 = \varphi_{\ell+1},\dots,\psi_k = \varphi_n$. Die Aussage folgt aus der Eindeutigkeit der Frobenius-Normalform (Satz 7.2.31). □

Definition 7.3.25. Zwei Matrizen $A, B \in R^{n\times n}$ heißen *unimodular ähnlich* falls es eine unimodulare Matrix $S \in R^{n\times n}$ gibt mit $SAS^{-1} = B$.

Nachrechnen: unimodulare Ähnlichkeit ist eine Äquivalenzrelation.

Korollar 7.3.26 (Satz von Frobenius)**.** *Seien $A, B \in \mathbb{K}^{n\times n}$. Dann sind die folgenden drei Aussagen logisch äquivalent.*

1. *A und B sind ähnlich;*
2. *$XE_n - A$ und $XE_n - B$ sind unimodular ähnlich;*
3. *$XE_n - A$ und $XE_n - B$ sind unimodular äquivalent.*

Beweis. Angenommen, A und B sind ähnlich. Dann gibt es eine invertierbare Matrix $C \in \mathbb{K}^{n\times n}$ so dass $A = CBC^{-1}$, und es gilt

$$XE_n - A = CXE_nC^{-1} - CBC^{-1} = C(XE_n - B)C^{-1}.$$

Die Implikation 2. ⇒ 3. ist trivial. Für die Implikation 3. ⇒ 1 verwenden wir, dass $XE_n - A$ und $XE_n - B$ die selbe Smith-Normalform S haben (Satz 7.3.20), und damit nach Satz 7.3.24 auch die gleichen Ähnlichkeitsinvarianten; die Aussage folgt dann aus der Eindeutigkeit der Frobenius-Normalform (Satz 7.2.31). □

Bemerkung 7.3.27. Korollar 7.3.26 hat einen direkten Beweis, der ohne die Smith-Normalform auskommt; siehe Serre [11] (Theorem 6.3.2).

Übung 51. Seien $A \in \mathbb{Q}[X]^{n\times n}$ und $b \in \mathbb{Q}[X]^n$. Erläutern Sie, wie man in polynomieller Zeit feststellen kann, ob es ein $x \in \mathbb{Q}[X]^n$ gibt, so dass $Ax = b$.

Übung 52. Die ganzzahligen Lösungen eines linearen Gleichungssystems lassen sich auch mit Hilfe der Smith-Normalform bestimmen. Sei $Ax = b$ für $A \in \mathbb{Z}^{m\times n}$ und $b \in \mathbb{Z}^m$ das gegebene lineare Gleichungssystem. Seien $P \in \mathbb{Z}^{m\times m}$ und $Q \in \mathbb{Z}^{n\times n}$ unimodular, so dass $S = PAQ$ in Smith-Normalform. Zeigen Sie die folgenden Aussagen:

- $Ax = b$ und $PAx = Pb$ haben die gleichen ganzzahligen Lösungen.
- $Ax = b$ hat genau dann eine ganzzahlige Lösung, wenn $Sy = Pb$ eine ganzzahlige Lösung hat. (Tipp: setze $x = Qy$).
- Die ganzzahligen Lösungen von $Sy = Pb$ lassen sich einfach ermitteln.

Übung 53 (Christian Zschalig). Finden Sie zwei ganzzahlige Matrizen, die unimodular äquivalent und spaltenäquivalent, aber nicht unimodular spaltenäquivalent sind.

Tipp. Es gibt bereits Beispiele aus $\mathbb{Z}^{2\times 2}$.

Übung 54. Kann man mit Hilfe der Smith-Normalform auch die Jordan-Normalform berechnen (so sie denn existiert)?[7]

[7] Vielen Dank an die Teilnehmer:innen der VL im SS24 für die Idee zu dieser Übung.

Kapitel 8

Euklidische und unitäre Vektorräume

Euklidische Vektorräume sind Vektorräume mit einem ausgezeichneten Skalarprodukt, und wurden bereits in Abschnitt 6.1.3 eingeführt. Wir wenden uns diesen jetzt systematischer zu. Ein großer Teil der Theorie kann parallel für Vektorräume über $\mathbb{C}$ entwickelt werden; die Modifikationen der Aussagen für die komplexe Variante wird immer grün hervorgehoben.

8.1 Bilinearformen

Wichtiger Spezialfall: Skalarprodukt (Abschnitt 6.1.3). Sei V ein $\mathbb{K}$-Vektorraum. Wir schreiben im folgenden wieder x anstatt $\vec{x}$ für Elemente von V.

Definition 8.1.1. Eine Abbildung

$$B\colon V \times V \to \mathbb{K}$$

heißt *Bilinearform (auf V)* wenn gilt: für alle $u, v, u_1, u_2, v_1, v_2 \in V$ und $\alpha \in \mathbb{K}$:

- Linearität in der ersten Stelle: $B(u_1 + u_2, v) = B(u_1, v) + B(u_2, v)$ und $B(\alpha u_1, v) = \alpha B(u_1, v)$.
- Linearität in der zweiten Stelle: $B(u, v_1 + v_2) = B(u, v_1) + B(u, v_2)$ und $B(u, \alpha v) = \alpha B(u, v)$.

Folgerung:

$$B(\mathbf{0}, v) = 0 = B(u, \mathbf{0})$$
$$B(-u, v) = -B(u, v) = B(u, -v)$$

Für $\mathbb{K} = \mathbb{C}$: eine Abbildung $B\colon V \times V \to \mathbb{C}$ heißt *Semibilinearform* (oder *Sesquilinearform*[1]) falls gilt

[1] Sesqui: lateinisch für eineinhalb.

- Linearität in der ersten Stelle (wie oben)
- *Semilinearität* in der zweiten Stelle:

$$B(u, v_1 + v_2) = B(u, v_1) + B(u, v_2)$$
$$B(u, \beta v) = \bar{\beta} B(u, v)$$

wobei $\bar{\beta} := a - bi$ für $\beta = a + bi$ die *konjugiert komplexe Zahl* zu $\beta \in \mathbb{C}$.

Bemerkung 8.1.2. In der Physik andere Konvention: das *erste* Argument erfüllt Semilinearität.

Bemerkung 8.1.3. Konjugation definiert ein *Automorphismus* des Körpers $(\mathbb{C}; +, *)$, d.h., eine bijektive Abbildung $\mathbb{C} \to \mathbb{C}$, die verträglich ist mit Addition und Multiplikation:

$$\begin{aligned} \overline{(a+bi)+(c+di)} &= (a+c)-(b+d)i = \overline{(a+bi)} + \overline{(c+di)} \\ \overline{(a+bi)*(c+di)} &= \overline{ac+(bc+ad)i-bd} = ac-adi-bci-bd \\ &= (a-bi)*(c-di) = \overline{(a+bi)} * \overline{(c+di)} \end{aligned}$$

Definition 8.1.4. Eine (Semi-) Bilinearform $B\colon V \times V \to \mathbb{K}$ heißt

- *nicht ausgeartet* falls

$$\forall u \neq \mathbf{0}\ \exists v \in V\colon B(u, v) \neq 0$$
$$\forall v \neq \mathbf{0}\ \exists u \in V\colon B(u, v) \neq 0$$

- *symmetrisch* falls

$$\forall u, v \in V:\ B(u, v) = B(v, u)$$

- *hermitisch* falls $\mathbb{K} = \mathbb{C}$ und

$$\forall u, v \in V : B(u, v) = \overline{B(v, u)}$$

- *positiv definit* falls $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und

$$\forall u \in V \setminus \{\mathbf{0}\} : B(u, u) \in \mathbb{R} \text{ und } B(u, u) > 0.$$

- *Skalarprodukt* falls $\mathbb{K} = \mathbb{R}$ (bzw. $\mathbb{K} = \mathbb{C}$) und B positiv definit und symmetrisch (bzw. hermitisch).

Beispiele für (Semi-) Bilinearformen.

1. Das Standardskalarprodukt im $\mathbb{R}^n$ (siehe Abschnitt 6.1.3):

$$B(x, y) := x * y = x^\top y = x_1 y_1 + \cdots + x_n y_n$$

2. das *Lorentz-Produkt* (Relativitätstheorie) auf $V = \mathbb{R}^4$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ t \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ s \end{pmatrix} \qquad \text{(Raum- und Zeitkoordinaten)}$$

$$B(x,y) := x_1y_1 + x_2y_2 + x_3y_3 - c^2ts = x^\top Ay^\top$$

wobei (c: Lichtgeschwindigkeit)

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

Kein Skalarprodukt, da nicht positiv definit.

3. Sei V der Vektorraum aller auf dem Intervall $[a,b] \subseteq \mathbb{R}$ stetigen Funktionen $f\colon \mathbb{R} \to \mathbb{R}$. Dann ist
$$B(f,g) := \int_a^b f(x)g(x)\,\mathrm{d}x$$
ein Skalarprodukt (Beispiel 6.1.2).

Allgemeine Sätze für Skalarprodukte (so wie z.B. Cauchy-Schwartz) gelten für alle diese Beispiele und brauchen nicht immer neu bewiesen zu werden.

8.1.1 Bilinearformen und Matrizen

Sei V ein $\mathbb{K}$-Vektorraum mit der Basis $C = (v_1, \dots, v_n)$, und $B\colon V \times V \to \mathbb{K}$ eine (Semi-) Bilinearform auf V. Die Matrix

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

mit $a_{ij} = B(v_i, v_j)$ heißt *Gramsche Matrix*[2] der Bilinearform. (Hängt von C ab!)

Durch die Gramsche Matrix ist die Bilinearform eindeutig festgelegt: seien $u = \sum_{i=1}^n \alpha_i v_i$ und $v = \sum_{i=1}^n \beta_i v_i$. Dann ist

$$\begin{aligned} B(u,v) &= B\left(\sum \alpha_i v_i, \sum \beta_i v_i\right) = \sum_{i,j} \alpha_i\beta_j B(v_i, v_j) \\ &= \sum_{i,j} \alpha_i a_{ij} \beta_j = (\alpha_1, \dots, \alpha_n) A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. \end{aligned} \tag{8.1}$$

[2] Jórger Perdersen Gram (1850-1916), dänischer Versicherungsmathematiker.

Spezialfall $V = \mathbb{K}^n$, $C = (e_1, \dots, e_n)$ Standardbasis:

$$B(u, v) = u^\top Av \tag{8.2}$$

Standardbilinearform auf $\mathbb{K}^n$: $A = E_n$, d.h., $B(u, v) = u^\top v$ (vgl. Standardskalarprodukt).

Umgekehrt gilt: für jede Matrix $A \in \mathbb{K}^{n\times n}$ ist durch (8.1) bzw. (8.2) eine Bilinearform auf V bzw. $\mathbb{K}^n$ gegeben.

Zusammenhang Eigenschaften von Bilinearformen und Matrizeneigenschaften: sei $B\colon V \times V \to \mathbb{K}$ Bilinearform und A die Gramsche Matrix von B (bzgl. irgendeiner Basis). Dann gilt

$$\begin{aligned}
B \text{ symmetrisch} &\iff A \textit{ symmetrisch, d.h., } A = A^\top \\
B \text{ hermitisch} &\iff A \textit{ hermitisch, d.h., } A = \bar{A}^\top \\
B \text{ positiv definit} &\iff A \textit{ positiv definit, d.h., } \forall x \neq \mathbf{0} : 0 < x^\top Ax \quad (\text{falls } \mathbb{K} = \mathbb{R} \text{ oder } \mathbb{C}) \\
B \text{ nicht ausgeartet} &\iff \operatorname{rg}(A) = n
\end{aligned}$$

Denn: es gibt kein $v \in V$ mit $u^\top Av \neq 0$ genau dann, wenn $u^\top A = \mathbf{0}$. Es gibt $u \in V \setminus \{0\}$ mit $u^\top A = \mathbf{0}$ genau dann, wenn A nicht invertierbar ist, also wenn $\operatorname{rg}(A) \neq n$.

Nun ein Vorgriff auf Abschnitt 8.3.5 (Hauptachsentransformation). Minoren (Determinanten von quadratischen Untermatrizen) haben wir bereits in Definition 7.3.18 kennengelernt. Der k-te *Hauptminor* einer Matrix ist die Determinanten der Untermatrix, die aus den ersten k Elementen der ersten k Zeilen von A besteht.

Proposition 8.1.5. *Sei B eine symmetrische Bilinearform B von $\mathbb{R}^n$ (oder $\mathbb{C}^n$). Dann ist B genau dann ein Skalarprodukt, wenn alle Hauptminoren der Gramschen Matrix A von B (reell und) positiv sind.*

Beweis. Kommt in Abschnitt 8.3.5. □

Beispiel 8.1.6. Die Diagonalmatrix

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \in \mathbb{R}^{n\times n}$$

ist genau dann Gramsche Matrix eines Skalarprodukts, wenn alle Eigenwerte λ_i positiv sind. Das zugehörige Skalarprodukt auf $\mathbb{R}^n$ ist dann

$$B(u, v) = u^\top Av = \lambda_1 u_1 v_1 + \cdots + \lambda_n u_n v_n \tag{8.3}$$

("gewichtetes Standardskalarprodukt").

Beweis: offenbar mit Proposition 8.1.5.

Direkter Beweis:

- Der Ausdruck in (8.3) ist sicher symmetrisch, und positiv definit falls alle $\lambda_1, \dots, \lambda_n$ positiv sind.
- Falls aber $\lambda_i \le 0$ für ein $i \le n$, dann ist $B(e_i, e_i) = \lambda_i \le 0$, und damit ist B nicht positiv definit. △

8.1.2 Zusammenhang zwischen Bilinearformen

"Kennt man eine, kennt man alle."
Bilinearformen unterscheiden sich nur durch einen Endomorphismus, genauer:

Satz 8.1.7. *Sei $B: V \times V \to \mathbb{K}$ eine nicht ausgeartete Bilinearform auf V und $f: V \to V$ lineare Abbildung (Endomorphismus). Dann ist*

$$B'(u, v) := B(f(u), v) \tag{8.4}$$

ebenfalls eine Bilinearform, und jede Bilinearform B' ensteht aus B auf diese Weise (für geeignetes f).

Für $V = \mathbb{K}^n$ lässt sich (8.4) schreiben als

$$B'(u, v) = B(Cu, v)$$

für ein $C \in \mathbb{K}^{n \times n}$.

Beweis. Sei $V = \mathbb{K}^n$. Sei A' die Gramsche Matrix von B' und A die Gramsche Matrix von B. Dann ist A invertierbar, da B nicht ausgeartet (Satz 3.2.25). Also

$$B'(u, v) = u^\top A' v = u^\top A' A^{-1} A v = (\underbrace{(A^{-1})^\top (A')^\top}_{=:C} u)^\top A v = B(Cu, v). \qquad \square$$

8.1.3 Beschreibung von Bilinearformen durch quadratische Formen

Sei $B: V \times V \to \mathbb{K}$ eine Bilinearform auf V. Die zugehörige *quadratische Form*

$$q: V \to \mathbb{K}$$

ist definiert durch

$$q(v) := B(v, v).$$

Achtung: keine lineare Abbildung!
q hat die Eigenschaften:

1. $q(\lambda v) = \lambda^2 q(v)$
2. $q(u + v) = q(u) + B(u, v) + B(v, u) + q(v)$

Eine symmetrische Bilinearform ist sogar durch q eindeutig bestimmt (falls $\mathrm{char}(\mathbb{K}) \neq 2$), denn aus der zweiten Eigenschaft folgt ('Polarisierung'):

$$B(u, v) = 2^{-1}(q(u+v) - q(u) - q(v))$$

Bemerkung: 2^{-1} existiert nicht in $\mathbb{F}_2$ oder allgemeiner falls $\mathrm{char}(\mathbb{K}) = 2$.

Die *Kennlinie* K einer Bilinearform ist definiert durch

$$K := \{u \in V \mid q(u) = 1\}.$$

Beispiel 8.1.8. $V = \mathbb{R}^2$. Anderes Skalarprodukt:

$$\begin{aligned} x * y &:= 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 \\ &= (x_1\ x_2) \underbrace{\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}}_{\text{Gramsche Matrix}} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned}$$

Kriterium aus Proposition 8.1.5 zeigt, dass Skalarprodukt vorliegt:

$$\det(2) > 0, \quad \det \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} > 0$$

Die zugehörige quadratische Form:

$$q(x) = x * x = 2x_1^2 + 2x_1x_2 + x_2^2$$

Kennlinie:

$$K := \{x \in \mathbb{R}^2 \mid 2x_1^2 + 2x_1x_2 + x_2^2 = 1\}$$

eine Ellipse. Bild: eine Ellipse durch Punkte $(0,1), (0,-1), (1,-1), (-1,1)$.
Bestimmung der Achsen: Hauptachsentransformation. △

Bemerkung 8.1.9. Die Kennlinie einer Bilinearform ist stets ein Kegelschnitt (Ellipse für Skalarprodukte).

Klassifikation von Bilinearformen durch Kennlinie: Kapitel 8.3.5, verwendet ebenfalls Hauptachsentransformation.

8.1.4 Bilinearformen und Dualraum

Erinnerung: V^* steht für den dualen Raum aller Linearformen von V. Zusammenhang zwischen Bilinearformen $B\colon V \times V \to \mathbb{K}$ und linearen Abbildungen $L\colon V \to V^*$:

- Sei $\phi\colon V \to V^*$ lineare Abbildung. Dann ist durch

  $$B_\phi(u, v) := \underbrace{\phi(u)}_{\in V^*}(v)$$

  eine Bilinearform definiert (die genau dann nicht ausgeartet ist, wenn ϕ injektiv ist).

- Jede Bilinearform entsteht auf diese Weise. Für die Bilinearform $B\colon V \times V \to \mathbb{K}$ ist durch
$$\phi_B\colon V \to V^* : u \mapsto f_u$$
mit $f_u(v) := B(u, v)$ eine lineare Abbildung definiert (die genau dann injektiv ist, wenn B nicht ausgeartet ist).

Durch diesen Zusammenhang ist eine Bijektion gegeben, denn es gilt

$$B_{\phi_B} = B \text{ und } \phi_{B_\phi} = \phi.$$

8.2 Euklidische und Unitäre Vektorräume

Wiederholung (Definition 6.1.1): ein *euklidischer* bzw. *unitärer* Vektorraum ist ein $\mathbb{R}$ (bzw. $\mathbb{C}$)-Vektorraum V mit einem Skalarprodukt

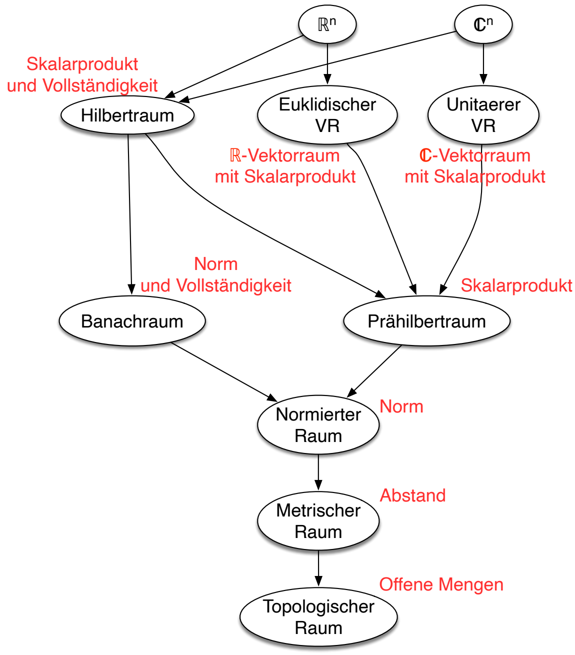
$$*\colon V \times V \to \mathbb{R} \qquad\qquad V \times V \to \mathbb{C}$$

Wiederholung:
$$\|u\| := \sqrt{u * u} = \sqrt{q(u)}$$
heißt *Norm* von $u \in V$.
u ist *normierter* Vektor falls $\|u\| = 1$.

Satz 8.2.1. *In einem euklidischen bzw. unitären Vektorraum gilt:*

- *Die Norm ist ein* vernünftiges Längenmaß:
  1. $\|x\| \geq 0$ *für alle* $x \in V$; $\|x\| = 0 \iff x = \mathbf{0}$.
  2. $\|\alpha x\| = |\alpha| \cdot \|x\|$ *für alle* $x \in V$ *und* $\alpha \in \mathbb{R}$ *bzw.* $\alpha \in \mathbb{C}$.
  3. *Die Dreiecksungleichung (siehe Abschnitt 6.1.5):*
  $$\|x + y\| \leq \|x\| + \|y\|$$

- *Die Cauchy-Schwarzsche Ungleichung:*
$$|x * y| \leq \|x\| \cdot \|y\|$$
äquivalent dazu:
$$(x * y)(x * y) \leq (x * x)(y * y)$$
bzw.
$$\begin{aligned} |x * y|^2 &= (x * y)\overline{(x * y)} \\ &= (x * y)(y * x) \leq (x * x)(y * y) \end{aligned}$$
Gleichheit gilt genau dann, wenn $x - \alpha y = \mathbf{0}$, *d.h., wenn* x, y *linear abhängig sind.*



Beweis. Wir konzentrieren uns auf den komplexen Fall der Cauchy-Schwarzen Ungleichung; den reellen Fall haben wir bereits in Abschnitt 6.1.4 betrachtet. Falls $y = \mathbf{0}$ ist die Aussage trivial. Sei nun $y \neq \mathbf{0}$. Setzen $\alpha := \frac{x*y}{||y||^2}$. Nun gilt

$$
\begin{aligned}
0 &\leq (x - \alpha y) * (x - \alpha y) && (* \text{ ist positiv definit})\\
&= x * (x - \alpha y) - \alpha y * (x - \alpha y) && (\text{Linearität in 1. Stelle})\\
&= (x * x) - \bar{\alpha}(x * y) - \alpha(y * x) + \alpha\bar{\alpha}(y * y) && (\text{Semilinearität in 2. Stelle})\\
&= ||x||^2 - \frac{\overline{(x*y)}(x*y)}{||y||^2}\\
&\quad - \frac{(x*y)(y*x)}{||y||^2} + \frac{(x*y)\overline{(x*y)}(y*y)}{||y||^2\,||y||^2} && (\text{Einsetzen von } \alpha)\\
&= ||x||^2 - \frac{(x*y)(y*x)}{||y||^2} && (\text{Vereinfachen})
\end{aligned}
$$

Damit ist $(x * y)^2 \leq ||x||^2\, ||y||^2$. □

Übung 55. Sei $A \in \mathbb{R}^{n \times n}$ so, dass $||Ax|| = 0$ für alle $x \in \mathbb{R}^n$. Zeigen Sie, dass $A = \mathbf{0}$.

Jede Norm auf einem Vektorraum induziert durch $d := ||x - y||$ eine *Metrik* (Abstandsbegriff ⇝ Analysis). Ist ein unitärer Raum *vollständig* bzgl. der Norm (i.e., jede Cauchyfolge konvergiert), so heißt er *Hilbert-Raum*. Ein Vektorraum, in dem ein Skalarprodukt definiert ist, heißt dagegen *Prähilbertraum*.

Beispiel 8.2.2. Betrachten die Menge der Folgen $(x_n)_{n\in\mathbb{N}}$ von komplexen Zahlen, so dass die Reihe $\sum_{n=1}^{\infty} |x_n|^2$ konvergiert. Hierauf definieren wir das Skalarprodukt

$$(x_n)_{n\in\mathbb{N}} * (y_n)_{n\in\mathbb{N}} := \sum_{n\in\mathbb{N}} x_n \bar{y}_n$$

Dieser Ausdruck konvergiert: zunächst ist

$$\sum_{i=1}^{n} |x_i \bar{y}_i| \le \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2}$$

wegen der Ungleichung von Cauchy-Schwarz für das Standardskalarprodukt im $\mathbb{C}^n$; die rechte Seite aber ist uniform beschränkt. Dieser Raum ist vollständig und damit ein Hilbertraum, und wird (wie die entsprechende Norm) mit ℓ_2 bezeichnet und auch "*der* Hilbertraum" genannt. Es ist bis auf Isometrie der einzige unendlich-dimensionale *separable* (d.h., mit abzählbarer dichter Teilmenge) Hilbertraum (Satz von Fischer-Riesz). △

8.2.1 Orthogonalität

Zwei Vektoren u, v eines euklidischen Vektorraums V heißen *orthogonal*, $u \perp v$, wenn $u * v = 0$. Für Teilmenge $U \subseteq V$ heißt

$$U^{\perp} := \{w \in V \mid \forall u \in U\colon w \perp u\}$$

das *orthogonale Komplement* von U.

Bemerkungen.

- $U^{\perp}$ ist stets Untervektorraum von V und es gilt

$$U^{\perp} = \langle U^{\perp}\rangle = \langle U\rangle^{\perp}$$

- Vorgriff: für $U \le V$ ist $U^{\perp}$ ein Komplement im Sinne von Definition 2.4.17, d.h., es gilt (für Begründung siehe Bemerkung 8.2.13)

$$V = U \oplus U^{\perp}$$

  (jedes $v \in V$ lässt sich *eindeutig* schreiben als $u + w$ für $u \in U$ und $w \in U^{\perp}$) Insbesondere:

$$\dim U + \dim U^{\perp} = \dim V$$

8.2.2 Orthogonalsysteme

Sei V euklidischer oder unitärer Vektorraum.

Definition 8.2.3. $(v_1, \ldots, v_r) \in V^r$ heißt

- *Orthogonalsystem* falls $v_1 \neq \mathbf{0}, \dots, v_r \neq \mathbf{0}$ und $v_i * v_j = 0$ für verschiedene $i, j \in \{1, \dots, r\}$;
- *Orthonormalsystem* falls $(v_1, \dots, v_n)$ ein *normiertes* Orthogonalsystem, d.h., falls zusätzlich $\|v_i\| = 1$ für alle $i \in \{1, \dots, r\}$. Anders geschrieben: $v_i * v_j = \delta_{ij}$ (Kroneckersymbol);
- *Orthogonalbasis* falls Basis und Orthogonalsystem;
- *Orthonormalbasis* (oder kurz *ON-Basis*) falls Basis und Orthonormalsystem.

Proposition 8.2.4. *Jedes Orthogonalsystem ist linear unabhängig.*

Folgerung: ein Orthogonalsystem $(v_1, \dots, v_r)$ ist genau dann eine ON-Basis, wenn $r = \dim V$.

Beweis. Sei $(v_1, \dots, v_r)$ Orthogonalsystem, und

$$\alpha_1 v_1 + \dots + \alpha_n v_n = \mathbf{0}$$

Zu zeigen: $\alpha_i = 0$. Skalarprodukt mit v_i auf beiden Seiten ergibt

$$\alpha_1(v_1 * v_i) + \dots + \alpha_r(v_r * v_i) = \mathbf{0}$$

also $\alpha_i(v_i * v_i) = \mathbf{0}$. Da $v_i \neq \mathbf{0}$ ist $v_i * v_i \neq 0$, also $\alpha_i = 0$. □

Bemerkung 8.2.5. Die Gramsche Matrix eines Skalarprodukts ist bezüglich einer ON-Basis stets die Einheitsmatrix.

Beispiel 8.2.6. Sei $V = \mathbb{R}^n$ und

$$x * y := x_1 y_1 + \dots + x_n y_n$$

das Standardskalarprodukt. Wegen $e_i * e_j = \delta_{ij}$ ist $(e_1, \dots, e_n)$ eine ON-Basis. △

8.2.3 Das Gram-Schmidtsche Orthonormalisierungsverfahren

Sei V ein euklidischer bzw. unitärer VR und seien $(v_1, \dots, v_r)$ linear unabhängig.

Satz 8.2.7 (Gram-Schmidt)**.** *Durch folgende rekursive Definitionen erhält man für $k \in \{1, \dots, r\}$ ein Orthonormalsystem $(\tilde{v}_1, \dots, \tilde{v}_r)$ mit $\langle v_1, \dots, v_k \rangle = \langle \tilde{v}_1, \dots, \tilde{v}_k \rangle$:*

$$\begin{aligned}
\tilde{v}_1 &:= \frac{1}{\|v_1\|} v_1 \\
v_k' &:= v_k - \sum_{i=1}^{k-1} (v_k * \tilde{v}_i)\tilde{v}_i \\
\tilde{v}_k &:= \frac{1}{\|v_k'\|} v_k' && \textit{(Normierung)}.
\end{aligned} \tag{8.5}$$

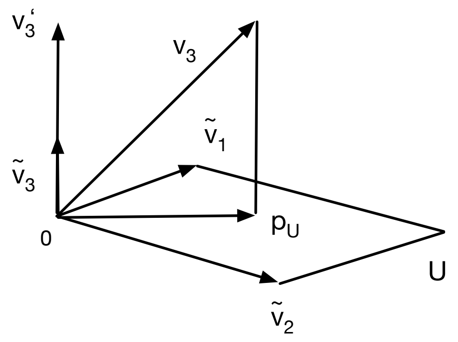


Abbildung 8.1: Illustration zum Verfahren von Gram-Schmidt für $k = 3$.

Idee für (8.5) für $k = 3$: sei $U = \langle \tilde{v}_1, \tilde{v}_2 \rangle$ der von $\tilde{v}_1, \tilde{v}_2$ aufgespannte Untervektorraum und p_U die Projektion von v_3 auf U. Siehe Abbildung 8.1 (und Abschnitt 6.2.4).

$$v_3' = v_3 - p_u = v_3 - \sum_{i=1}^{2} (v_3 * \tilde{v}_i)\tilde{v}_i$$

Folgerung 1. Jeder n-dimensionale euklidische (unitäre) VR hat eine ON-Basis (man starte Verfahren mit Basis $v_1, \dots, v_n$).

Folgerung 2. Jede Orthonormalbasis eines Untervektorraums $U \leq V$ läßt sich zu einer ON-Basis von V ergänzen.

Beweis von Satz 8.2.7. Per Induktion über k. Fall $k = 1$: klar, $\tilde{v}_1 * \tilde{v}_1 = 1$. Induktionsschritt: Behauptung sei für alle $j < k$ bereits bewiesen.

$$\begin{aligned} v_k' * v_j &= \left(v_k - \sum_{i=1}^{k-1} (v_k * \tilde{v}_i)\tilde{v}_i \right) * v_j \\ &= v_k * \tilde{v}_j - \sum_{i=1}^{k-1} (v_k * \tilde{v}_i) \underbrace{(\tilde{v}_i * \tilde{v}_j)}_{=\delta_{ij}} \\ &= v_k * \tilde{v}_j - v_k * \tilde{v}_j = 0 \end{aligned}$$

also $\tilde{v}_k * \tilde{v}_j = \frac{1}{\|v_k'\|}(v_k', \tilde{v}_j) = 0$. Weiterhin gilt $\tilde{v}_k * \tilde{v}_k = 1$ und

$$\langle \tilde{v}_1, \dots, \tilde{v}_k \rangle = \langle v_1, \dots, v_k \rangle$$

Dies folgt aus

$$\begin{aligned} \langle \tilde{v}_1, \dots, \tilde{v}_k \rangle &\subseteq \langle \tilde{v}_1, \dots, \tilde{v}_{k-1}, v_k \rangle && \text{(wegen (8.5))} \\ &= \langle v_1, \dots, v_{k-1}, v_k \rangle && \text{(Induktionsvoraussetzung))} \end{aligned}$$

Es gilt sogar Gleichheit, da $\tilde{v}_1, \dots, \tilde{v}_k$ nach Proposition 8.2.4 linear unabhängig. □

Beispiel 8.2.8. Sei $V = \mathbb{R}^4$ mit Standardskalarprodukt.

$$v_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \quad v_2 := \begin{pmatrix}-1\\0\\1\\2\end{pmatrix} \quad v_3 := \begin{pmatrix}0\\0\\0\\-1\end{pmatrix}$$

sind linear unabhängig. Gram-Schmidtsches ON-Verfahren liefert:

$$\tilde{v}_1 := \frac{v_1}{\|v_1\|} = \frac{v_1}{\sqrt{4}} = \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

und

$$v_2' := v_2 - \underbrace{(v_2 * \tilde{v}_1)}_{-\frac{1}{2}+0+\frac{1}{2}+\frac{1}{2}2=1} \tilde{v}_1 = \begin{pmatrix}-1\\0\\1\\2\end{pmatrix} - 1 \cdot \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}-3\\-1\\1\\3\end{pmatrix}.$$

Weiterhin $\tilde{v}_2 = \frac{v_2'}{\|v_2\|} = \frac{1}{2\sqrt{5}}\begin{pmatrix}-3\\-1\\1\\3\end{pmatrix}$. (Probe: $v_1 * v_2 = 0$.)

$$v_3' := v_3 - \underbrace{(v_3 * \tilde{v}_1)}_{=-\frac{1}{2}} \tilde{v}_1 - \underbrace{(v_3 * \tilde{v}_2)}_{=-\frac{3}{2\sqrt{5}}} \tilde{v}_2$$

$$= \begin{pmatrix}0\\0\\0\\-1\end{pmatrix} \underbrace{-(-\frac{1}{2})\frac{1}{2}}_{=\frac{5}{20}} \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \underbrace{-(-\frac{3}{2\sqrt{5}})\frac{1}{2\sqrt{5}}}_{\frac{3}{20}} \begin{pmatrix}-3\\-1\\1\\3\end{pmatrix} = \frac{1}{20}\begin{pmatrix}-4\\2\\8\\-6\end{pmatrix} = \frac{1}{10}\begin{pmatrix}-2\\1\\4\\-3\end{pmatrix}$$

$\tilde{v}_3 := \frac{v_3'}{\|v_3'\|} = \frac{1}{\sqrt{30}}\begin{pmatrix}-2\\1\\4\\-3\end{pmatrix}$. (Probe: $v_i * v_j = \delta_{ij}$.) $\triangle$

Bemerkung 8.2.9. Manchmal kann es von Vorteil sein, zunächst nur ein Orthogonalsystem zu berechnen, und die Normierung der Vektoren erst am Ende des Verfahrens durchzuführen, da dann im Anfangsteil des Algorithmus keine Wurzeln gezogen werden müssen.

Motivation: Mit ON-Basis wird das Rechnen mit Koordintenvektoren, Skalarprodukt, und orthogonalem Komplement besonders einfach.

Bemerkung 8.2.10. Sei $(v_1, \dots, v_n)$ ON-Basis von euklidischen/unitären VR V. Dann gilt folgender Entwicklungssatz für $v \in V$ (gilt auch für unendliche ON-Basen):

$$v = \sum_{i=1}^{n} \underbrace{(v * v_i)}_{\text{"Fourierkoeffizienten"}} v_i$$

Denn: falls $v = \sum_{j=1}^{n} \alpha_j v_j$, dann gilt

$$v * v_i = \sum_{j=1}^{n} \alpha_j \underbrace{(v_j, v_i)}_{=\delta_{ij}} = \alpha_i.$$

Bemerkung 8.2.11. Bezüglich einer ON-Basis $(v_1, \dots, v_n)$ ist jedes Skalarprodukt das Standardskalarprodukt der Koordinatenvektoren: für $u = \sum_{i=1}^{n} \alpha_i v_i$ und $w = \sum_{i=1}^{n} \beta_i v_i$ gilt

$$u * w = \alpha_1 \beta_1 + \dots + \alpha_n \beta_n$$

bzw., in unitären VR,

$$u * w = \alpha_1 \bar{\beta}_1 + \dots + \alpha_n \bar{\beta}_n$$

Denn:

$$\sum \alpha_i v_i * \sum \beta_j v_j = \sum_{i,j} \alpha_i \bar{\beta}_j \underbrace{(v_i * v_j)}_{\delta_{ij}}$$

Bemerkung 8.2.12. Es gilt die *Parsevalsche Gleichung*:

$$v * v = \|v\|^2 = \sum_{i=1}^{n} |v * v_i|^2$$

(rechte Seite: Summe der Quadrate der Fourrierkoeffizienten)

Denn: für $v = \sum_{i=1}^{n} \alpha_i v_i$ ist

$$v * v \overset{(2.)}{=} \alpha_1 \bar{\alpha}_1 + \dots + \alpha_n \bar{\alpha}_n = |\alpha_1|^2 + \dots + |\alpha_n|^2 \overset{(1.)}{=} \sum_{i=1}^{n} |v * v_i|^2$$

In Hilberträumen gilt für orthonormales System $(e_1, e_2, \dots)$:

$$v * v = \|v\|^2 \geq \sum_{i \in \mathbb{N}} |v * e_i|^2 \qquad \textit{(Besselsche Ungleichung)}$$

Bemerkung 8.2.13. Sei $U \leq V$ und $(u_1, \dots, u_m)$ ON-Basis von U.
Sei $(u_1, \dots, u_m, u_{m+1}, \dots, u_n)$ Ergänzung zu ON-Basis von V (Verfahren von Gram-Schmidt, Folgerung 2). Dann ist

$$U^{\perp} = \langle u_{m+1}, \dots, u_n \rangle$$

und $(u_{m+1}, \dots, u_n)$ ist ON-Basis von $U^\perp$.
Also: $V = U \oplus U^\perp$ (siehe Abschnitt 2.4.17).

Denn: $v_i * v_j = 0$ fur $i \le m$ und $m+1 \le j$, also $v_{m+1}, \dots, v_n \in U^\perp$ also

$$\langle u_{m+1}, \dots, u_n \rangle \subseteq U^\perp$$

Umgekehrt sei $v \in \sum \alpha_i v_i \in U^\perp$. Dann ist $\alpha_i \overset{(1)}{=} v * v_i = 0$ für $i \le m$ (da $v_i \in U$). Also $v = \sum_{i=m+1}^n \alpha_i v_i \in \langle v_{m+1}, \dots, v_n \rangle$.

8.2.4 Orthogonalprojektion

Sei V euklidischer (oder unitärer) VR, und $U \le V$ Untervektorraum.

$$\begin{aligned} & V = U \oplus U^\perp \\ \exists u, w : \quad & v = u + w \qquad \text{(eindeutig!)} \end{aligned}$$

Bezeichnung: $p_U(v) := u$.

Definition 8.2.14. Die *Orthogonalprojektion* eines Vektors $v \in V$ auf einen Unterraum U ist der (eindeutig bestimmte) Vektor $p_U(v) \in U$, so dass eine Zerlegung $v = p_U(v) + w$ mit $w \in U^\perp$ existiert (insbesondere $p_U(v) \perp w$).

Einfache Berechnung von $p_U(v)$ mit ON-Basis $(u_1, \dots, u_m)$ von U:

$$p_U(v) = \sum_{i=1}^m (v * u_i) u_i$$

(erster Teil der Fourierentwicklung, siehe Abschnitt 8.2.3)

Satz 8.2.15. *Sei V euklidischer VR und $U \le V$.*

1. *Für alle $u \in U$ und $v \in V$ gilt*

$$\|v - p_U(v)\|^2 \le \|v - u\|^2$$

2. *Das Gleichheitszeichen gilt* nur *für $u = p_U(v)$.*

Beweis. Zu (1). Sei $v \in V$. Dann gibt es $w \in U^\perp$ mit $v = p_U(v) + w$ (Abschnitt 8.2.1). Sei $u \in U$. Dann gilt

$$\begin{aligned} \|v-u\|^2 = \|w + p_U(v) - u\|^2 &= \|w\|^2 + \|p_U(v) - u\|^2 \qquad \text{(da } w \perp (p_U(v) - u)) \\ &\ge \|w\|^2 = \|v - p_U(v)\|^2. \end{aligned}$$

Zu (2). Nach dem Entwicklungssatz aus Abschnitt 8.2.3:

$$\begin{aligned} v &= \underbrace{\sum_{i=1}^m (v * v_i) v_i}_{\in U} + \underbrace{\sum_{i=m+1}^n (v * v_i) v_i}_{\in U^\perp} \\ &= p_U(v) \qquad \text{(nach Definition 8.2.14).} \end{aligned}$$

□

8.2.5 Anwendung: Methode der kleinsten Fehlerquadrate

Entwickelt von Gauß zur Berechnung von Planetenbahnen.

Annahme: theoretisch bekannter (oder vermuteter) Zusammenhang f zwischen zwei (Mess-) Größen

$$y = f(x) = a_0 + a_1 g_1(x) + \cdots + a_m g_m(x)$$

wobei $g_1, \ldots, g_m$ bekannte Funktionen, z.B. $g_i(x) := x^i$, also

$$y = f(x) = a_0 + a_1 x + \cdots + a_m x^m$$

Gegeben: Messreihe $(x_1, y_1), \ldots, (x_n, y_n)$ (i.A. fehlerbehaftet)

Gesucht: Möglichst genaue Approximation für f, d.h., für $a_0, a_1, \ldots, a_m$.
Was ist gute Approximation?

Ansatz:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \underset{\text{exakt}}{\overset{\text{falls Messwerte}}{=}} \begin{pmatrix} 1 & g_1(x_1) & \cdots & g_m(x_1) \\ 1 & g_1(x_2) & \cdots & g_m(x_2) \\ \vdots & \vdots & & \vdots \\ 1 & g_1(x_n) & \cdots & g_m(x_n) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Maß für Abweichung der Kurve f von den Messpunkten ist

$$\underset{\text{Norm in } \mathbb{R}^{m+1}}{\|y - Aa\|} \quad \underset{\text{z.B. für Standardskalarprodukt.}}{= \sqrt{(y_1 - c_1)^2 + \cdots + (y_n - c_n)^2}}$$

→ *Methode der kleinsten Fehlerquadrate.*
Große Abweichungen der Modellfunktion von den Daten werden stärker gewichtet.

Gegeben: A, y.
Gesucht: $a \in \mathbb{R}^{m+1}$, so dass $\|y - Aa\| \leq \|y - Ab\|$ für alle $b \in \mathbb{R}^{m+1}$.

Sei

$$U := \{Ab \mid b \in \mathbb{R}^{m+1}\} = \text{Bild } A \leq \mathbb{R}^n$$

Lösung $Aa = p_U(y)$ liefert beste Approximation gemäß Definition 8.2.14,
da $\|y - p_U(y)\|$ minimal.

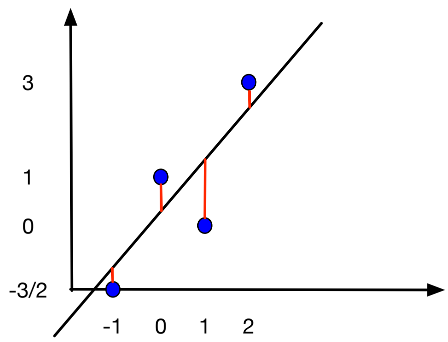
Lösungsmethode:

1. Bestimmung einer Basis $(v_1, \ldots, v_r)$ von $U :=$ Bild A.
2. Bestimmung einer ON-Basis $(\tilde{v}_1, \ldots, \tilde{v}_r)$ von U (siehe Abschnitt 8.2.3).
3. Berechne $p_U(y) = \sum_{i=1}^{r} (y * \tilde{v}_i)\tilde{v}_i$ (siehe Satz 8.2.15).

4. Berechnung der Lösung $a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}$ des Gleichungssystems
$$Aa = p_U(y)$$
liefert beste Approximation
$$f(x) = a_0 + a_1 g_1(x) + \cdots + a_m g_m(x).$$
Lösbarkeit garantiert, da $p_U(y) \in U = \text{Bild } A$, Abschnitt 3.3.2.
System eindeutig lösbar falls $\text{rg}(A) = r = n$, Korollar 3.3.7.

Beispiel. Messreihe:

Versuch Nr.	x_i	y_i
1	-1	-3/2
2	0	1
3	1	0
4	2	3



Theoretisch gegebener Zusammenhang sei lineare Funktion (Gerade)
$$y = f(x) = a_0 + \underbrace{a_1 x}_{g_1(x)}$$
"Ausgleichsrechnung": Fehler $\sqrt{\sum_{i \le 4} (f(x_i) - y_i)^2}$ minimieren.
$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, y = \begin{pmatrix} -3/2 \\ 1 \\ 0 \\ 3 \end{pmatrix}$$

Gesucht: Lösung $a = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$ für $Aa = p_U(y)$, $U := \text{Bild } A$.

1. Basis von U: haben $\text{rg}(A) = 2$ und wählen
$$v_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 := \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

2. ON-Basis von U:

$$\tilde{v}_1 = \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

$$v_2' := v_2 - (v_2 * \tilde{v}_1)\tilde{v}_1 = \begin{pmatrix}-1\\0\\1\\2\end{pmatrix} - \frac{1}{4}\left(\begin{pmatrix}-1\\0\\1\\2\end{pmatrix} * \begin{pmatrix}1\\1\\1\\1\end{pmatrix}\right)\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}-3\\-1\\1\\3\end{pmatrix}$$

$$\|v_2'\| = \sqrt{(9+1+1+9)/4} = \sqrt{5}$$

$$\tilde{v}_2 := \frac{1}{\|v_2'\|}v_2' = \frac{1}{2\sqrt{5}}\begin{pmatrix}-3\\-1\\1\\3\end{pmatrix}$$

3. Berechnung von $p_U(y)$:

$$p_U(y) = (y * \tilde{v}_1)\tilde{v}_1 + (y * \tilde{v}_2)\tilde{v}_2 \overset{\text{Ü}}{=} \frac{5}{4}\tilde{v}_1 + \frac{5\sqrt{5}}{4}\tilde{v}_2 \overset{\text{Ü}}{=} \frac{5}{4}\begin{pmatrix}-1\\0\\1\\2\end{pmatrix}$$

4. Lösung des Gleichungssystems $A\begin{pmatrix}a_0\\a_1\end{pmatrix} = p_U(y)$ nach a_0 und a_1:

$$\begin{pmatrix}1 & -1\\1 & 0\\1 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}a_0\\a_1\end{pmatrix} = \begin{pmatrix}-5/4\\0\\5/4\\5/2\end{pmatrix}$$

   Hat Lösung $a_0 = 0, a_1 = 5/4$.

Ergebnis: die beste Approximation für die Messreihe ist die Gerade

$$y = f(x) = \frac{5}{4}x$$

Jede andere Gerade liefert größeren Fehler!

8.3 Klassifikation bis auf orthogonale und unitäre Ähnlichkeit

Was sind strukturverträgliche Abbildungen für euklidische Vektorräume? Mit Struktur ist gemeint: Skalarprodukt, Längen, Orthogonalität, Winkel, ... Die Antwort lautet: orthogonale Abbildungen.

8.3.1 Orthogonale und unitäre Abbildungen

In diesem Abschnitt behandeln wir orthogonale (und unitäre) Abbildungen in etwas allgemeinerer Form als in Abschnitt 6.5).

Es seien V und W euklidische (unitäre) Vektorräume, $*_V$ und $*_W$ die dazugehörigen Skalarprodukte, und $\|.\|_V$, $\|.\|_W$ die zugehörigen Normen.

Definition 8.3.1. Eine lineare Abbildung $f\colon V \to W$ heißt *orthogonal* (bzw. unitär) falls für alle $u, v \in V$:

$$u *_V v = f(u) *_W f(v)$$

Satz 8.3.2 (Charakterisierung Orthogonalität)**.** *Es sei $f\colon V \to W$ lineare Abbildung. Dann sind äquivalent:*

1. *f ist orthogonal;*
2. $\forall x \in V : \|x\|_V = 1 \Rightarrow \|f(x)\|_W = 1$
3. $\forall x \in V : \|x\|_V = \|f(x)\|_W$ *(f ist* längentreu*)*
4. *falls $(u_1, \ldots, u_r)$ ON-System in V, so ist $(f(u_1), \ldots, f(u_r))$ ein ON-System in W.*

Beweis. 1. $\Rightarrow$ 2.: Aus $\|x\| = 1$ folgt

$$\begin{aligned}\|f(x)\|_W^2 &= f(x) * f(x)\\ &= x * x && \text{(wegen (1))}\\ &= \|x\|_V^2 = 1\end{aligned}$$

also auch $\|f(x)\|_W = 1$.

2. $\Rightarrow$ 3.: Ist $x = \mathbf{0}$, so gilt $f(x) = \mathbf{0}$, also $\|f(x)\| = 0 = \|x\|$.
Sei nun $x \neq \mathbf{0}$. Für $\tilde{x} = \frac{x}{\|x\|}$ gilt $\|\tilde{x}\| = 1$. Also $\|f(\tilde{x})\| = 1$ wegen (2) und es folgt

$$\begin{aligned}\|f(x)\|_W &= \|f(\|x\|_V \tilde{x})\|\\ &= \|x\|_V \cdot \|f(\tilde{x})\|_W && \text{(Linearität von } f \text{ und Eigenschaft von Normen)}\\ &= \|x\|_V.\end{aligned}$$

3. $\Rightarrow$ 4.: Sei $(u_1, \ldots, u_r)$ ein ON-System. Dann

$$\|u_j\| = 1 \overset{(3)}{\Rightarrow} \|f(u_j)\| = 1$$

Sei $j \neq k$, zu zeigen bleibt: $f(u_j) * f(u_k) = 0$.

$$\|u_j + u_k\|^2 = \|u_j\|^2 + \|u_k\|^2 + \overbrace{2(u_j * u_k)}^{=0}$$
$$\|f(u_j + u_k)\|^2 = \|f(u_j) + f(u_k)\|^2 = \|f(u_j)\|^2 + \|f(u_k)\|^2 + 2(f(u_j) * f(u_k))$$

Also $f(u_j) * f(u_k) = 0$.

4. $\Rightarrow$ 1.: Es gelte (4), z.z. ist $u * v = f(u) * f(v)$.

1. Fall: u, v sind linear abhängig, o.B.d.A: $v = \alpha u$ für $\alpha \in \mathbb{K}$. Dann ist $\tilde{u} := \frac{u}{\|u\|}$ ein ON-System (bestehend aus nur einem Vektor) also auch $f(\tilde{u})$ nach (4). Es folgt

$$u * v = \|u\|\tilde{u} * \alpha\|u\|\tilde{u} = \underbrace{\tilde{u} * \tilde{u}}_{=1}\|u\|^2\alpha$$

$$\text{und} \quad f(u) * f(v) = f(\|u\|\tilde{u}) * f(\alpha\|u\|\tilde{u}) = \underbrace{(f(\tilde{u}) * f(\tilde{u}))}_{=1}\|u\|^2\alpha.$$

2. Fall: u, v sind linear unabhängig. Verfahren aus Abschnitt 8.2.3 liefert ON-System $(\tilde{u}, \tilde{v})$ mit $\langle u, v\rangle = \langle \tilde{u}, \tilde{v}\rangle$, d.h. es gibt $\alpha_1, \alpha_2, \beta_1, \beta_2$ so dass $u = \alpha_1\tilde{u} + \alpha_2\tilde{v}$ und $v = \beta_1\tilde{u} + \beta_2\tilde{v}$. Nach (4) ist $(f(\tilde{u}), f(\tilde{v}))$ ein ON-System, also

$$\begin{aligned} f(u) * f(v) &= (\alpha_1 f(\tilde{u}) + \alpha_2 f(\tilde{v})) * (\beta_1 f(\tilde{u}) + \beta_2 f(\tilde{v})) \\ &= \alpha_1\beta_1 + \alpha_2\beta_2 \\ &= (\alpha_1\tilde{u} + \alpha_2\tilde{v}) * (\beta_1\tilde{u} + \beta_2\tilde{v}) = u * v. \end{aligned}$$ □

Wegen Satz 8.3.2 werden unitäre und orthogonale Abbildungen auch häufig *Isometrien* genant. Wir sehen also: längentreu impliziert winkeltreu. Die Umkehrung gilt aber nicht ($x \mapsto 2x$ ist winkeltreue, aber nicht längentreue linear Abbildung).

Folgerungen:

1. Sind $B = (v_1, \ldots, v_n)$ und $B' = (w_1, \ldots, w_n)$ ON-Basen von V bzw. von W, so ist die durch $f\colon V \to W : v_i \mapsto w_i$ definierte lineare Abbildung orthogonal.

2. Wenn $f\colon V \to W$ orthogonal, dann ist f injektiv, denn

$$\begin{aligned} f(x) = \mathbf{0} &\Rightarrow \|f(x)\| = 0 \\ &\Rightarrow \|x\| = 0 \\ &\Rightarrow x = \mathbf{0}. \end{aligned}$$

8.3.2 Darstellungsmatrizen orthogonaler Abbildungen

Häufig wird Orthogonalität nur für quadratische Matrizen definiert; wir machen das gleich etwas allgemeiner, damit auch der Zusammenhang zu orthogonalen Abbildungen in voller Allgemeinheit formuliert werden kann.

Definition 8.3.3. Eine Matrix $A \in \mathbb{R}^{m\times n}$ (bzw. $\mathbb{C}^{m\times n}$) heiße *orthogonal* (bzw. unitär) wenn

$$A^\top A = E_n \quad \text{bzw. } \bar{A}^\top A = E_n.$$

Bemerkung 8.3.4. Falls $A \in \mathbb{R}^{n\times n}$ orthogonal (unitär), so ist A invertierbar und es gilt $A^{-1} = A^\top$ ($A^{-1} = \bar{A}^\top$), also insbesondere auch $AA^\top = AA^{-1} = E_n$ ($A\bar{A}^\top = AA^{-1} = E_n$).

Rechtfertigung für diese Definition: Satz 8.3.5.

Satz 8.3.5. *Seien V und W euklidische (unitäre) Vektorräume mit ON-Basen $B = (v_1, \dots, v_n)$ beziehungsweise $C = (w_1, \dots, w_m)$. Sei $f\colon V \to W$ eine lineare Abbildung und sei $A := M_C^B(f) \in \mathbb{R}^{m\times n}$ ($\mathbb{C}^{m\times n}$) die Darstellungsmatrix von f. Dann sind äquivalent:*

1. *A ist orthogonal (unitär);*
2. *f ist orthogonal (unitär);*
3. *die Spalten von A bilden eine ON-Basis von $U \le \mathbb{R}^m$ ($\mathbb{C}^m$) mit $\dim(U) = n$ bezüglich des Standardskalarproduktes.*

Beweis. (1) $\Rightarrow$ (2): Für das Standardskalarprodukt haben wir diese Implikation in Proposition 6.5.1 gezeigt. Der allgemeine Fall geht im Prinzip genauso; wir zeigen ihn gleich für den unitären Fall. Seien $u = x_1u_1 + \dots + x_nu_n \in V$ und $v = y_1v_1 + \dots + y_nv_n \in V$. Da $*$ hermitesch und B orthonormal gilt

$$\begin{aligned} u * v &= x_1\overline{y_1} + \dots + x_n\overline{y_n} \\ &= x^\top \overline{E_n}\bar{y} = x^\top \overline{\bar{A}^\top A}\bar{y} && \text{(da } A \text{ unitär)} \\ &= x^\top (A^\top \bar{A})\bar{y} = (Ax)^\top \overline{(Ay)} = f(u) * f(v). \end{aligned}$$

(2) $\Rightarrow$ (3): nach Satz 8.3.2 ist $f(v_1), \dots, f(v_n)$ ein ON-System. Die i-te Spalte von A ist der Koordinatenvektor von $f(v_i)$ bezüglich der ON-Basis C; also sind die Spalten von A eine ON-Basis von $U \le \mathbb{R}^m$ ($\mathbb{C}^m$) mit $\dim(U) = n$.
(3) $\Rightarrow$ (1): Seien $s_1, \dots, s_n$ die Spalten von A. Dann gilt $s_i * s_j = s_i^\top s_j = \delta_{ij}$ für alle $i, j \in \{1, \dots, n\}$. Also $A^\top A = E_n$. □

$O(n) \subset \mathbb{R}^{n\times n}$: Menge aller quadratischen orthogonalen Matrizen.
$U(n) \subset \mathbb{C}^{n\times n}$: Menge aller quadratischen unitären Matrizen.

Bemerkung 8.3.6. Für orthogonale (bzw. unitäre) Matrix $A \in \mathbb{R}^{n\times n}$ (bzw. $\mathbb{C}^{n\times n}$) gilt

- $|\det A| = 1$ (denn $\frac{1}{\det(A)} = \det(A^{-1}) = \det(A^\top) = \det(A)$).
- $|\lambda| = 1$ für jeden Eigenwert λ von A (denn f_A ist längentreu!).

Beispiel 8.3.7. *Permutationsmatrizen* (Beispiel 7.3.2) sind orthogonal: Bezeichnet P_π die zu einer Permutation π zugehörige Permutationsmatrix, dann gilt

$$P_\pi^\top P_\pi = P_{\pi^{-1}} P_\pi = P_{\pi^{-1}\circ\pi} = P_{\mathrm{id}} = E$$

denn

- die transponierte Permutationsmatrix ist gleich der Permutationsmatrix der inversen Permutation, und

- das Produkt von Permutationsmatrizen entspricht der Hintereinanderausführung der Permutationen. △

Übung 56. Wie würden Sie die Lösung des linearen Gleichungssystems $Ax = b$ berechnen, wenn $A \in \mathbb{Q}^{n\times n}$ orthogonal ist?

Übung 57. Zeigen Sie: die vorzeichenbehafteten Permutationsmatrizen, bei denen in jeder Zeile und Spalte genau ein Eintrag plus oder minus eins ist und alle übrigen Einträge null sind, sind genau die ganzzahligen orthogonalen Matrizen.

8.3.3 Orthogonale und unitäre Ähnlichkeit

Zwei Matrizen $A, A' \in \mathbb{R}^{n\times n}$ $(\mathbb{C}^{n\times n})$ heißen *orthogonal ähnlich* (bzw. unitär ähnlich; in der Literatur bisweilen auch: *unitär äquivalent*, das ist aber im Hinblick auf die Definition von (gewöhnlicher) Ähnlichkeit und Äquivalenz irreführend) falls es eine orthogonale (unitäre) Matrix S gibt so dass $A = S^{-1}A'S$. Eine Äquivalenzrelation.

Eine Klassifikation aller Matrizen bis auf orthogonale (unitäre) Ähnlichkeit ist für diese Vorlesung zu ehrgeizig. Dazu ein Beispiel. Sei $n > 2$, und betrachten

$$A = \begin{pmatrix} 1 & 1 & * & * & * & \cdots & * \\ 0 & 2 & 1 & * & * & \cdots & * \\ 0 & 0 & 3 & 1 & * & \cdots & * \\ 0 & 0 & 0 & 4 & 1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & * \\ \vdots & & & & \ddots & \ddots & 1 \\ 0 & \cdots & & & \cdots & 0 & n \end{pmatrix} \tag{8.6}$$

Alle diese Matrizen sind ähnlich, denn die jordansche Normalform ist immer die gleiche, da allesamt diagonalisierbar wegen $\chi_A = (X-1)(X-2)\cdots(X-n)$.

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & & \\ \vdots & & \ddots & \\ 0 & & & n \end{pmatrix}$$

Auf der anderen Seite sind zwei Matrizen von der Form wie in (8.6) nur dann unitär ähnlich, wenn sie die gleichen Einträge haben (ohne Beweis; hier wird verwendet, dass auf der oberen Nebendiagonale von A alle Einträge 1 sind. Originalliteratur dazu: Heydar Radjavi, *On Unitary Equivalence of Arbitrary Matrices*, Transactions of the AMS, 1962).

Wir erwähnen ohne Beweis:

Satz 8.3.8 (Normalform orthogonaler Matrizen)**.** *Eine reelle Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann orthogonal $(A^{-1} = A^\top)$, wenn sie zu einer Matrix der folgenden Gestalt orthogonal* *ähnlich ist*

$$\begin{pmatrix} +1 &&&&&&&& \\ & \ddots &&&&&&& \\ && +1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& K_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& K_m \end{pmatrix}$$

wobei $K_1, \dots, K_m$ Drehmatrizen (Abschnitt 6.5.2), d.h., jeweils von der Gestalt

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

für ein $\alpha \in \mathbb{R}$.

8.3.4 Selbstadjungierte Abbildungen

Für die wichtige Klasse der symmetrischen (hermiteschen) Matrizen wird uns eine Klassifikation bis auf orthogonale Ähnlichkeit gelingen (in Abschnitt 8.3.5). Im folgenden sei V ein euklidischer (bzw. unitärer) Vektorraum.

Definition 8.3.9. Ein Endomorphismus $\varphi\colon V \to V$ heißt *selbstadjungiert* wenn für alle $u, v \in V$

$$f(u) * v = u * f(v).$$

Satz 8.3.10. *Sei $f \in \mathrm{End}(V)$ und $A := M_B^B(f)$ Darstellungsmatrix von f bezüglich einer ON-Basis B. Dann ist f genau dann selbstadjungiert wenn A symmetrisch (bzw. hermitesch, $A = \bar{A}^\top$) ist.*

Zum Namen: $\bar{A}^\top =: A^*$ heißt *Adjungierte* zu A.

Beweis. Den Koordinatenvektor von $v \in V$ bezüglich B bezeichnen wir mit v_B. Dann ist $(f(v))_B = Av_B$ und $u * v = u_B^\top v_B$ (die Gramsche Matrix ist E_n weil B eine ON-Basis, siehe Abschnitt 8.2.3). (Bzw.: $u * v = u_B^\top \bar{v}_B$) Also

$$\begin{aligned} & f(u) * v = u * f(v) && \forall u, v \in V \\ \iff\ & (Au_B)^\top v_B = u_B^\top A v_B && \forall u, v \in V \\ \iff\ & u_B^\top A^\top v_B = u_B^\top A v_B && \forall u, v \in V \\ \iff\ & A^\top = A \end{aligned}$$

(Komplexe Variante: Striche über v_B und die A's auf der rechten Seite.) □

Bemerkung 8.3.11. Adjazenzmatrizen von (ungerichteten) Graphen (siehe Abschnitt 3.2) sind symmetrisch.

Bemerkung 8.3.12. Symmetrische Matrizen treten auch bei der Beschreibung quadratischer Formen auf (siehe Abschnitt 8.3.7).

Satz 8.3.13. *Sei $f \in \mathrm{End}(V)$ selbstadjungiert. Dann gilt:*

1. *f hat nur reelle Eigenwerte, die Nullstellen von χ_f (interessant wenn V unitärer Vektorraum);*

2. *Das charakteristische Polynom χ_f zerfällt in Linearfaktoren (interessant wenn V euklidischer Vektorraum);*

3. *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis. Zu 1. Sei v ein Eigenvektor zum EW λ:

$$f(v) = \lambda v, \quad v \neq \mathbf{0}$$

Dann gilt

$$\lambda(v * v) = \lambda v * v = f(v) * v = v * f(v) = v * \lambda v = \bar{\lambda}(v * v)$$

also $\lambda = \bar{\lambda}$, und daher $\lambda \in \mathbb{R}$.

Zu 2. Falls V unitär: Fundamentalsatz der Algebra. Falls V euklidisch: Sei $A := M_B^B(f)$ (symmetrische!) Darstellungsmatrix bzgl ON-Basis B. Fassen A als hermitesche Matrix $A \in \mathbb{C}^{n \times n}$ auf. Dann ist

$$\chi_f(X) = \chi_A(X) = \det(XE - A) = (X - \lambda_1) \cdots (X - \lambda_n)\,.$$

Wegen Teil 1 sind $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, also zerfällt $\chi_f(X)$ auch über $\mathbb{R}$.

Zu 3. Sei $f(u) = \lambda u$, $f(v) = \mu v$, $\lambda \neq \mu$. Dann

$$\begin{aligned}
\lambda(u * v) = \lambda u * v = f(u) * v & \\
= u * f(v) & \qquad \text{(da } f \text{ selbstadjungiert)} \\
= u * \mu v = \bar{\mu}(u * v) & \\
= \mu(u * v) & \qquad \text{(da } \mu \in \mathbb{R} \text{ nach Teil 1)}
\end{aligned}$$

Also $(\lambda - \mu)(u * v) = 0$. Da $\lambda \neq \mu$, ist $u * v = 0$, also $u \perp v$. □

8.3.5 Spektralzerlegung (selbstadjungierter Fall)

Titel wird erst später klar. In diesem Abschnitt Lösung des Klassifikationsproblems von symmetrischen/hermiteschen Matrizen bis auf orthogonale/unitäre Ähnlichkeit.

Satz 8.3.14. *Sei V ein endlichdimensionaler euklidischer (unitärer) VR, und $f \in$ End(V) selbstadjungiert. Dann existiert eine ON-Basis B von V aus Eigenvektoren von f (ein* Hauptachsensystem*); es gilt*

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{8.7}$$

wobei $\lambda_1, \dots, \lambda_n$ reelle Eigenwerte von f.

Beweis. Beweis per Induktion über $n := \dim V$.
Induktionsanfang: $n = 1$. Jeder Vektor $\neq \mathbf{0}$ ist Eigenvektor. Normieren liefert ON-Basis aus (einem) Eigenvektor.
Induktionsschritt: Sei $\dim V = n+1$ und Satz sei für Dimension n schon bewiesen. Nach Satz 8.3.13 (2) zerfällt χ_f in Linearfaktoren. Sei v_{n+1} Eigenvektor zu Eigenwert λ_{n+1}; o.B.d.A. $\|v_{n+1}\| = 1$ (sonst normieren).

$$U := \langle v_{n+1} \rangle$$

Behauptung: $U^\perp$ ist f-invariant, d.h., $x \in U^\perp \Rightarrow f(x) \in U^\perp$. Denn:

$$\begin{aligned} f(x) * v_{n+1} &= x * f(v_{n+1}) && (f \text{ selbstadjungiert}) \\ &= x * \lambda_{n+1} v_{n+1} && (v_{n+1} \text{ ist EV zu EW } \lambda) \\ &= \lambda_{n+1}(x * v_{n+1}) \end{aligned}$$

Also:

$$\begin{aligned} x \in U^\perp &\Rightarrow x * v_{n+1} = 0 \\ &\Rightarrow f(x) * v_{n+1} = 0 && \text{(siehe oben)} \\ &\Rightarrow f(x) \in U^\perp \end{aligned}$$

Wegen der Behauptung ist

$$f_0 := f|_{U^\perp} \in \mathrm{End}(U^\perp)$$

f_0 ist wie f selbstadjungiert. Da $\dim U^\perp = n$ hat $U^\perp$ nach Induktionsvoraussetzung eine Basis $B = (v_1, \dots, v_n)$ aus EW von f_0, und von f:

$$f(v_i) = f_0(v_i) = \lambda_i v_i$$

für $i \in \{1, \dots, n\}$. Da $v_i \perp v_{n+1}$ ist $(v_1, \dots, v_n, v_{n+1})$ eine ON-Basis von V aus Eigenvektoren von f. Eigenwerte sind reell nach Satz 8.3.13 (1). Aussage (8.7) folgt aus Lemma 4.3.15. □

Sei $A \in \mathbb{R}^{n \times n}$ ($A \in \mathbb{C}^{n \times n}$) und $B = (u_1, \dots, u_n)$ ein Hauptachsensystem für f_A, d.h., es gibt $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ so dass u_i ist EV von f_A zum EW λ_i. Sei S die Matrix mit den Spalten $u_1, \dots, u_n$. Dann ist

- S orthogonal (Satz 8.3.5), und
- $S^\top AS$ ($S^\top A\bar{S}$) Diagonalmatrix (Satz 8.3.14).

"A ist orthogonal diagonalisierbar". Es gilt

$$\begin{aligned}
A &= S\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}S^\top\\
&= \begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\begin{pmatrix}| & & |\\ u_1 & \dots & u_n\\ | & & |\end{pmatrix}\begin{pmatrix}- & u_1 & -\\ & \vdots & \\ - & u_n & -\end{pmatrix}\\
&= \begin{pmatrix}| & & |\\ \lambda_1 u_1 & \dots & \lambda_n u_n\\ | & & |\end{pmatrix}\begin{pmatrix}- & u_1 & -\\ & \vdots & \\ - & u_n & -\end{pmatrix} = \lambda_1 u_1 u_1^T + \dots + \lambda_n u_n u_n^\top
\end{aligned}$$

$$\text{und analog}\quad A = S\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\bar{S}^\top = \lambda_1 u_1 \bar{u}_1^T + \dots + \lambda_n u_n \bar{u}_n^\top .$$

Man spricht von der *Spektralzerlegung von A* (Spektrum: Eigenwerte).

Für $x \in \mathbb{R}^n$, setze $y := S^\top x$. Dann gilt

$$x^\top Ax = y^\top Dy = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2. \tag{8.8}$$

Die $n \times n$-Matrizen $P_i := u_i u_i^\top$ heißen *Projektionsmatrizen*: für $v \in \mathbb{R}^n$ ist

$$P_i v = p_{U_i}(v)$$

die Projektion von v auf die Gerade $U_i = \langle u_i \rangle$. Denn:

$$p_{U_i} = (v * u_i)u_i = (u_i * v)u_i = (u_i^\top v)u_i \;=\; \underbrace{u_i}_{\in\mathbb{R}^{n\times 1}}\;\underbrace{(u_i^\top v)}_{\in\mathbb{R}^{1\times 1}} = (u_i u_i^\top)v = P_i v.$$

Spektralzerlegung:

$$Av = \lambda_1 p_{U_1}(v) + \dots + \lambda_n p_{U_n}(v) \tag{8.9}$$

Jeder Summand liefert Anteil bezüglich $U_i = \mathbb{R}u_i$, den *Hauptachsen* des Systems.

Korollar 8.3.15. *Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn alle Eigenwerte positiv sind.*

Beweis. Nach Satz 8.3.14 gibt es eine ON-Basis B aus Eigenvektoren $u_1, \dots, u_n$; sei S die Matrix mit den Spalten $u_1, \dots, u_n$. Für $x \in \mathbb{R}^n$ setze $y := S^\top x$. Dann gilt (wie in 8.8)

$$x^\top Ax = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 \tag{8.10}$$

und daraus liest man die Aussage ab. □

Bemerkung 8.3.16. Aus Korollar 8.3.15 folgt insbesondere Proposition 8.1.5.

Übung 58. Eine Bilinearform $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ heißt *positiv semidefinit* (*negative semidefinit*) falls für alle $u \in V$ gilt $B(u,u) \geq 0$ ($B(u,u) \leq 0$). Zeigen Sie: eine symmetrische Bilinearform B ist genau dann positiv semidefinit, wenn alle Eigenwerte der Gramschen Matrix A von B reell und positiv sind.

Übung 59. Zeigen oder widerlegen Sie: eine symmetrische Bilinearform $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist genau dann positiv semidefinit, wenn für die Gramschen Matrix A von B gilt: für jedes $k \in \{1,\dots,n\}$ hat die Matrix, die aus den ersten k Elementen der ersten k Zeilen von A besteht, eine nicht-negative Determinante.

Übung 60. Gibt es $x, y \in \mathbb{R}$, so dass die Matrix

$$\begin{pmatrix} x & 1 & 0 \\ 1 & y & 0 \\ 0 & 0 & -x \end{pmatrix}$$

positiv semidefinit ist?

Übung 61. Eine symmetrische Bilinearform $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist genau dann positiv semidefinit, wenn sich die Gramsche Matrix A von B schreiben lässt als $A = C^\top C$ für ein $C \in \mathbb{R}^{n\times n}$.

Übung 62. Sei $G = (V,E)$ ein ungerichteter Graph. Die *Laplace-Matrix* $L \in \mathbb{Z}^{n\times n}$ von G ist definiert als $D - A$ wobei A die Adjazenzmatrix von G (siehe Abschnitt 3.2) und $D = (d_{ij})_{i,j\in V}$ die Gradmatrix von G, d.h., die Diagonalmatrix mit Einträgen

$$d_{i,j} = \begin{cases} \operatorname{grad}(v_i) := |\{v_k \mid \{v_i, v_k\} \in E\}| & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

Zeigen Sie:

- L ist stets positiv semidefinit.
- L hat stets den Eigenwert 0.

Hinweis: Zeigen Sie zunächst, dass für beliebiges $v \in \mathbb{R}^n$ gilt, dass

$$(Lv)_i = \sum_{j:(i,j)\in E} (v_i - v_j),$$

und dann, dass

$$v^\top L v = \sum_{(i,j)\in E, i<j} (v_i - v_j)^2.$$

8.3.6 Hauptachsentransformation

Gegeben: symmetrische (hermitesche) Matrix $A \in \mathbb{R}^{n\times n}$ ($\mathbb{C}^{n\times n}$).
Gesucht: Hauptachsensystem, i.e., eine ON-Basis von $V = \mathbb{R}^n$ ($\mathbb{C}^n$) bestehend aus Eigenvektoren von A.

Die Matrix S mit diesen Vektoren als Spalten ist dann orthogonal (unitär) und liefert Diagonalmatrix $D = S^\top AS$ ($D = \bar{S}^\top AS$).

Lösung: Wie bei Diagonalisierung (Abschnitt 4.3.4) blos mit Orthonormalisierung.

1. Berechnung der Eigenwerte $\lambda_1, \dots, \lambda_r$ von A.
2. a) Zu jedem λ_i Berechnung einer Basis des Eigenraums
$$\mathrm{Eig}_{\lambda_i}(A) = \mathrm{Kern}(A - \lambda_i E) = \mathrm{Lös}(A - \lambda_i E, \mathbf{0})$$
   b) Gram-Schmidtsches ON-Verfahren liefert ON-Basis für $\mathrm{Eig}_{\lambda_i}(A)$
3. Aneinanderreihung aller ON-Basen aus Schritt 2 (b) liefert ON-Basis $(u_1, \dots, u_n)$ von V, die nur aus Eigenvektoren besteht:
$$Au_i = \mu_i u_i \qquad \text{mit } \{\mu_1, \dots, \mu_n\} = \{\lambda_1, \dots, \lambda_r\}$$
$(u_1, \dots, u_n)$: Hauptachsensystem.

Bemerkung 8.3.17. Verfahren führt stets zur Lösung, denn

- A ist diagonalisierbar (da A symmetrisch / hermitesch);
- die zusammengesetzten Basen aus 2 (b) ergeben ON-Basis nach Abschnitt 8.2.3 (Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal, Satz 8.3.13).

Beispiel 8.3.18. Sei
$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$
(aus Abschnitt 4.3).

1. Eigenwerte. Nullstellen von $\det(XE_2 - A) = (X-3)^2 - 1$: $\lambda_1 = 2$ und $\lambda_2 = 4$.
2. a) Eigenräume. Algebraische Vielfachheit = geometrische Vielfachheit = 1.
$$\mathrm{Eig}_{\lambda_1}(A) = \langle v_1 \rangle \text{ für } v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mathrm{Eig}_{\lambda_2}(A) = \langle v_2 \rangle \text{ für } v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
   b) ON-Basen.
$$u_1 := \frac{v_1}{\|v_1\|} = \sqrt{2}/2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad u_2 := \frac{v_2}{\|v_2\|} = \sqrt{2}/2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
3. Hauptachsensystem ist $B = (u_1, u_2)$. △

Bemerkung 8.3.19. Effiziente Algorithmen zur Berechnung der Hauptachsentransformation und deren exakte Komplexität sind Gegenstand aktueller Forschung; es sei wieder verwiesen auf [4].

8.3.7 Kurven 2ter Ordnung und Kegelschnitte

Eine *Kurve 2ter Ordnung* (in der Ebene $\mathbb{R}^2$) ist eine Menge der Gestalt

$$K := \{(x_1, x_2) \in \mathbb{R}^2 \mid ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2 + f = 0\} \tag{8.11}$$

wobei $a, b, c, d, e, f \in \mathbb{R}$, und a, b, c nicht alle Null. Treten z.B. auf als Kennlinien von Bilinearformen, Abschnitt 8.1.3.

Beispiel 8.3.20. Die leere Menge für zum Beispiel für $a = 1$, $b = c = d = e = 0, f = -1$, denn $\emptyset = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 = -1\}$. △

Beispiel 8.3.21. $\mathbb{R}^2$ zum Beispiel für $a = b = c = d = e = f = 0$. △

Beispiel 8.3.22. Ein einzelner Punkt für zum Beispiel $a = c = 1$ und $b = d = e = f = 0$, denn $\{(0, 0)\} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 0\}$. △

Beispiel 8.3.23. Eine Gerade: zum Beispiel beschrieben durch $x_1^2 = 0$. △

Beispiel 8.3.24. Eine Ellipse: zum Beispiel, für $a, c > 0$,

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid ax_1^2 + cx_2^2 = 1\}.$$

Spezialfall $a = c$: Kreis. △

Beispiel 8.3.25. Eine Parabel: zum Beispiel

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 - x_2 = 0\}$$

(Scheitel im Punkt $(0, 0)$ und Achse auf der x_2-Achse). △

Beispiel 8.3.26. Eine Hyperbel: Zum Beispiel

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 - x_2^2 = 1\}$$

(Mittelpunkt $(0, 0)$ und Hauptachse x_2). △

Beispiel 8.3.27. Die Vereinigung von zwei sich schneidenden Geraden: z.B. beschrieben durch $x_1^2 - x_2^2 = 0$. △

Beispiel 8.3.28. Die Vereinigung von zwei parallelen Geraden: z.B. beschrieben durch $x_1^2 = 1$. △

Dies sind im wesentlichen alle Möglichkeiten. Präzisierung mit Hilfe der Hauptachsentransformation.

Geometrisch Kegelschnitte: Schnitt einer Ebene mit Doppelkegel.
Experiment: Taschenlampe auf Wand, welche Fläche sieht man?

Fall 1 (leere Menge) und 8 (Parallele Geraden): Schnitt von Ebene mit *Kreiszylinder* (Grenzfall eines Kegels mit Kegelspitze im Unendlichen).

Matrixdarstellung:

$$x^\top Ax + (d\ e)x + f = 0$$

für symmetrische Matrix

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

Alternative:

$$x^\top B x = 0$$

für

$$B = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}$$

Hauptachsentransformation: es gibt orthogonale Matrix S mit

$$D = S^\top A S = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

wobei λ_1, λ_2 Eigenwerte von A. Koordinatentransformation

$$x = Sx', \; x' = S^\top x$$

liefert (8.11) in neuen Koordinaten ($x^\top A x = x'^\top D x'$)

$$\lambda_1 (x_1')^2 + \lambda_2 (x_2')^2 + (g_1 x_1' + g_2 x_2') + h = 0$$

Falls $\lambda_i \neq 0$ kann durch Koordinatenwechsel auch noch das lineare Glied $g_i x_i'$ zum Verschwinden gebracht werden:

$$\lambda_i (x_i')^2 + g_i x_i' = \lambda_i \underbrace{(x_i' + \frac{g_i}{2\lambda_i})}_{=:x_i''}{}^2 - \frac{g_i^2}{4\lambda_i^2}$$

Fallunterscheidung:

1. $\lambda_1 \lambda_2 > 0$. Ellipse oder degenerierte Fälle: die leere Menge oder ein Punkt.
2. $\lambda_1 \lambda_2 = 0$: Parabel oder degenerierte Fälle: die leere Menge, eine Gerade, oder zwei parallele Geraden.
3. $\lambda_1 \lambda_2 < 0$. Hyperbel oder degenerierte Fälle: eine Gerade oder zwei sich schneidende Geraden.

Bemerkung 8.3.29. $\det A$ ändert sich nicht, wenn wir drehen und verschieben.

Bemerkung 8.3.30. Kurven zweiter Ordnung im $\mathbb{R}^2$ (siehe (8.11)) haben folgende natürliche Verallgemeinerung im $\mathbb{R}^n$: eine *Quadrik* ist eine Menge der Gestalt

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i,j=1}^{n} a_{i,j} x_i x_j + 2 \sum_{i=1}^{n} b_i x_i + c\}$$

für $a_{1,1}, \ldots, a_{n,n}, b_1, \ldots, b_n, c \in \mathbb{R}$, wobei mindestes einer der Koeffizienten $a_{1,1}, \ldots, a_{n,n}$ ungleich Null sein muss.

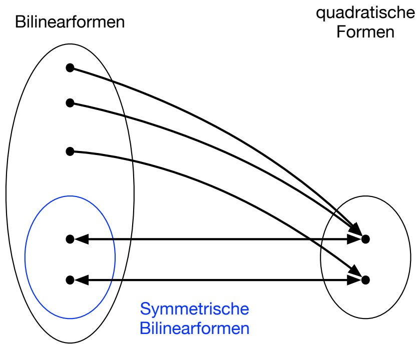


Abbildung 8.2: Zusammenhang Bilinearformen, quadratische Formen, und symmetrische Bilinearformen.

8.3.8 Klassifikation von quadratischen Formen

Klassifizieren quadratische Formen $q: \mathbb{R}^n \to \mathbb{R}$, und daher auch symmetrische Bilinearformen $B: \mathbb{R}^{n\times n} \to \mathbb{R}$ (Abschnitt 8.1.3 zum Zusammenhang quadratische Formen und Bilinearformen). Können q schreiben als

$$q(x) = x^T Ax = Ax * x$$

wobei $*$ das Standardskalarprodukt und $A \in \mathbb{R}^{n\times n}$ symmetrisch (siehe Abbildung 8.2). Denn: sei B Bilinearform mit

$$q(x) = B(x,x) = \sum_{1\leq i\leq j\leq n} B(e_i,e_j)x_i x_j.$$

Sei A die Matrix mit Einträgen $a_{i,i} := B(e_i,e_i)$ und $a_{i,j} = B(e_i,e_j)/2$ für $i \neq j$. Dann gilt

$$q(x) = B(x,x) = \sum_{1\leq i,j\leq n} a_{i,j}x_i x_j = x^\top Ax.$$

Also können wir eine ON-Basis von $\mathbb{R}^n$ finden, die A diagonalisiert, d.h.

$$A = SDS^{-1} = S\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} S^\top$$

für eine orthogonale Matrix $S \in \mathbb{R}^{n\times n}$. Schreiben y für $S^\top x$ (Koordinatenwechsel), und erhalten (wie in (8.8))

$$q(x) = x^\top SDS^\top = y^\top Dy = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

Klassifikation:

- Alle $\lambda_1, \ldots, \lambda_n$ positiv oder alle negativ: q ist *elliptisch*.
- Mindestens ein λ_i ist Null: q heißt *parabolisch*.
- Sonst (also alle λ_i ungleich Null, und es gibt sowohl positive als auch negative Werte): q heißt *hyperbolisch* (oder auch *indefinit*).

Wie kann man den Typ von q entscheiden, ohne die Nullstellen des charakteristischen Polynoms zu berechnen?

Lemma 8.3.31 (Vorzeichenregel von Descartes)**.** *Sei $\varphi \in \mathbb{R}[X]$ ein Polynom so dass*

$$\varphi(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 = (X - \lambda_1)\cdots(X - \lambda_n) \tag{8.12}$$

für $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

1. *0 ist genau dann Nullstelle von φ, wenn $a_0 = 0$.*
2. *Alle Nullstellen von φ sind negativ $\Leftrightarrow$ $a_{n-1}, \ldots, a_0 > 0$.*
3. *Falls n gerade ist:*
   alle Nullstellen von φ positiv $\Leftrightarrow$ $\underbrace{a_{n-1} < 0, a_{n-2} > 0, \ldots, a_1 < 0, a_0 > 0}_{\textit{alternierend}}$.
4. *Falls n ungerade ist:*
   alle Nullstellen von φ positiv $\Leftrightarrow$ $a_{n-1} < 0, a_{n-2} > 0, \ldots, a_1 > 0,\ a_0 < 0$.

Beispiel 8.3.32. Die Nullstellen von $X^3 - X^2 + X - 1$ und von $X^4 - X^3 + X^2 - X + 1$ sind alle positiv. $\triangle$

Beweis.

- Die erste Aussage ist klar (X ausklammern).
- Beweis der zweiten Aussage.
  $\Rightarrow$ folgt aus (8.12): falls $\lambda_1, \ldots, \lambda_n$ negativ sind, dann sind $a_{n-1}, \ldots, a_0$ positiv, denn ausmultiplizierte positive Ausdrücke haben positive Koeffizienten.
  $\Leftarrow$: wenn $a_0, a_1, \ldots, a_{n-1}$ positiv sind, dann ist $\varphi(t) > 0$ für alle nicht-negativen $t \in \mathbb{R}$, also sind alle Nullstellen von φ negativ.
- Die dritte Aussage:

$$\begin{aligned} & \text{Alle Nullstellen von } \varphi(X) \text{ positiv} \\ \Leftrightarrow\ & \text{Alle Nullstellen von } \varphi(-X) \text{ negativ} \\ \Leftrightarrow\ & \text{Koeffizienten von } \varphi(-X) \text{ positiv} && \text{(nach Teil 2)} \\ \Leftrightarrow\ & a_{n-1} < 0, a_{n-2} > 0, \ldots, a_1 < 0, a_0 > 0 \end{aligned}$$

  Hier ist Teil 2 anwendbar, da $\varphi(-X)$ weiterhin normiert, wenn n gerade ist.

- Die vierte Aussage: analog zur dritten. □

Mit dem einfachen ersten Kriterium in Lemma 8.3.31 können wir also einfach entscheiden, ob eine quadratische Form q parabolisch ist. Mit den übrigen Kriterien läßt sich feststellen, ob q elliptisch ist. Ansonsten ist q hyperbolisch.

Übung 63. Erklären Sie, warum die Terminologie zu *elliptischen, parabolischen, und hyperbolischen* quadratischen Formen zusammenpasst mit der Klassifikation von Kurven zweiter Ordnung in *Ellipsen, Parabeln, und Hyperbeln* (oder degenerierten Fällen).

8.3.9 Anwendung: Hauptkomponentenanalyse

Experiment: Probanden beantworten Persönlichkeitsfragen zu Ihnen bekannten Personen, z.B.: “Nimmt sich die Person Zeit für Andere?”, “Wird die Person schnell zornig”, etc., auf einer Skala von $\{1,\dots,10\}$ (“trifft überhaupt nicht zu”, ..., bis “trifft voll und ganz zu”)

Ω: Grundmenge aller möglichen Versuchsergebnisse (Annahme: endlich).
$P\colon \mathcal{P}^{\Omega} \to [0,1] \subset \mathbb{R}$: Wahrscheinlichkeitsmaß.

Seien $X_1,\dots,X_n\colon \Omega \to \mathbb{R}$ Zufallsvariablen.
In der Anwendung etwa: $X_i(\omega) = a$ falls im Versuchsergebnis $\omega \in \Omega$ die Frage i mit a geantwortet wird.
Schreiben $\vec{X}$ für die Zufallsvariable $(X_1,\dots,X_n)$, die $\omega \in \Omega$ abbildet auf

$$(X_1(\omega),\dots,X_n(\omega)) \in \mathbb{R}^n.$$

Summe und Produkt von Zufallsvariablen, oder von Zufallsvariablen und reellen Zahlen, sind punktweise definiert: Beispielsweise ist XY die Zufallsvariable, die $\omega \in \Omega$ abbildet auf $X(\omega)Y(\omega)$. Wir schreiben $X = i$ als Abkürzung für $\{\omega \in \Omega \mid X(\omega) = i\}$.

Definitionen für Zufallsvariablen X, Y:

- *Erwartungswert* von X:

  $$E[X] := \sum_{\omega\in\Omega} X(\omega)P(\{\omega\})$$

  Schätzung von $E[X_i]$ im Experiment: Für Stichproben $\omega_1,\omega_2,\dots,\omega_n$ berechne arithmetisches Mittel $\frac{1}{m}\sum_{j\in\{1,\dots,m\}} X_i(\omega_j)$.

- *Varianz* von X:

  $$V[X] := E[(X - E[X])^2]$$

- *Kovarianz* von X und Y:

  $$\mathrm{Cov}[X,Y] := E[(X - E[X]) \cdot (Y - E[Y])]$$

  Verallgemeinerung der Varianz $V[X] = \mathrm{Cov}[X,X]$.
  Erhalten Information über ‘Korrelation’ zwischen X und Y. Beispiel: die Frage

"Wird die Person schnell zornig" und "Hupt die Person häufig im Straßenverkehr" sind vermutlich positiv korreliert.

$$\begin{aligned}
\mathrm{Cov}[X,Y] &= E[(X - E[X])(Y - E[Y])] \\
&= E[XY - XE[Y] - E[X]Y + E[X]E[Y]] \\
&= E[XY] - E[XE[Y]] - E[E[X]Y] + E[E[X]E[Y]] && \text{(Linearität von } E[.]) \\
&= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] && (E[\mathrm{const}] = \mathrm{const}) \\
&= E[XY] - E[X]E[Y] && (8.13)
\end{aligned}$$

Insbesondere gilt also

$$V[X] = E[XX] - E[X]^2. \tag{8.14}$$

Falls X und Y unabhängig sind, gilt $\mathrm{Cov}[X,Y] = 0$, denn

$$\begin{aligned}
E[XY] &= \sum_{\omega\in\Omega} X(\omega)\cdot Y(\omega)\,\Pr[\{\omega\}] && \text{(Definition)} \\
&= \sum_i \big(i\cdot \Pr(XY = i)\big) && \text{(Summe endlich da } \Omega \text{ endlich)} \\
&= \sum_{k,l} \big(kl\cdot \Pr(X = k \text{ und } Y = l)\big) \\
&= \sum_{k,l} \big(kl\cdot \Pr(X = k)\Pr(Y = l)\big) && (X \text{ und } Y \text{ sind unabhängig}) \\
&= E[X]E[Y].
\end{aligned}$$

Erhalten Schätzung von $\mathrm{Cov}[X_i, X_j]$ aus dem Experiment: für Stichproben $\omega_1, \omega_2, \dots, \omega_m$ berechne

$$\frac{1}{m}\sum_{k\in\{1,\dots,m\}} (X_i(\omega_k) - E[X_i])(Y_j(\omega_k) - E[X_j]).$$

- Die *Kovarianzmatrix*: Matrix aller paarweisen Kovarianzen von $\vec{X} = (X_1, \dots, X_n)$

$$\mathrm{Cov}[\vec{X}] := \begin{pmatrix} \mathrm{Cov}[X_1,X_1] & \cdots & \mathrm{Cov}[X_1,X_n] \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_n,X_1] & \cdots & \mathrm{Cov}[X_n,X_n] \end{pmatrix} \in \mathbb{R}^{n\times n}$$

  $\mathrm{Cov}[\vec{X}]$ ist symmetrisch!
  Schätzung der Kovarianzmatrix ebenfalls symmetrisch.

Satz 8.3.14 liefert: $\mathrm{Cov}[\vec{X}]$ (bzw. Schätzung von $\mathrm{Cov}[\vec{X}]$) ist orthogonal diagonalisierbar! Bedeutung der Eigenwerte und -vektoren?

Lemma 8.3.33. *Die normierten Eigenvektoren zum größten Eigenwert von* $\mathrm{Cov}[\vec{X}]$ *sind genau die normierten Vektoren* $u \in \mathbb{R}^n$*, die die Varianz* $V[u^\top \vec{X}]$ *maximieren.*

Beweis. Es gilt

$$\begin{aligned} V[u^\top \vec{X}] &= E[u^\top \vec{X} u^\top \vec{X}] - E[u^\top \vec{X}]^2 && \text{(siehe (8.14))} \\ &= u^\top (E[\vec{X} u^\top \vec{X}] - E[\vec{X}]E[u^\top \vec{X}]) \\ &= u^\top (E[\vec{X}\vec{X}^\top] - E[\vec{X}]E[\vec{X}]^\top) u \\ &= u^\top \operatorname{Cov}[\vec{X}] u && \text{(see (8.13)).} \end{aligned}$$

Nach Satz 8.3.14 gibt es ein Hauptachsensystem $(v_1, \ldots, v_n)$ für $\operatorname{Cov}(\vec{X})$, so dass v_i ein EV von $\operatorname{Cov}(\vec{X})$ zum EW λ_i mit $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Sei S die Matrix mit Spalten $v_1, \ldots, v_n$. Also

$$\operatorname{Cov}(\vec{X}) = S \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} S^\top .$$

Sei $w := S^\top u$. Da u normiert und S orthogonal, ist w normiert! Dann gilt

$$\begin{aligned} u^\top \operatorname{Cov}(\vec{X}) u = u^\top S D S^\top u = w^\top D w \\ &= \sum_{i=1}^{n} w_i^2 \lambda_i && \text{(siehe (8.9))} \\ &\leq \lambda_1 (w_1^2 + \cdots + w_n^2) = \lambda_1 \|w\| = \lambda_1 . \end{aligned}$$

Also gilt für alle $u \in \mathbb{R}^n$, dass $V[u^\top \vec{X}] \leq \lambda_1$. Für den Vektor $u = v_1$ gilt Gleichheit, da dann

$$w = S^\top v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ und } w^\top D w = \lambda_1 .$$

Die normierten Eigenvektoren u zum größten Eigenwert von $\operatorname{Cov}[\vec{X}]$ maximieren also $V[u^\top \vec{X}]$. Umgekehrt, falls $u \in \mathbb{R}$ normiert so dass $V[u^\top \vec{X}]$ maximal, dann muss gelten $w^\top D w = \lambda_1$, also $w_i = 0$ falls $\lambda_i < \lambda_1$. Also ist w eine Linearkombination von $\{v_i \mid \lambda_i = \lambda_1\}$. Also ist auch u ein Eigenvektor zum Eigenwert λ_1. □

Betrachten in unserer Anwendung die fünf größten Eigenwerte und zugehörige Eigenvektoren. Haben *im wesentlichen*[3] psychologische Interpretation:

1. Extraversion (zurückhaltend und reserviert vs. gesellig),
2. Neurotizismus (selbstsicher und ruhig vs. emotional und verletzlich),

[3] Psychologische Modelle sind noch etwas komplizierter; ausgehend von den fünf größten Eigenwerten und deren Eigenvektoren wird im ‘Big Five’ Modell noch nach weiteren psychologisch relevanten Kriterien optimiert.

3. Offenheit für Erfahrungen (konservativ und vorsichtig vs. erfinderisch und neugierig),
4. Verträglichkeit (wettbewerbsorientiert und antagonistisch vs. kooperativ, freundlich, mitfühlend), und
5. Gewissenhaftigkeit (unbekümmert und nachlässig vs. effektiv und organisiert).

"The big five". Klassiker in der Psychologie. Ergebnis sehr stabil, z.B. bzgl. Veränderungen bei den Details des Experiments:

- andere Fragen,
- andere Skalen für die Antworten,
- andere Proband:innen,
- Fragen nicht über andere Personen, sondern über sich selbst, etc.

Zudem ist Ergebnis weitgehend kulturstabil.

Eine weitere psychologische Entdeckung ist, dass sich mit dem gleichen Ansatz zu Fragebögen für Intelligenztests ein mit Abstand größter Eigenwert findet. Der zugehörige Eigenvektor ist die *Definition* von Intelligenz (und der Begriff 'Intelligenz' erst danach in die Alltagssprache übergegangen).

Das Verfahren hat ebenfalls Anwendungen in Bilderkennung, Spracherkennung, maschinellem Lernen, etc. ("Clustering"). Die Hauptkomponentenanalyse ist eine *"explorative Faktorenanalyse"*, und nicht zu verwechseln mit *"konfirmatorischer Faktorenanalyse"*. Der Autor dankt Timo von Oertzen für Erklärungen zu den Anwendungen der Hauptkomponentenanalyse in der Psychologie.

Übung 64. Zeigen Sie, dass Kovarianzmatrizen stets positiv semidefinit sind (siehe Übung 58).

8.3.10 Spektralsatz

Klären nun: orthogonale Diagonalisierbarkeit.
Symmetrische Matrizen sind orthogonal diagonalisierbar, aber welche noch?
$f\colon V \to V$ diagonalisierbar gdw. V eine Basis hat aus Eigenvektoren von V.
Wann hat V eine ON-Basis aus Eigenvektoren von f?

Definition 8.3.34. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *normal* falls

$$A^\top \cdot A = A \cdot A^\top .$$

Analog heißt $B \in \mathbb{C}^{n\times n}$ *normal* falls

$$\bar{A}^\top \cdot A = A \cdot \bar{A}^\top .$$

Bemerkung 8.3.35.
- Symmetrische Matrizen mit $A^\top = A$ sind offensichtlich normal
$$A^\top \cdot A = A \cdot A = A \cdot A^\top$$
- Orthogonale (und hermitesche) Matrizen $A^{-1} = A^\top$ sind normal
$$A^\top A = A^{-1} A = E_n = AA^{-1} = AA^\top$$

Für $A \in \mathbb{C}^{n\times n}$ heißt die Matrix $\bar{A}^\top$ die *Adjungierte* von A. Falls $A \in \mathbb{R}^{n\times n}$ ist die Adjungierte gleich $A^\top$.

Lemma 8.3.36. *Sei $A \in \mathbb{C}^{n\times n}$ und $*$ das Standardskalarprodukt von $\mathbb{C}$. Dann gilt*
$$Ax * y = x * (\bar{A}^\top y)$$
(Übrigens: diese Eigenschaft charakterisiert die Adjungierte bereits eindeutig).

Beweis. Es gilt
$$\begin{aligned} Ax * y &= (Ax)^\top \bar{y} && \text{(Definition Standardskalarprodukt)}\\ &= x^\top A^\top \bar{y} && \text{(Rechenregel für Transposition)}\\ &= x^\top \overline{(\bar{A})^\top y)} && \text{(Rechenregel für Konjugation)}\\ &= x^\top * (\bar{A}^\top y) && \text{(Definition Standardskalarprodukt)} \end{aligned}$$
□

Folgendes geht natürlich auch wieder unitär ...

Für $A \in \mathbb{R}^{n\times n}$ gilt:
$$A = \mathbf{0} \Leftrightarrow (Ax * y = 0 \text{ für alle } x, y \in \mathbb{R}^n)$$
$\Rightarrow$ ist trivial, $\Leftarrow$: mit $y := Ax$ haben wir $Ax * Ax = 0$, und damit $Ax = \mathbf{0}$ für alle $x \in \mathbb{R}^n$, und damit $A = \mathbf{0}$ (Übung 12). Falls A symmetrisch ist, lässt sich mehr sagen:

Lemma 8.3.37. *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann gilt:*
$$A = \mathbf{0} \iff \forall x \in \mathbb{R}^n : Ax * x = 0$$

Beweis. $\Rightarrow$ ist trivial, $\Leftarrow$:
$$\begin{aligned} 0 &= A(x+y) * (x+y)\\ &= \underbrace{Ax * x}_{=0} + Ax * y + Ay * x + \underbrace{Ay * y}_{=0}\\ &= Ax * y + y * Ax && \text{da } A^\top = A\\ &= 2Ax * y \end{aligned}$$
Setze $y = Ax$, dann erhalten wir $0 = Ax * Ax = ||Ax||^2$ für alle $x \in \mathbb{R}^n$, also $A = \mathbf{0}$ (Übung 55). □

Bemerkung 8.3.38. Falls A nicht symmetrisch ist, gilt Lemma 8.3.37 im Allgemeinen nicht: für

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \text{‘schiefsymmetrische’ Matrix}$$

gilt für alle $x \in \mathbb{R}^2$:

$$Ax * x = -x * Ax = -Ax * x$$

also $Ax * x = 0$.

Proposition 8.3.39. *Eine Matrix $A \in \mathbb{R}^{n\times n}$ ($A \in \mathbb{C}^{n\times n}$) ist genau dann normal, wenn für alle $x \in \mathbb{R}^n$ ($x \in \mathbb{C}^n$)*

$$\|Ax\| = \|A^\top x\| \qquad \|Ax\| = \|\bar{A}^\top x\|$$

Beweis. Für alle $x \in \mathbb{R}^n$ gilt

$$\begin{aligned} &\|Ax\| = \|A^\top x\| \\ \Leftrightarrow\ &\|Ax\|^2 = \|A^\top x\|^2 \\ \Leftrightarrow\ &Ax * Ax = A^\top x * A^\top x \\ \Leftrightarrow\ &x * A^\top Ax = x * AA^\top x && \text{(Lemma 8.3.36)} \\ \Leftrightarrow\ &x * (A^\top A - AA^\top)x = 0 \end{aligned}$$

was genau dann der Fall ist, wenn $A^\top A = AA^\top$: Denn $\Leftarrow$ ist trivial, und $\Rightarrow$ folgt aus Lemma 8.3.37, da $A^\top A - AA^\top = (A^\top A)^\top - (AA^\top)^\top$ selbstadjungiert. □

Lemma 8.3.40. *Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann normal, wenn $BC = CB$ wobei $B = (A + A^\top)/2$ und $C = (A - A^\top)/2$.*

Beweis.

$$\begin{aligned} BC &= (A + A^\top)/2(A - A^\top)/2 \\ &= (A^2 - (A^\top)^2 + A^\top A - AA^\top)/4 \\ CB &= (A^2 - (A^\top)^2 - A^\top A + AA^\top)/4 \end{aligned}$$

Also

$$\begin{aligned} BC = CB \iff\ & A^\top A - AA^\top = -A^\top A + AA^\top \\ \iff\ & AA^\top = A^\top A \end{aligned}$$ □

Lemma 8.3.41. *Sei $A \in \mathbb{R}^{n\times n}$ und sei $M \subseteq \mathbb{R}^n$ invariant unter sowohl $f := f_A$ als auch $f^\top := f_{A^\top}$. Dann ist $M^\top$ ebenfalls f- und $f^\top$-invariant.*

Beweis. Sei $x \in M$ und $y \in M^\top$. Zu zeigen ist:

$$x * f(y) = 0 \qquad (\Rightarrow M^\top \text{ ist } f\text{-invariant})$$
$$x * f^\top y = 0 \qquad (\Rightarrow M^\top \text{ ist } f^\top\text{-invariant})$$

Einfach:

$$x * f(y) = f^\top(x) * y = 0 \qquad (\text{da } M\ f^\top\text{-invariant})$$
$$x * f^\top y = f(x) * y = 0 \qquad (\text{da } M\ f\text{-invariant})$$

□

Satz 8.3.42 (Spektralsatz). *Sei V euklidischer VR,* $\dim V = n$*, und* $f \in \operatorname{End}(V)$*, und* $A := M_B^B(f)$ *bezüglich einer Basis B von V. Die folgenden Aussagen sind äquivalent:*

1. *A ist orthogonal diagonalisierbar.*
2. *Es existiert eine ON-Basis von $\mathbb{R}^n$ aus Eigenvektoren von f;*
3. *Das charakteristische Polynom χ_f zerfällt in Linearfaktoren und A ist normal;*

Beweis. (2) ⇒ (1): wissen bereits (Satz 4.3.19): f ist diagonalisierbar.
φ_B für $B = (b_1, \ldots, b_n)$ ist der kanonische Basisisomorphismus

$$(x_1, \ldots, x_n) \mapsto x_1 b_1 + \cdots + x_n b_n$$

Für ON-Basis $(w_1, \ldots, w_n)$ aus Eigenvektoren von f seien

$$u_1 := \varphi_B^{-1}(w_1), \ldots, u_n := \varphi_B^{-1}(w_n)$$

die Koordinatenvektoren und $S := \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}$ ist die gesuchte Transformationsmatrix mit $D = S^{-1}AS$. Es gilt nun

$$\delta_{ij} = w_i * w_j \qquad (\text{da } (w_1, \ldots, w_n) \text{ ON-Basis})$$
$$= u_i^\top u_j \qquad (\text{da } * \text{ Standardskalarprodukt})\,.$$

Also $S^\top S = E$, d.h., $S^\top = S^{-1}$.

(1) ⇒ (2): wenn $u_1, \ldots, u_n$ Spalten von S, dann ist $w_1 := \varphi_B(u_1), \ldots, w_n := \varphi_B(u_n)$ ON-Basis aus Eigenvektoren wegen $u_i^\top u_j = w_i * w_j$.

(1) ⇒ (3): Falls $D = S^{-1}AS = S^\top AS$ diagonal, dann ist auch $D^\top$ diagonal. Für solche Matrizen gilt $D^\top D = D^\top D$. Da $D^\top = S^\top A^\top S$ und damit $A^\top = SD^\top S^\top$ haben wir

$$\begin{aligned} AA^\top = SDS^\top SD^\top S^\top &= SDD^\top S^\top \\ &= SD^\top DS^\top = SD^\top S^\top SDS^\top = A^\top A \end{aligned}$$

Also ist A normal. Und: $\chi_A = \chi_D$ zerfällt in Linearfaktoren (Satz 4.3.19).

(3) ⇒ (2): Ähnlich zum Beweis von Satz 8.3.14.
(Idee war: zeige f_A-Invarianz vom orthogonalen Komplement zu einem Eigenvektor.)
Setze $f^\top := f_{A^\top}$.
Ziel: Finde Eigenvektor x so dass $M := \langle x \rangle$ sowohl f_A als auch $f^\top$-invariant.

- Setze $B := (A + A^\top)/2$ und $C := (A - A^\top)/2$. Lemma 8.3.40: $BC = CB$.
- Wenden Satz 8.3.14 (Spektralzerlegung) auf die *symmetrische* Matrix B an: finden $\alpha \in \mathbb{R}$ und $y \in K := \mathrm{Kern}(B - \alpha E_n)$. Dann gilt $f_C(K) \subseteq K$. Sei $x \in K$. Zu zeigen ist, dass $Cx \in K$.
$$\begin{aligned}(B - \alpha E_n)(Cx) &= BCx - \alpha Cx \\ &= CBx - C\alpha x \\ &= C(B - \alpha E_n)x = 0\,.\end{aligned}$$
- Wenden Satz 8.3.14 auf symmetrische Matrix $(f_C)|_K$ an: finden $x \in K$ mit $Cx = \beta x$. Dann gilt
$$Ax = Bx + Cx = \alpha x + \beta x = (\alpha + \beta)x,$$
also ist x ein EV von f.
- Nach Lemma 8.3.41 ist $M^\top$ sowohl f_A- als auch $f^\top$-invariant.
- $f|_{M^\top} : M^\top \to M^\top$ wieder normal.

Induktion wie im Beweis von Satz 8.3.14. □

Analog erhält man den folgenden Satz für unitäre Vektorräume.

Satz 8.3.43. *Es sei $A \in \mathbb{C}^{n\times n}$. Die folgenden Aussagen sind äquivalent:*

1. *A ist unitär ähnlich zu einer Diagonalmatrix. (Man sagt, A ist* unitär diagonalisierbar*.)*
2. *Es existiert eine ON-Basis von $\mathbb{C}^n$ aus Eigenvektoren von A.*
3. *A ist normal.*

Korollar 8.3.44 (Klassifikation bis auf unitäre Ähnlichkeit)**.** *Zwei normale Matrizen sind genau dann unitär ähnlich, wenn sie die gleichen Eigenwerte (mit den gleichen Vielfachheiten) haben.*

Korollar 8.3.45 (Normalform unitärer Matrizen)**.** *Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann unitär ($A^{-1} = \bar{A}^\top$), wenn sie zu einer Diagonalmatrix unitär ähnlich ist, deren Diagonalelemente alle den Betrag 1 haben, d.h., $\exists S \in U(n)$ mit*
$$\bar{S}^\top AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
mit $|\lambda_i| = 1$.

Beweis. $\Rightarrow$: Falls A unitär ist, dann auch normal, und unitäre Diagonalisierbarkeit folgt direkt aus Satz 8.3.43. Aussage folgt, da alle Eigenwerte von unitären Matrizen Betrag 1 haben (Satz 8.3.2).

$\Leftarrow$: Für Diagonalmatrizen D ist $\bar{D}^\top = \bar{D}$. Falls alle Diagonalelemente Betrag 1 haben, gilt $D\bar{D} = E$: denn für alle $d \in \mathbb{C}$ gilt $d\bar{d} \in \mathbb{R}$, und falls $\|d\| = 1$ so folgt $|d| = 1$. Also ist D unitär. Und damit auch $A = SD\bar{S}^\top$ als Produkt unitärer Matrizen. □

8.4 Der Silverstersche Trägheitssatz

Im Abschnitt 7.1.4 haben wir Matrizen bis auf Ähnlichkeit klassifiziert:

$$A \approx B \text{ genau dann, wenn es invertierbares } S \text{ gibt mit } B = S^{-1}AS.$$

Im Abschnitt 8.3.3 dann bis auf orthogonale (beziehungsweise unitäre) Ähnlichkeit:

$$A \approx_{\text{ortho}} B \text{ genau dann, wenn es orthogonales } S \text{ gibt mit } B = S^{-1}AS = S^\top AS.$$

Eine weitere interessante Äquivalenzrelation ist auf $\mathbb{R}^{n\times n}$ definiert durch

$$A \approx_\top B \text{ genau dann, wenn es invertierbares } S \text{ gibt mit } B = S^\top AS.$$

In diesem Fall werden A und B auch *kongruent* genannt. Die Bedeutung von Kongruenz aus der Sicht der Bilinearformen $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ erklärt sich durch den folgenden Zusammenhang.

Lemma 8.4.1. *Seien $A, A' \in \mathbb{R}^{n\times n}$. Es gilt genau dann $A \approx_\top A'$, wenn A und A' die Gramschen Matrizen von derselben Bilinearform sind.*

Beweis. Für den Beweis der Rückrichtung sei $S = (e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^n$, und $T = (v_1, \ldots, v_n)$ eine andere Basis, d.h., $P := \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ ist invertierbar. Sei $A = (a_{ij})$ Gramsche Matrix von B bezüglich S, also $a_{ij} = B(e_i, e_j)$. Dann ist $A' \approx_\top P^\top AP$ die Gramsche Matrix von B bezüglich T, denn

$$\begin{aligned} B(v_i, v_j) &= v_i^\top A v_j \\ &= (Pe_i)^\top A(Pe_j) \\ &= e_i^\top (P^\top AP)e_j. \end{aligned}$$

Der Beweis der Vorwärtsrichtung ergibt sich aus der gleichen Rechnung. □

Bemerkung 8.4.2. Kongruente Matrizen haben den gleichen Rang.

Bemerkung 8.4.3. Die Eigenwerte einer quadratischen Matrix sind im Allgemeinen nur unter Ähnlichkeit von Matrizen invariant, nicht aber unter $\approx_\top$.

Sei V ein n-dimensionaler euklidischer Vektorraum und $A \in \mathbb{R}^{n\times n}$ symmetrisch (bzw. $f\colon V \to V$ selbstadjungierte Abbildung). Dann sind alle Eigenwerte $\lambda_1, \ldots, \lambda_n$ reell und

$$\chi_f(X) = (\lambda_1 - X) \cdot \cdots \cdot (\lambda_n - \lambda).$$

Seien

$$\begin{aligned} n_+ &:= \text{Anzahl der } \lambda_i > 0 \\ n_- &:= \text{Anzahl der } \lambda_i < 0 \\ n_0 &:= \text{Anzahl der } \lambda_i = 0 \end{aligned}$$

Dann heißt (n_+, n_-, n_0) bzw. (n_+, n_-) die *Signatur* (oder der *Typ*, oder die *Trägheit*[4]) von A (bzw f), bzw. die Signatur der Bilinearform $B(x,y) = x^T Ay$, bzw. die Signatur der quadratischen Form $q(x) := x^T Ax$.

$$\begin{aligned} n_+ + n_- &= \mathrm{rg}(A) \\ n_+ + n_- + n_0 &= n = \dim V \end{aligned}$$

Satz 8.4.4 (Sylvesterscher Trägheitssatz[5])**.** *Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann gilt*

$$A \approx_\top D_{n_+,n_-,n_0} := \begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} n_+ \\ \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} n_- \\ \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} n_0 \end{array}$$

Weiterhin: falls $(n'_+, n'_-, n'_0) \neq (n_+, n_-, n_0)$, dann gilt $D_{n'_+,n'_-,n'_0} \not\approx_\top D_{n_+,n_-,n_0}$. Insbesondere gilt also genau dann $A \approx_\top A'$, wenn A und A' dieselbe Signatur besitzen (Klassifikation durch charakteristische Daten, Abschnitt 7.1.1).

Das heißt, es gibt eine invertierbare Matrix P so dass $D := P^\top AP$. Die Spalten von P bilden eine Basis von $\mathbb{R}^n$; diese heißt *Sylvesterbasis* der Bilinearform $x * y := x^\top Ay$.

Beweis. Nach Satz 8.3.14 gibt es eine orthogonale Matrix S mit

$$D := S^\top AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

wobei $\lambda_1, \dots, \lambda_n$ die EW von A. Ohne Beschränkung der Allgemeinheit sind $\lambda_1, \dots, \lambda_{n_+}$

[4] *"This constant number of positive signs which attaches to a quadratic function under all its transformations (...) may be termed conveniently its inertia, until a better word is found."* (from Sylvester's article "On the Theory of the Syzygetic Relations".

[5] *"(...) my view of the physical meaning of quantity of matter inclines me, upon the ground of analogy, to give [this law] the name of the Law of Inertia for Quadratic forms, as expressing the fact of the existence of an invariable number inseparably attached to such forms."* from Sylvester's article "On a Theory of the Syzygetic Relations of Two Rational Integral Functions, Comprising an Application to the Theory of Sturm's Functions, and That of the Greatest Algebraical Common Measure".

positiv, $\lambda_{n_++1}, \ldots, \lambda_{n_++n_-}$ negativ, und $\lambda_{n_++n_-+1} = \cdots = \lambda_n = 0$. Seien

$$\begin{aligned}
\alpha_i &:= \frac{1}{\sqrt{\lambda_i}} && \text{für } i \in \{1, \ldots, n_+\} \\
\alpha_i &:= \frac{1}{\sqrt{-\lambda_i}} && \text{für } i \in \{n_+ + 1, \ldots, n_+ + n_-\} \\
\alpha_i &:= 1 && \text{für } i \in \{n_+ + n_- + 1, \ldots, n\}
\end{aligned}$$

Setze

$$Q := \begin{pmatrix} \alpha_1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \alpha_{n_+} & & & & & & \\ & & & \alpha_{n_++1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \alpha_{n_++n_-} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 1 \end{pmatrix}.$$

Q ist invertierbar (denn alle Werte auf der Diagonalen sind ungleich 0) und $Q^\top = Q$ (aber $Q^\top \neq Q^{-1}$!). Es gilt

$$Q^\top D Q = \begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 0 \end{pmatrix}$$

denn

- für $i \in \{1, \ldots, n_+\}$ gilt $\alpha_i \lambda_i \alpha_i = \frac{\lambda_i}{\sqrt{\lambda_i}^2} = 1$,
- für $i \in \{n_+ + 1, \ldots, n_-\}$ gilt $\alpha_i \lambda_i \alpha_i = \frac{\lambda_i}{\sqrt{-\lambda_i}^2} = -1$,
- für $i \in \{n_-, \ldots, n\}$ gilt $\alpha_i \lambda_i \alpha_i = \frac{\lambda_i}{1} = 0$.

Also folgt für $P := SQ \in \mathrm{GL}(n, \mathbb{R})$

$$P^\top A P = (SQ)^\top A (SQ) = Q^\top S^\top A S Q = Q^\top D Q$$

die angegebene Normalform.

Für die Eindeutigkeit nehmen wir an, dass $D_{n'_+,n'_-,n'_0} \approx_\top D_{n_+,n_-,n_0}$. Dann gilt $n'_0 = n_0$, da kongruente Matrizen denselben Rang haben (Bemerkung 8.4.2). Da

$$n_+ + n_- + n_0 = \dim V = n'_+ + n'_- + n'_0,$$

genügt es zu zeigen, dass $n_+ = n'_+$. Die Matrizen D_{n_+,n_-,n_0} und $D_{n'_+,n'_-,n'_0}$ sind nach Lemma 8.4.1 Gramschen Matrizen derselben Bilinearform $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Wir stellen fest, dass n_+ gleich der Dimension d ist des größten Untervektorraums $U \leq \mathbb{R}^n$, so dass $B|_{U\times U}$ positiv definit ist (siehe Korollar 8.3.15): Klarerweise ist B positiv definit auf $\langle e_1, \ldots, e_{n_+}\rangle$, also gilt $d \geq n_+$. Auf $V := \langle e_{n_++1}, \ldots, e_n\rangle$ dagegen ist B negativ semidefinit (siehe Übung 58), also gilt $U \cap V = \{\mathbf{0}\}$. Das impliziert, dass $d \leq n - \dim(V) = n - n_- - n_0 = n_+$. Es gilt also $p = n_+$. Also auch $n'_+ = p$, und $(n_+, n_-, n_0) = (n'_+, n'_-, n'_0)$. □

8.5 Singulärwertzerlegung

Zur Einordnung:

	Klassisch, S (und T) invertierbar	S (und T) orthogonal/unitär
$S^{-1}AS$	Ähnlichkeit Abschnitt 7.1.4	Orthogonale/unitäre Ähnlichkeit Abschnitt 8.3
SAT	Äquivalenz Abschnitt 7.1.2	Orthogonale/unitäre Äquivalenz

Vorteile:

- Auch anwendbar für lineare Abbildungen zwischen Vektorräumen V, W (mit Skalarprodukt) verschiedener (endlicher) Dimension;
- effiziente numerische Verfahren, große Bedeutung in der numerischen Mathematik;
- mathematischer Kern der Hauptkomponentenanalyse in der multivariaten Statistik, mit Anwendungen in der Datenkompression.

Satz 8.5.1. *Sei $A \in \mathbb{R}^{n\times m}$ ($A \in \mathbb{C}^{n\times m}$). Dann gibt es ON-Basis $B = (e_1, \ldots, e_m)$ von $\mathbb{R}^m$ und $C = (f_1, \ldots, f_n)$ von $\mathbb{R}^n$ so dass für ein $k \leq m$ gilt:*

$$f(e_1) = \lambda_1 f_1, \ldots, f(e_k) = \lambda_k f_k$$
$$f(e_{k+1}) = \cdots = f(e_m) = 0$$
$$M_C^B(f_A) = \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_n & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}.$$

$\lambda_1, \dots, \lambda_n$: sind natürlich keine Eigenwerte! Aber ein guter Ersatz dafür.

Beweis. Wenden den Spektralsatz (Satz 8.3.14) auf die symmetrische Matrix $A^\top A$ an. Erhalten ON-Basis $e_1, \dots, e_m$ für $\mathbb{R}^m$ so dass $A^\top A e_i = \sigma_i e_i$. Dann gilt:

$$Ae_i * Ae_j = A^\top Ae_i * e_j = \sigma_i e_i * e_j = \sigma_i \delta_{ij}\,.$$

Sortieren um, so dass $\sigma_1, \dots, \sigma_k \neq 0$.
Definieren $f_i := \frac{Ae_i}{\|Ae_i\|}$ für $i \in \{1, \dots, k\}$.
Ergänzen $f_1, f_2, \dots, f_k$ zu einer ON-Basis $f_1, \dots, f_n$ von $\mathbb{R}^n$. □

8.6 Übersicht Äquivalenzrelationen

In Abbildung 8.3 findet sich eine Übersicht zu Äquivalenzrelationen auf Matrizen, in der Reihenfolge ihrer Einführung im Skript.

Symbol	Definition	Name	Normalform / charakteristische Daten	Motivation
$A \sim B$	$A = SBT$, $S, T \in \mathrm{GL}(n, \mathbb{R})$	Äquivalenz	Rang	fundamental, Abschnitt 3.2.2
$A \approx B$	$A = S^{-1}BS$, $S \in \mathrm{GL}(n, \mathbb{R})$	Ähnlichkeit	Frobenius-NF Jordan-NF	$f \in \mathrm{End}(V)$ klassifizieren
	$A = SB$, $S \in \mathrm{GL}(n, \mathbb{R})$	Zeilen-äquivalenz	Reduzierte Zeilenstufenform	$\mathrm{Lös}(A, b)$ über $\mathbb{K}$, Abschnitt 3.3.4
	$A = BS$, $\det S$ Einheit	Unimodulare Spalten-äquivalenz	Hermite-NF	$\mathrm{Lös}(A, b)$ über $\mathbb{Z}$, Ideen für Smith-NF
	$A = SBT$, $\det S$ und $\det T$ Einheiten	Unimodulare Äquivalenz	Smith-NF	AL10, Frobenius-NF ausrechnen
$A \approx_{\mathrm{orth}} B$	$A = S^\top BS = S^{-1}BS$, S orthogonal	Orthogonale Ähnlichkeit	für normale Matrizen: Spektralsatz	Quadratische Formen klassifizieren, Hauptkomponentenanalyse, Optimierung, Datenreduktion, ...
$A \approx_\top B$	$A = S^\top BS$, $S \in \mathrm{GL}(n, \mathbb{R})$	Kongruenz	Signatur	Symmetrische Bilinearformen klassifizieren
	$A = S^\top BT$, S, T orthogonal	Orthogonale Äquivalenz	Singulärwert-zerlegung	Wie bei $\approx_{\mathrm{orth}}$, aber numerisch robuster

Abbildung 8.3: Äquivalenzrelationen auf Matrizen, entsprechende Normalformen, und Anwendungen

Literaturverzeichnis

[1] A. Beutelspacher. *Lineare Algebra.* Vieweg Lehrbuch Mathematik, 1994.

[2] M. Bodirsky. Introduction to mathematical logic, 2023. Course notes, TU Dresden, `https://wwwpub.zih.tu-dresden.de/~bodirsky/Logic.pdf`.

[3] T.-W. J. Chou and G. E. Collins. Algorithms for the solution of systems of linear Diophantine equations. *SIAM Journal on Computing*, 11:687–708, 1982.

[4] P. Dey, R. Kannan, N. Ryder, and N. Srivastava. Bit Complexity of Jordan Normal Form and Polynomial Spectral Factorization. In Y. Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 42:1–42:18, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[5] G. Fischer. *Lineare Algebra.* Grundkurs Mathematik. Vieweg Verlag, Braunschweig, 7 edition, Sept. 1981. Von diesem Buch gibt es diverse aktuellere Auflagen: Seit Oktober 2009 ist die 17. Auflage (ISBN 978-3834809964) erhältlich.

[6] 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.

[7] R. Kannan and A. Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. *SIAM Journal on Computing*, 8(4):499–507, 1979.

[8] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[9] P. Petersen. *Linear Algebra.* Undergraduate Texts in Mathematics. Springer Science & Business Media, 2012. 390 pages.

[10] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley - Interscience Series in Discrete Mathematics and Optimization, 1998.

[11] D. Serre. *Matrices: Theory and Applications*. Springer, 2nd ed., 2010 edition, 2002. Graduate Texts in Mathematics 216.

[12] J. von zur Gathen and D. Panario. Factoring polynomials over finite fields: A survey. *Journal of Symbolic Computation*, 31(1):3–17, 2001.