

Vorlesungsskript

LinA I* WiSe 23/24

Inhaltsverzeichnis

1. Motivation und mathematische Grundlagen	3
1.1. Mengen	3
1.2. Relationen	8
1.3. Abbildungen	12
2. Algebraische Strukturen	17
2.1. Gruppen	17
2.2. Ringe	21
2.3. Körper	23
2.4. Vektorräume	26
3. Basen und Dimensionen von Vektorräumen	32
3.1. Lineare Unabhängigkeit	32
3.2. Basen	39
3.3. Dimensionen	43
3.4. Direkte Summen	45
4. Lineare Abbildungen	50
4.1. Definitionen und grundlegende Eigenschaften	50
4.2. Kern, Bild und Rang von linearen Abbildungen	54
5. Matrizen	62
5.1. Definitionen und Basisoperationen	62
5.2. Matrizengruppen und -ringe	67
5.3. Matrizen und lineare Abbildungen	69
5.4. Basiswechsel	74
5.5. Elementarumformung von Matrizen	77
6. Lineare Gleichungssysteme	80
6.1. Existenz von Lösungen und Lösungsmengen	80
6.2. Der Gauß-Algorithmus	84
7. Determinanten von Matrizen	92
7.1. Definitionen und grundlegende Eigenschaften	92
7.2. Existenz und Eindeutigkeit	100

Definitionen

1 .

- 1.1: Mengen
- 1.3: Teilmenge
- 1.6: Mengenoperationen
- 1.10: Kardinalität, Potenzmenge
- 1.12: Kartesisches Produkt
- 1.14: Relationen
- 1.16: reflexiv, symmetrisch,
antisymmetrisch, transitiv
- 1.19: Äquivalenzklassen,
Quotientenmenge
- 1.23: maximales Element, obere
Schranke
- 1.27: Abbildungen
- 1.29: Bild, Urbild
- 1.31: injektiv, surjektiv, bijektiv
- 1.33: Komposition
- 1.36: inverse Abbildung

3 .

- 3.1: lineare Unabhängigkeit
- 3.6: Span für unendliche Mengen
- 3.9: Erzeugendensystem
- 3.11: Basis
- 3.15: endlichdimensional,
unendlichdimensional
- 3.27: Dimension eines Vektorraums
- 3.32: Summe von Mengen
- 3.36: Direkte Summe

5 .

- 5.1: Matrix
- 5.3: Addition von Matrizen

2 .

- 2.1: innere Verknüpfung, Halbgruppe
- 2.3: neutrales Element
- 2.7: Gruppen
- 2.11: Homomorphismus
- 2.14: Kern
- 2.15: Ring
- 2.18: invertierbar
- 2.20: Körper
- 2.26: Vektorraum
- 2.31: Untervektorraum
- 2.34: Linearkombination

4 .

- 4.1: Lineare Abbildung
- 4.11: Kern, Bild
- 4.14: Rang einer Abbildung

6 .

- 6.1: Lineares Gleichungssystem
- 6.5: Affiner Unterraum

- 5.4: Multiplikation einer Matrix mit einem Skalar
- 5.5: Multiplikation von Matrizen
- 5.8: Transposition von Matrizen
- 5.19: Matrixdarstellung
- 5.22: Spaltenrang, Zeilenrang
- 6.7: Dimension affiner Unterräume

7 .

- 7.1: Determinante
- 7.4: Permutation
- 7.5: Transposition
- 7.7: Fehlstand
- 7.11: Charakteristik

1. Motivation und mathematische Grundlagen

Was ist lineare Algebra bzw. analytische Geometrie?

- analytische Geometrie:
Beschreibung von geometrischen Fragen mit Hilfe von Gleichungen, Geraden, Ebenen sowie die Lösungen von Gleichungen als geometrische Form
- lineare Algebra:
die Wissenschaft der linearen Gleichungssysteme bzw der Vektorräume und der linearen Abbildungen zwischen ihnen

Wozu braucht man das?

- mathematische Grundlage für viele mathematische Forschung z.B. in der algebraischen Geometrie, Numerik, Optimierung
- viele Anwendungen z.B. Page-Rank-Algorithmus, lineare Regression
- oder Optimierung:
linear: Beschreibung zulässiger Punkte als Lösung von (Un)-Gleichungen
nichtlinear: notwendige Optimalitätsbedingungen

1.1. Mengen

Der Mengenbegriff wurde von Georg Cantor (dt. Mathematiker, 1845-1918) eingeführt.

Definition 1.1: Mengen

Unter einer **Menge** verstehen wir jede Zusammenfassung M von bestimmten, wohlunterschiedenen Objekten x unserer Anschauung oder unseres Denkens, welche **Elemente** von M genannt werden, zu einem Ganzen.

Bemerkungen:

Für jedes Objekt x kann man eindeutig feststellen, ob es zu einer Menge M gehört oder nicht.

$$\begin{aligned}x \in M &\rightarrow x \text{ ist Element von } M \\x \notin M &\rightarrow x \text{ ist nicht Element von } M\end{aligned}$$

Beispiel 1.2: Beispiel für Mengen

- {rot, gelb, grün}
- {1, 2, 3, 4}
- $\mathbb{N} = \{1, 2, 3, \dots\}, \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$
- $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$
- $\mathbb{Q} = \{x \mid x = \frac{a}{b} \text{ mit } a \in \mathbb{Z} \text{ und } b \in \mathbb{N}\}$
- $\mathbb{R} = \{x \mid x \text{ ist reelle Zahl}\}$
- \emptyset bzw. $\{\} \hat{=}$ leere Menge

Definition 1.3: Teilmenge

Seien M, N Mengen.

1. M heißt **Teilmenge** von N , wenn jedes Element von M auch Element von N ist.

Notation: $M \subseteq N$

2. M und N heißen gleich, wenn $M \subseteq N$ und $N \subseteq M$ gilt.

Notation $M = N$

Falls das nicht gilt, schreiben wir $M \neq N$

M heißt **echte Teilmenge** von N , wenn $M \subseteq N$ und $M \neq N$ gilt.

Notation: $M \subset N$

Nutzt man die Aussagenlogik, kann man diese Definitionen Umformulieren zu:

- $M \subseteq N \iff (\forall x : x \in M \implies x \in N)$
- $M = N \iff (M \subseteq N \wedge N \subseteq M)$
- $M \subset N \iff (M \subseteq N \wedge M \neq N)$

Kommentare:

- \iff heißt "genau dann, wenn"
- \forall heißt "für alle"
- \wedge heißt "und"
- $:$ heißt "mit der Eigenschaft"

Satz 1.4: Für jede Menge M gilt:

$$1) M \subseteq M$$

$$2) \emptyset \subseteq M$$

$$3) M \subseteq \emptyset \implies M = \emptyset$$

Beweis:

zu 1) Direkter Beweis (verwenden der Definitionen um Aussage zu folgern). Die Aussage:

$$x \in M \implies x \in M$$

folgt aus Def. 1.1. Daraus folgt aus Def 1.3, 1, dass $M \subseteq M$.

zu 2) Widerspruchsbeweis

Beweis der Aussage durch Annahme des Gegenteils und Herleitung eines Widerspruchs. Annahme: Es existiert eine Menge M , sodass $\emptyset \not\subseteq M$. Dann gilt: es existiert ein $x \in \emptyset$ mit $x \notin M$.

Aber: Die leere Menge enthält keine Elemente $\Rightarrow \nexists \Rightarrow$ Es existiert keine Menge M mit $\emptyset \subsetneq M \Rightarrow$ Behauptung

zu 3) Nach 2. $\emptyset \subseteq M$, wir wissen $M \subseteq \emptyset$. Nach Def. 1.3, 2 $\Rightarrow M = \emptyset$

□

Beispiel 1.5: Ob ein Objekt ein Element oder eine Teilmenge einer Mengen ist, ist vom Kontext abhängig. Betrachten wir folgende Menge:

$$M := \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$$

D.h. die Elemente dieser Menge M sind die natürlichen, ganzen, rationalen und reellen Zahlen. Damit gilt $\mathbb{N} \in M$ aber $\mathbb{N} \subset \mathbb{Z}$ und $\mathbb{N} \subset \mathbb{Q}$.

Definition 1.6: Mengenoperationen

Seien M, N Mengen.

1. Man bezeichnet die Menge der Elemente, die sowohl in M als auch in N enthalten sind, als **Durchschnitt** von M und N

$$M \cap N = \{x \mid (x \in M) \wedge (x \in N)\}$$

2. Man bezeichnet die Menge der Elemente, die entweder in M oder in N enthalten sind oder in beiden enthalten sind, als **Vereinigung** von M und N

$$M \cup N = \{x \mid (x \in M) \vee (x \in N)\}$$

3. Man bezeichnet die Menge der Elemente, die in M aber nicht in N enthalten sind, als **Differenz** von M und N

$$\begin{aligned} M \setminus N &= \{x \mid (x \in M) \wedge (x \notin N)\} \\ &= \{x \in M \mid x \notin N\} \end{aligned}$$

Beispiel 1.7:

Für $-\mathbb{N} := \{-n \mid n \in \mathbb{N}\}$ gilt:

- $\mathbb{N} \cup -\mathbb{N} = \mathbb{Z} \setminus \{0\}$
- $\mathbb{N} \cap -\mathbb{N} = \emptyset$

Wichtiges Beispiel für Mengen sind Intervalle reeller Zahlen

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}, a, b \in \mathbb{R}, a \leq b$$

Dies nennt man ein abgeschlossenes Intervall (die Grenzen sind enthalten). Sei jetzt $a, b \in \mathbb{R}, a \leq b$

$$[a, b[:= \{x \in \mathbb{R} \mid a \leq x < b\} \text{ oder }]a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$$

Diese Intervalle nennt man halboffene Intervalle (genau eine der Grenzen ist enthalten). Das Intervall

$$]a, b[:= \{x \in \mathbb{R} \mid a < x < b\}$$

heißt offenes Intervall (keine der Grenzen ist enthalten).

Für $M := \{4, 6, 8\}$ und $N := \{8, 10\}$ gilt:

- $M \cup N = \{4, 6, 8, 10\}$
- $M \cap N = \{8\}$
- $M \setminus N = \{4, 6\}$
- $N \setminus M = \{10\}$

Satz 1.8: Für zwei Mengen M, N gelte $M \subseteq N$. Dann sind folgende Aussagen äquivalent:

$$1) M \subset N$$

$$2) N \setminus M \neq \emptyset$$

Beweis:

Behauptung: 1) \iff 2)

zu zeigen: $1) \implies 2)$ und $2) \implies 1)$

1) \Rightarrow 2): Es gilt: $M \neq N$. Dann existiert $x \in N$ mit $x \notin M$. Dann gilt $x \in N \setminus M$. Also $N \setminus M \neq \emptyset$.

2) \implies 1): Es gilt $N \setminus M \neq \emptyset$. Dann existiert ein $x \in N$ mit $x \notin M$. Daher gilt $N \neq M$. Es gilt außerdem: $M \subseteq N$. Daraus folgt $M \subset N$.

9

Satz 1.9: Seien M, N, L Mengen. Dann gelten folgende Aussagen:

1. $M \cap N \subseteq M$ und $M \subseteq M \cup N$
2. $M \setminus N \subseteq M$
3. Kommutativgesetze:

$$M \cap N = N \cap M \text{ und } M \cup N = N \cup M$$

4. Assoziativgesetz:

$$\begin{aligned} M \cap (N \cap L) &= (M \cap N) \cap L \\ M \cup (N \cup L) &= (M \cup N) \cup L \end{aligned}$$

5. Distributivgesetze:

$$\begin{aligned} M \cup (N \cap L) &= (M \cup N) \cap (M \cup L) \\ M \cap (N \cup L) &= (M \cap N) \cup (M \cap L) \\ M \setminus (N \cap L) &= (M \setminus N) \cup (M \setminus L) \\ M \setminus (N \cup L) &= (M \setminus N) \cap (M \setminus L) \end{aligned}$$

Beweis: Es gilt $x \in M \cap N$ genau dann, wenn $x \in M \wedge x \in N$. Die Konjunktion zweier Aussagen ist symmetrisch bezüglich der Aussage. D.h. $A \wedge B \iff B \wedge A$. Es gilt also

$$(x \in M) \wedge (x \in N) \iff (x \in N) \wedge (x \in M)$$

Verwenden wir die Definition der Schnittmenge (1.6) so erhalten wir

$$(x \in N) \wedge (x \in M) \iff x \in N \cap M$$

Aus der Kette der Äquivalenzumformungen folgt $M \cap N = N \cap M$.

□

Etwas kompakter für das erste Distributivgesetz:

$$\begin{aligned}
 x \in M \cup (N \cap L) &\iff (x \in M) \vee (x \in N \cap L) \\
 &\iff (x \in M) \vee ((x \in N) \wedge (x \in L)) \\
 &\iff (x \in M \vee x \in N) \wedge (x \in M \vee x \in L) \\
 &\iff (x \in M \cup N) \wedge (x \in M \cup L) \\
 &\iff x \in (M \cup N) \cap (M \cup L)
 \end{aligned}$$

Damit ist $M \cup (N \cap L) = (M \cup N) \cap (M \cup L)$.

□

Die übrigen Aussagen zeigt man analog. *Übung*

Damit ist $M \cup N \cup L$ für die Mengen M, N, L wohldefiniert. Dies kann auf “viele” Mengen verallgemeinert werden:

Ist $I \neq \emptyset$ eine Menge und ist für jedes $i \in I$ eine Menge M_i gegeben, dann sind:

$$\begin{aligned}
 \bigcup_{i \in I} M_i &:= \{x \mid \exists i \in I \text{ mit } x \in M_i\} \\
 \bigcap_{i \in I} M_i &:= \{x \mid \forall i \in I \text{ mit } x \in M_i\}
 \end{aligned}$$

Die Menge I heißt auch **Indexmenge**. Für $I = \{1, \dots, n\}$ verwendet man auch die Notation

$$\begin{aligned}
 \bigcup_{i=1}^n M_i &:= \{x \mid \exists i \in I \text{ mit } x \in M_i\} \\
 \bigcap_{i=1}^n M_i &:= \{x \mid \forall i \in I \text{ mit } x \in M_i\}
 \end{aligned}$$

Definition 1.10: Kardinalität, Potenzmenge

Sei M eine endliche Menge, d.h. M enthält endlich viele Elemente.

Die **Mächtigkeit** oder **Kardinalität** von M , bezeichnet mit $|M|$ oder $\#M$ ist die Anzahl von Elementen in M .

Die **Potenzmenge** von M , bezeichnet mit $\mathcal{P}(M)$ ist die Menge aller Teilmengen von M . D.h.

$$\mathcal{P}(M) := \{N \mid N \subseteq M\}$$

Beispiel 1.11:

Die leere Menge \emptyset hat die Kardinalität Null. Es gilt $\mathcal{P}(\emptyset) = \{\emptyset\}$, $|\mathcal{P}(\emptyset)| = 1$.

Für $M = \{2, 4, 6\}$ gilt $|M| = 3$. $\mathcal{P}(M) = \{\emptyset, \{2\}, \{4\}, \{6\}, \{2, 4\}, \{2, 6\}, \{4, 6\}, \{2, 4, 6\}\}$.

Man kann zeigen: $|\mathcal{P}(M)| = 2^{|M|}$. Deswegen wird auch die Notation 2^M für die Potenzmenge von M verwendet.

1.2. Relationen

Definition 1.12: Kartesisches Produkt

Sind M und N zwei Mengen, so heißt die Menge

$$M \times N := \{(x, y) \mid x \in M \wedge y \in N\}$$

das **kartesische Produkt** von M und N .

Sind n Mengen M_1, \dots, M_n gegeben, so ist deren kartesisches Produkt gegeben durch:

$$M_1 \times \dots \times M_n := \{(x_1, \dots, x_n) \mid x_1 \in M_1 \wedge \dots \wedge x_n \in M_n\}$$

Das n -fache kartesische Produkt einer Menge von M ist:

$$M^n := M \times \dots \times M := \{(x_1, \dots, x_n) \mid x_i \in M \text{ für } i = 1, \dots, n\}$$

Ein Element $(x, y) \in M \times N$ heißt geordnetes Paar und ein Element

$(x_1, \dots, x_n) \in M_1 \times \dots \times M_n$ heißt **(geordnetes) n -Tupel**.

Ist mindestens eine der auftretenden Mengen leer, so ist auch das resultierende kartesische Produkt leer, d.h. die leere Menge. Das kartesische Produkt wurde nach Rene Decartes benannt. Rene Decartes war ein französische Mathematiker (1596-1650) und ein Begründer der analytischen Geometrie.

Beispiel 1.13: Das kartesische Produkt zweier Intervalle.

Seien $[a, b] \subset \mathbb{R}$ und $[c, d] \subset \mathbb{R}$ zwei abgeschlossene Intervalle von reellen Zahlen. Dann ist das kartesische Produkt beider Intervalle gegeben durch:

$$[a, b] \times [c, d] := \{(x, y) \mid x \in [a, b] \wedge y \in [c, d]\}$$

Das kartesische Produkt ist nicht kommutativ. Beweis durch Gegenbeispiel.

Definition 1.14: Relationen

Seien M und N nichtleere Mengen. Eine Menge $R \subseteq M \times N$ heißt **Relation** zwischen M und N . Ist $M = N$, so nennt man R **Relation auf M** . Für $(x, y) \in R$ schreibt man $x \sim_R y$ oder $x \sim y$, wenn die Relation aus dem Kontext klar ist. Ist mindestens eine der beiden Mengen leer, dann ist auch jede Relation zwischen den beiden Mengen die leere Menge.

Beispiel 1.15: Sei $M = \mathbb{N}$ und $N = \mathbb{Z}$. Dann ist

$$R := \{(x, y) \in M \times N \mid x + y = 1\}$$

eine Relation zwischen M und N . Es gilt

$$R = \{(1, 0), (2, -1), (3, -2), \dots\} = \{(n, -n + 1) \mid n \in \mathbb{N}\}$$

Definition 1.16: reflexiv, symmetrisch, antisymmetrisch, transitiv

Es sei M eine nicht leere Menge. Eine Relation auf M heißt:

1. reflexiv:

$$\forall x \in M : x \sim x$$

2. symmetrisch:

$$\forall x, y \in M : x \sim y \implies y \sim x$$

3. antisymmetrisch:

$$\forall x, y \in M : x \sim y \wedge y \sim x \implies x = y$$

4. transitiv:

$$\forall x, y, z \in M : x \sim y \wedge y \sim z \implies x \sim z$$

Falls die Relation R reflexiv, transitiv und symmetrisch ist, so nennt man R eine **Äquivalenzrelation** auf M . Ist R reflexiv, transitiv und antisymmetrisch, so nennt man R eine **partielle Ordnung** auf M .

Beispiel 1.17: $M = \mathbb{R}$

- Die Relation $<$ auf $M = \mathbb{R}$ ist transitiv, aber weder reflexiv noch symmetrisch und auch nicht antisymmetrisch.
- Die Relation \leq auf $M = \mathbb{R}$ ist reflexiv, antisymmetrisch und transitiv. Sie ist nicht symmetrisch. \leq ist somit eine partielle Ordnung.
- Die Relation $=$ auf \mathbb{R} ist reflexiv, symmetrisch und transitiv. Also ist $=$ eine Äquivalenzrelation. (Äquivalenzrelationen können auch antisymmetrisch sein)

Beispiel 1.18: Interpretiert man “Pfeile” als Objekte mit gleicher Orientierung und Länge, erhält man die Äquivalenzrelation

$$x \sim y :\iff x \text{ und } y \text{ haben die gleiche Länge und Orientierung}$$

Auf Grund der Transitivität sind somit alle Pfeile einer vorgegebenen Orientierung und Länge äquivalent zu dem Pfeil, der im Koordinatenursprung startet und die gleiche Länge sowie Orientierung besitzt. Somit können wir Vektor $x = (x_1, x_2) \in \mathbb{R}^2$ als Repräsentant einer ganzen Klasse von Pfeilen interpretieren. Alle zueinander äquivalente Pfeile haben gemeinsam, dass die Differenz zwischen End- und Anfangspunkt genau den Vektor x ergeben.

Als Formalisierung erhält man:

Definition 1.19: Äquivalenzklassen, Quotientenmenge

Sei \sim eine Äquivalenzrelation auf einer nichtleeren Menge M . Die Äquivalenzklasse eines Element $\bar{a} \in M$ ist definiert durch:

$$[\bar{a}] := \{a \in M \mid a \sim \bar{a}\}$$

Ist die Relation nicht aus dem Kontext klar, schreibt man $[\bar{a}]_{\sim}$.

Elemente einer Äquivalenzklasse werden als **Vertreter** oder **Repräsentanten** der Äquivalenzklasse bezeichnet. Die Menge aller Äquivalenzklassen einer Äquivalenzrelation \sim in einer Menge M , d.h.

$$M / \sim := \{[a]_{\sim} \mid a \in M\}$$

wird als **Faktormenge** oder **Quotientenmenge** bezeichnet.

Beispiel 1.20: (Fortsetzung von Beispiel 1.18)

Die Menge aller Pfeile gleicher Länge und Orientierung bilden eine solche Äquivalenzklasse, welche durch den Vektor $x = (x_1, x_2) \in \mathbb{R}^2$ repräsentiert wird. Die Menge der Vektoren $x = (x_1, x_2) \in \mathbb{R}^2$ bilden die Quotientenklasse.

Beispiel 1.21: Für eine gegebene Zahl $x \in \mathbb{N}$ ist die Menge:

$$R_n := \{(a, b) \in \mathbb{Z}^2 \mid a - b \text{ ist ohne Rest durch } n \text{ teilbar}\}$$

eine Äquivalenzrelation auf \mathbb{Z} , denn

- reflexiv: $a \sim a, (a, a) \in R_n : \Leftrightarrow a - a = 0 \checkmark$
- symmetrie:
 $a \sim b \Rightarrow (a, b) \in R_n \Rightarrow a - b \text{ ist ohne Rest teilbar} \Rightarrow a - b = k \cdot n$
 $\Rightarrow b - a = -k \cdot n \Rightarrow (b, a) \in R_n \Rightarrow b \sim a \checkmark$
- transitiv: zz: $a \sim b \wedge b \sim c \Rightarrow a \sim c$

$$a \sim b \Rightarrow a - b = k \cdot n$$

$$b \sim c \Rightarrow b - c = l \cdot n$$

$$\text{Gleichungen addieren: } a - c = n(k + l) \Rightarrow a \sim c \checkmark$$

Für a wird die Äquivalenzklasse $[a]$ auch die Restklasse von a modulo n genannt.

$$[a] = a + n \cdot z := \{a + nz \mid z \in \mathbb{Z}\}$$

Die Äquivalenzrelation R_n definiert auch eine Zerlegung der Menge \mathbb{Z} in disjunkte Teilmengen, nämlich

$$[0] \cup [1] \cup \dots \cup [n-1] = \bigcup_{a=0}^{n-1} [a] = \mathbb{Z}$$

Es gilt allgemein: Ist \sim eine Äquivalenzrelation auf M , so ist M die Vereinigung aller Äquivalenzklassen.

- “ \subseteq ”:

$$M = \bigcup_{a \in M} \{a\} \subseteq \bigcup_{a \in M} [a] \quad \checkmark$$

- “ \supseteq ”:

$$[a] \subset M \implies \bigcup_{a \in M} [a] \subseteq M \quad \checkmark$$

□

Satz 1.22: Ist R eine Äquivalenzrelation auf der Menge M und sind $a, b \in M$, dann sind folgende Aussagen äquivalent:

$$1) [a] = [b] \qquad 2) [a] \cap [b] \neq \emptyset \qquad 3) a \sim b$$

Beweis: Durch Ringschluss

zu zeigen: $1 \implies 2, 2 \implies 3, 3 \implies 1$

$1 \implies 2$:

$$\text{Wegen } a \sim a \implies a \in [a] = [b] \implies a \in [a] \cap [b] \implies [a] \cap [b] \neq \emptyset$$

$2 \implies 3$:

Aus $[a] \cap [b] \neq \emptyset \implies$ es existiert $c \in [a] \cap [b]$. Nach Definition gilt dann $c \sim a$ wegen der Symmetrie von $a \sim c$. Nach Definition auch $c \sim b$. Wegen der Transitivität der Relation gilt dann auch $a \sim b$

$3 \implies 1$:

$$\begin{aligned} \text{Es gilt } a \sim b. \text{ Sei } c \in [a] \implies c \sim a. \text{ Wegen der Transitivität folgt} \\ c \sim b \implies c \in [b] \implies [a] \subseteq [b]. \text{ Analog folgt } [b] \subseteq [a]. \end{aligned}$$

□

Aus Satz 1.22 2) folgt, dass die Äquivalenzklassen eine disjunkte Zerlegung der Menge M darstellen.

Definition 1.23: maximales Element, obere Schranke

Sei M eine Menge und sei für jedes Element $m \in M$ eine weitere Menge S_m gegeben. Für $\mathcal{S} := \{S_m \mid m \in M\}$ ist die Teilmengenrelation \subseteq eine partielle Ordnung. Die Menge \mathcal{S} heißt dann **partiell geordnet**. Eine Menge $\hat{S} \in \mathcal{S}$ heißt **maximales Element** von \mathcal{S} (bezüglich \subseteq), wenn aus $S \in \mathcal{S}$ und $\hat{S} \subseteq S$ folgt, dass $S = \hat{S}$ ist. Eine nichtleere Teilmenge $\mathcal{K} \subseteq \mathcal{S}$ heißt **Kette** (bezüglich \subseteq), wenn für alle $K_1, K_2 \in \mathcal{K}$ gilt, dass $K_1 \subseteq K_2$ oder $K_2 \subseteq K_1$. Ein Element $\hat{K} \in \mathcal{S}$ heißt **obere Schranke** der Kette \mathcal{K} , wenn $K \subseteq \hat{K}$ für alle $K \in \mathcal{K}$ gilt.

Beispiel 1.24: Sei $\mathcal{S} = P(\{2, 4, 6, 8, 10\})$

Dann ist $\mathcal{K} = \{\emptyset, \{2\}, \{2, 6\}, \{2, 6, 10\}\} \subseteq \mathcal{S}$ eine Kette.

Die Menge $K = \{2, 6, 10\}$ ist die obere Schranke von \mathcal{K} , das maximale Element von \mathcal{S} ist $\hat{S} = \{2, 4, 6, 8, 10\}$.

Lemma 1.25: Zornsche Lemma

Sei M eine Menge und sei $\mathcal{S} \subseteq \mathcal{P}(M)$ eine nichtleere Menge mit der Eigenschaft, dass für jede Kette $\mathcal{K} \subseteq \mathcal{S}$ auch ihre Vereinigungsmenge in \mathcal{S} liegt, d.h.

$$\bigcup_{A \in \mathcal{K}} A \in \mathcal{S}$$

Dann besitzt \mathcal{S} ein maximales Element.

Beweis: Das Zornsche Lemma ist ein fundamentales Resultat aus der Mengenlehre, hier ohne Beweis

□

Lemma 1.26: Sei M eine Menge und $\mathcal{K} \subseteq \mathcal{P}(M)$ eine Kette. Dann gibt es zu je endlich vielen $A_1, \dots, A_n \in \mathcal{K}$ ein $\hat{i} \in \{1, \dots, n\}$ mit $A_i \subseteq A_{\hat{i}}$ für alle $i \in \{1, \dots, n\}$.

Beweis: Durch vollständige Induktion über n .

Induktionsanfang: $n = 1$

D.h. wir haben $A_1 \in \mathcal{K}$ und für $\hat{i} = 1$ gilt $A_1 \subseteq A_{\hat{i}} = A_1$ ✓

Induktionsschritt: $n - 1 \mapsto n$

Für $A_1, \dots, A_{n-1} \in \mathcal{K}$ existiert ein $\hat{j} \in \{1, \dots, n-1\}$ mit $A_i \subseteq A_{\hat{j}}$ für alle $i \in \{1, \dots, n-1\}$. Mit

$$\hat{i} := \begin{cases} \hat{j} & \text{für } A_n \subseteq A_{\hat{j}} \\ n & \text{für } A_{\hat{j}} \subseteq A_n \end{cases}$$

folgt die Behauptung.

□

1.3. Abbildungen

Definition 1.27: Abbildungen

Es Seien X und Y beliebig, nichtleere Mengen. Eine **Abbildung** von X nach Y ist eine Vorschrift f , die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet. Man schreibt

$$f : X \rightarrow Y, \quad x \mapsto y = f(x)$$

Die Menge X heißt **Definitionsbereich** von f , die Menge Y heißt **Wertebereich** von f

Achtung: Jede Abbildung besteht aus drei "Teilen". Angabe des Definitionsbereichs, Angabe des Wertebereichs, Angabe der Zuordnungsvorschrift.

Beispiel 1.28: Sei M eine nichtleere Menge. Dann ist

$$f : M \rightarrow N, \quad x \mapsto x = f(x)$$

eine Abbildung f **Identität** von M mit der Notation I_m/Id_m .

Sei $X = Y = \mathbb{R}$, dann ist $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto f(x) := 7x + 2$ eine Abbildung.

Definition 1.29: Bild, Urbild

Seien X, Y beliebige nichtleere Mengen und $f : X \rightarrow Y$. Es gelte $M \subseteq X$ und $N \subseteq Y$. Dann heißen die Mengen:

$$f(M) := \{f(x) \in Y \mid x \in M\} \subseteq Y \text{ das } \mathbf{Bild} \text{ von } M \text{ unter } f.$$

$$f^{-1}(N) := \{x \in X \mid f(x) \in N\} \subseteq X \text{ das } \mathbf{Urbild} \text{ von } N \text{ unter } f.$$

Ist $\emptyset \neq M \subseteq X$, dann heißt $f|_M : M \rightarrow Y$, $x \mapsto f(x)$, die **Einschränkung** von f auf M .

Beispiel 1.30: Sei $X = Y = \mathbb{R}$ und $x \mapsto f(x) = x^4$. Dann ist \mathbb{R} Definitions- und Wertebereich von f .

- $f(\mathbb{R}) = \mathbb{R}_+ := [0, \infty[$ das Bild von f
- $f([0, 2]) = [0, 16]$
- $f^{-1}([16, 81]) = [-3, -2] \cup [2, 3]$ das Urbild des Intervalls $[16, 81]$ unter f .

Definition 1.31: injektiv, surjektiv, bijektiv

Seien X, Y zwei beliebige, nichtleere Mengen und $f : X \rightarrow Y$ eine Abbildung. Dann heißt f :

- **injektiv:** falls für alle $x, \tilde{x} \in X$ gilt:

$$f(x) = f(\tilde{x}) \implies x = \tilde{x}$$

- **surjektiv:** falls für jedes $y \in Y$ gilt:

$$\exists x \in X : f(x) = y$$

- **bijektiv:** falls f injektiv und surjektiv ist

Man kann sich anhand der Definition leicht überlegen, dass eine Abbildung $f : X \rightarrow Y$ genau dann bijektiv ist, wenn es für jedes $y \in Y$ genau ein $x \in X$ gibt, sodass $f(x) = y$ gilt.

Beispiel 1.32: Betrachte die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \max(0, x)$

- $f : \mathbb{R} \rightarrow \mathbb{R}$, f ist weder injektiv noch surjektiv
- $f : \mathbb{R} \rightarrow \mathbb{R}_+$, f ist surjektiv, aber nicht injektiv
- $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, f ist injektiv aber nicht surjektiv
- $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, f ist bijektiv

Definition 1.33: Komposition

Seien X, Y, Z nichtleere Mengen und die Abbildungen $f : X \rightarrow Y$, $x \mapsto f(x)$ sowie $g : Y \rightarrow Z$, $y \mapsto g(y)$ gegeben. Dann ist die **Komposition** oder **Hintereinanderausführung** von f und g die Abbildung

$$g \circ f : X \rightarrow Z, \quad x \mapsto g(f(x)) \in Z$$

Satz 1.34: Seien W, X, Y und Z nichtleere Mengen, und die Abbildungen $f : W \rightarrow X$, $g : X \rightarrow Y$, $h : Y \rightarrow Z$ gegeben. Dann gilt:

1. $h \circ (g \circ f) = (h \circ g) \circ f$, d.h. die Komposition von Abbildungen ist Assoziativ
2. Sind beide Abbildungen f und g injektiv/surjektiv/bijektiv, dann ist auch die Komposition $g \circ f$ injektiv / surjektiv / bijektiv.
3. Ist $g \circ f$ injektiv, dann ist f injektiv
4. Ist $g \circ f$ surjektiv, dann ist g surjektiv

Beweis: (Übungsaufgabe, Blatt 4 Aufgabe 1)

1. $h \circ (g \circ f)(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x)$
2. Für jedes $x_1, x_2 \in X$ folgt aus g injektiv: $g(f(x_1)) = g(f(x_2)) \implies f(x_1) = f(x_2)$. Aus f injektiv folgt wiederum: $f(x_1) = f(x_2) \implies x_1 = x_2$. Also gilt $g(f(x_1)) = g(f(x_2)) \implies x_1 = x_2$. Somit ist $g \circ f$ injektiv.

Für jedes $z \in Z$ folgt aus g surjektiv: $\exists y \in Y : f(y) = z$. Für jedes $y \in Y$ folgt aus f surjektiv wiederum: $\exists x \in X : f(x) = y$. Also folgt $\forall z \in Z \exists x \in X : g(f(x)) = z$. Somit ist $g \circ f$ surjektiv.

Sind f und g bijektiv, folgt aus den obigen Beweisen, dass $g \circ f$ injektiv und surjektiv ist. Somit ist $g \circ f$ auch bijektiv.

3. Ist f nicht injektiv, dann existieren $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$ aber $x_1 \neq x_2$. Wegen $g \circ f$ injektiv gilt $g(f(x_1)) = g(f(x_2)) \implies x_1 = x_2$. Damit $g(f(x_1)) = g(f(x_2))$ gilt, muss auch $f(x_1) = f(x_2)$ gelten, dann gilt aber auch $f(x_1) = f(x_2) \implies x_1 = x_2$. Dies ist ein Widerspruch \nmid . f ist also injektiv.
4. Ist g nicht surjektiv, dann existiert ein $z \in Z$ für das kein $y \in Y$ mit $g(y) = z$ existiert. Wegen $g \circ f$ surjektiv gilt $\forall z \in Z \exists x \in X : g(f(x)) = z$. Dann gilt auch $g(f(x)) = g(y) = z$, $y \in Y$, also existiert ein $y \in Y$ mit $g(y) = z$. Dies ist ein Widerspruch \nmid . Also ist g surjektiv.

□

Satz 1.35: Seien X, Y nichtleere Mengen und $f : X \rightarrow Y$ eine Abbildung. Die Abbildung ist genau dann bijektiv, wenn es eine Abbildung $g : Y \rightarrow X$ existiert, so dass $g \circ f = \text{Id}_X$ und $f \circ g = \text{Id}_Y$ gilt.

Beweis:

“ \implies ”

Zu jedem $y \in Y$ existiert genau ein $x_y \in X$ mit $f(x_y) = y$. Damit kann man eine Abbildung g definieren durch:

$$g : Y \rightarrow X, \quad g(y) = x_y$$

Für $y \in Y$ folgt dann $(f \circ g)(y) = f(g(y)) = f(x_y) = y \implies f \circ g = \text{Id}_Y$. Sei $x \in X \implies f(x) = y \in Y$. Wegen der Bijektivität von f folgt $x = x_y \in X$

Dann gilt:

$$(g \circ f)(x) = g(f(x)) = g(y) = x_y = x \implies g \circ f = \text{Id}_X$$

“ \Leftarrow ”:

Es gilt: $g \circ f = \text{Id}_X$, Id_X ist injektiv. Wegen Satz 1.34, 3) ist dann auch f injektiv. Des weiteren gilt $f \circ g = \text{Id}_Y$ ist surjektiv. Wegen 1.34, 4) ist dann auch f surjektiv $\implies f$ ist bijektiv

Frage: Gibt es eine weitere Abbildung, $\tilde{g} : Y \rightarrow X$ mit den gleichen Eigenschaften wie im letzten Satz? Wegen Satz 1.34, 1) gilt:

$$\tilde{g} = \text{Id}_X \circ \tilde{g} = (g \circ f) \circ \tilde{g} = g \circ (f \circ \tilde{g}) = g \circ \text{Id}_Y = g$$

Definition 1.36: inverse Abbildung / Umkehrabbildung

Seien X, Y zwei nichtleere Mengen und $f : X \rightarrow Y$ eine Abbildung. Ist f bijektiv, dann heißt die in Satz 1.35 definierte, eindeutige Abbildung $g : Y \rightarrow X$ **inverse Abbildung** oder **Umkehrabbildung** von f und wird f^{-1} bezeichnet.

Beispiel 1.37: Die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 3x - 5$ ist bijektiv. Die zu f inverse Abbildung erhält man durch Umformung.

$$y = 3x - 5 \iff y + 5 = 3x \iff x = \frac{1}{3}(y + 5)$$

Also $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$, $y \mapsto \frac{1}{3}(y + 5)$

Achtung: $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ ist nicht bijektiv.

$$\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}_+, \quad \tilde{f}(x) = x^2 \qquad \tilde{f}^{-1}(y) = \sqrt{y}$$

Achtung: Die Notation f^{-1} ist doppelt Belegt! Zum einen für die Notation der Umkehrabbildung und zum Anderen für die Notation des Urbilds.

Satz 1.38: Seien X, Y und Z nichtleere Mengen und die Abbildungen $f : X \rightarrow Y$ sowie $g : Y \rightarrow Z$ bijektiv.

Dann gilt:

1. f^{-1} ist bijektiv $(f^{-1})^{-1} = f$
2. $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

Beweis: (1. Übungsaufgabe)

1.

$$\begin{aligned}
 f &= f \circ \text{Id}_X \\
 &= f \circ (f^{-1} \circ (f^{-1})^{-1}) \\
 &= (f \circ f^{-1}) \circ (f^{-1})^{-1} \\
 &= \text{Id}_Y \circ (f^{-1})^{-1} \\
 &= (f^{-1})^{-1}
 \end{aligned}$$

2. Aus Satz 1.34 folgt, dass $g \circ f$ bijektiv ist. $\implies (g \circ f)^{-1}$ existiert und ist eindeutig bestimmt.
Es gilt

$$\begin{aligned}
 (f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ ((g^{-1} \circ g) \circ f) \\
 &= f^{-1} \circ (\text{Id}_Y \circ f) \\
 &= f^{-1} \circ f = \text{Id}_X
 \end{aligned}$$

Analog zeigt man: $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \text{Id}_Y$

$$\implies (g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

□

2. Algebraische Strukturen

Algebraische Strukturen sind Mengen und Verknüpfungen, die auf den Elementen der Menge definiert sind. Ein Beispiel dafür ist die Menge aller ganzen Zahlen mit der Addition als Verknüpfung.

Algebraische Strukturen besitzen wichtige Eigenschaften:

- Die Summe zweier ganzer Zahlen ist wieder eine ganze Zahl $\hat{=}$ Abgeschlossenheit der Menge bezüglich der Verknüpfung
- Es gibt die ganze Zahl 0, sodass für jede ganze Zahl $a \in \mathbb{Z}$ gilt: $0 + a = a$. Dieses Element nennt man das neutrale Element
- Für jede ganze Zahl $a \in \mathbb{Z}$ gibt es ein $-a \in \mathbb{Z}$, sodass gilt: $(-a) + a = 0$. Dieses Element nennt man das inverse Element von a

Algebraische Strukturen erlauben es uns, abstrakte Konzepte aus konkreten Beispielen zu extrahieren und später komplexe Zusammenhänge mit diesen Konzepten zu analysieren und Stück für Stück zu erweitern.

2.1. Gruppen

Definition 2.1: innere Verknüpfung, Halbgruppe

Sei M eine nichtleere Menge. Eine Abbildung $\circ : M \times M \rightarrow M$, $(a, b) \mapsto a \circ b$ heißt (**innere**) **Verknüpfung** auf M . Gilt: $(a \circ b) \circ c = a \circ (b \circ c)$, dann heißt die Verknüpfung **assoziativ** und (M, \circ) eine **Halbgruppe**. Gilt für eine Halbgruppe, dass $a \circ b = b \circ a$, so heißt die Halbgruppe **abelsch** oder **kommutativ**.

Je nach Kontext kann die Notation einer Verknüpfung variieren. ($a \circ b$, $a \cdot b$, ab)

Beispiel 2.2:

- $(\mathbb{N}, +)$ und $(\mathbb{N}, *)$ sind kommutative Halbgruppen
- Sei X eine nichtleere Menge. Dann ist $M := \text{Abb}(X, X) = \{\text{Abbildungen } f : X \rightarrow X\}$ eine Halbgruppe mit der Verknüpfung \circ als Komposition von Abbildungen (Def. 1.33). Diese Halbgruppe ist nicht abelsch.

Beweis durch Gegenbeispiel:

Sei $a, b, c \in X$, $a \neq b$, $a \neq c$, $b \neq c$. Definiere $f, g \in M$ mit

$$f(x) := \begin{cases} b & \text{für } x = a \\ a & \text{für } x = b \\ x & \text{sonst} \end{cases} \quad g(x) := \begin{cases} c & \text{für } x = a \\ a & \text{für } x = c \\ x & \text{sonst} \end{cases}$$

Dann folgt:

$$\begin{aligned}(f \circ g)(a) &= f(g(a)) = f(c) = c \\(g \circ f)(a) &= g(f(a)) = g(b) = b \\&\quad \checkmark\end{aligned}$$

Die Halbgruppe ist ein relativ “schwaches” Konzept. Deswegen braucht man weitere Eigenschaften:

Definition 2.3: neutrales Element

Sei M eine nichtleere Menge und \circ eine innere Verknüpfung auf M . Existiert ein Element $e \in M$ mit

$$a \circ e = e \circ a = a \quad \forall a \in M$$

so heißt e **neutrales Element** für die Verknüpfung \circ .

Eine Halbgruppe, die ein neutrales Element besitzt heißt **Monoid**.

Beispiel 2.4: Kein Monoid

Gegeben sei die Menge $M = \{a, b\}$ und die folgende Verknüpfung

\circ	a	b
a	a	b
b	a	b

Man kann nachrechnen, dass (M, \circ) eine Halbgruppe ist. Man kann auch prüfen, dass a linksneutral aber nicht rechtsneutral ist, sowie dass b rechtsneutral aber nicht linksneutral ist. Somit besitzt die Halbgruppe kein neutrales Element, (M, \circ) ist also kein Monoid.

Bemerkung: In der Definition eines Monoids wird nur die Existenz aber nicht die Eindeutigkeit des neutralen Elements gefordert. Ist dies sinnvoll?

Lemma 2.5: Sei (M, \circ) ein Monoid und $e_1, e_2 \in M$ neutrale Elemente, dann gilt

$$e_1 = e_2$$

Beweis:

$$\begin{aligned}e_1 &= e_1 \circ e_2 \\&= e_2\end{aligned}$$

□

Beispiel 2.6:

- $(\mathbb{N}, +)$ ist kein Monoid, da kein neutrales Element existiert ($0 \notin \mathbb{N}$ in LinA)
- (\mathbb{N}, \cdot) ist ein Monoid mit dem neutralen Element $e = 1$
- Für $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ ist $(\mathbb{N}_0, +)$ ein Monoid mit dem neutralen Element $e = 0$

Definition 2.7: Gruppen

Ein Monoid (M, \circ) ist eine **Gruppe**, wenn für jedes $a \in M$ ein $b \in M$ existiert, so dass

$$a \circ b = b \circ a = e$$

wobei e das neutrale Element des Monoids ist. Wir nennen b das **inverse Element** zu dem gegebenen Element a und bezeichnen es mit $a^{-1} = b$.

Bemerkung: Für $\circ = +$, d.h. additiv geschriebene Gruppen schreibt man auch $-a := b$.

Beispiel 2.8:

- $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ sind kommutative Gruppen
- $(\mathbb{N}, +)$ ist keine Gruppe, da kein neutrales Element und keine inversen Elemente existieren
- Rechnen mit binären Zahlen

Betrachtet wird $\mathbb{F}_2 = \{0, 1\}$ und die Verknüpfungen

$+$	0	1
0	0	1
1	1	0

\cdot	0	1
0	0	0
1	0	1

Anhand der Verknüpfungstabellen erkennt man, dass $(\mathbb{F}_2, +)$ mit dem neutralen Element $e = 0$ eine abelsche Gruppe ist. Jedoch ist (\mathbb{F}_2, \cdot) keine Gruppe, da zwar ein neutrales Element $e = 1$, aber das Element 0 kein inverses Element besitzt.

Satz 2.9: Sei (M, \circ) eine Gruppe, dann gilt:

1. Es gibt **genau ein** neutrales Element in M .
2. Jedes Element der Menge M besitzt **genau ein** inverses Element.
3. Jedes linksinverse Element ist gleichzeitig auch rechtsinvers.
4. Jedes linksneutrale Element ist gleichzeitig auch rechtsneutral.

Beweis:

1. Folgt aus Lemma 2.5, da (M, \circ) nach Definition ein Monoid ist.
2. Annahme: Seien $b \in M$ und $\tilde{b} \in M$ inverse Elemente zu $a \in M$.

zu zeigen: $b = \tilde{b}$

Dann gilt:

$$\begin{aligned}
 b &= b \circ e \\
 &= b \circ (a \circ \tilde{b}) \\
 &= (b \circ a) \circ \tilde{b} \\
 &= e \circ \tilde{b} \\
 &= \tilde{b}
 \end{aligned}$$

3. Es sei $b \in M$ ein linksinverses Element zu $a \in M$. D.h. $b \circ a = e$. Sei $\tilde{b} \in M$ ein linksinverses Element zu $b \in M$. D.h. $\tilde{b} \circ b = e$.

$$\begin{aligned} a \circ b &= e \circ (a \circ b) \\ &= \tilde{b} \circ b \circ (a \circ b) \\ &= \tilde{b} \circ (b \circ a) \circ b \\ &= \tilde{b} \circ e \circ b \\ &= \tilde{b} \circ b \\ &= e \end{aligned}$$

4. Es gelte $e \circ a = a$ und $b \circ a = a \circ b = e$

$$\text{Dann gilt: } a \circ e = a \circ (b \circ a) = (a \circ b) \circ a = a \quad \checkmark$$

Lemma 2.10: Sei (M, \circ) eine Gruppe. Gilt für ein $a \in M$, dass $c \circ a = a$ für ein $c \in M$; dann ist c das neutrale Element der Gruppe.

Beweis: Sei e das neutrale Element (es gibt genau 1) der Gruppe (M, \circ) und für $a, c \in M$ gelte: $c \circ a = a$. Sei b das inverse Element zu a .

$$\begin{aligned} c &= c \circ e \\ &= c \circ (a \circ b) \\ &= (c \circ a) \circ b \\ &= a \circ b \\ &= e \end{aligned}$$

□

Besonders wichtig in der linearen Algebra sind Abbildungen zwischen Gruppen, die bezüglich der Verknüpfung "kompatibel" sind.

Definition 2.11: Homomorphismus

Seien (M, \circ) und (N, \oplus) Gruppen. Eine Abbildung $f : M \rightarrow N$ heißt **Homomorphismus** (oder **Gruppenhomomorphismus**) falls:

$$f(x \circ y) = f(x) \oplus f(y) \quad \forall x, y \in M$$

Ein Homomorphismus heißt **Isomorphismus**, wenn er bijektiv ist.

Beispiel 2.12: Die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}_{>0}$ mit $f(x) = e^{2x}$ ist ein Homomorphismus zwischen $(\mathbb{R}, +)$ und $(\mathbb{R}_{>0}, \cdot)$ mit $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$ denn

$$f(x + y) = e^{2(x+y)} = e^{2x} \cdot e^{2y} = f(x) \cdot f(y)$$

Satz 2.13: Sei $f : M \rightarrow N$ für die Gruppen (M, \circ) und (N, \oplus) ein Homomorphismus sowie e_M und e_N jeweils die neutralen Elemente. Dann gilt $f(e_M) = e_N$.

Beweis: Sei $a \in M$ beliebig gewählt, dann folgt

$$f(a) = f(e_M \circ a) = f(e_M) \oplus f(a) = e_N \oplus f(a)$$

$$\stackrel{2.10}{\implies} f(e_M) \text{ ist ein neutrales Element} \stackrel{2.9}{\implies} f(e_M) = e_N$$

□

Homomorphismen bilden das neutrale Element im Definitionsbereich immer auf das neutrale Element des Wertebereichs ab. Später werden wir sehen, dass wenn der Homomorphismus nicht bijektiv ist, noch mehr Elemente auf das neutrale Element im Wertebereich e_N abgebildet werden können. Dies motiviert folgende Definition:

Definition 2.14: Kern

Ist $f : (M, \circ) \rightarrow (N, \oplus)$ ein Homomorphismus, so nennt man:

$$\ker(f) := \{a \in M \mid f(a) = e_N\}$$

den **Kern** von f .

2.2. Ringe

Ringe sind eine Erweiterung der algebraischen Strukturen von einer auf zwei Verknüpfungen.

Definition 2.15: Ring

Seien R eine Menge und “+” sowie “·” zwei Verknüpfungen auf R . Das Tripel $(R, +, \cdot)$ heißt **Ring**, falls gilt:

1. $(R, +)$ ist eine kommutative Gruppe, deren neutrales Element wir mit 0 bezeichnen.
2. (R, \cdot) ist eine Halbgruppe, d.h. es gilt das Assoziativgesetz.
3. Es gelten die Distributivgesetze: $\forall a, b, c \in R$

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

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

Ein Ring heißt kommutativ, wenn \cdot kommutativ ist. D.h. $a \cdot b = b \cdot a$, $\forall a, b \in R$. Ein Element $1 \in R$ heißt **Einselement**, wenn es das neutrale Element bezüglich der Multiplikation ist. Das heißt wenn für alle $a \in R$ gilt: $1 \cdot a = a \cdot 1 = a$.

Achtung: Die Formulierung der Distributivitätsgesetze impliziert, dass die Multiplikation stärker bindet als die Addition (“Punkt vor Strich”).

Beispiel 2.16:

- $(\mathbb{N}, +, \cdot)$ ist kein Ring
- $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit Einselement

- $(\mathbb{F}_2, +, \cdot)$ ist ein kommutativer Ring mit Einselement, denn
 - das neutrale Element bezüglich der Addition ist die 0, denn $0 + 0 = 0$, $0 + 1 = 1$, $1 + 0 = 1$
 - das additive inverse Element zu 0 ist 0 und zu 1 die 1, denn $1 + 1 = 0$
 - die Addition ist kommutativ
 - die Addition ist assoziativ, zeigt man durch nachrechnen für alle 8 Möglichkeiten
 - das neutrale Element für die Multiplikation ist 1, denn $0 \cdot 1 = 0$ und $1 \cdot 1 = 1$
 - die Multiplikation ist kommutativ
 - die Multiplikation ist assoziativ, zeigt man durch nachrechnen
 - die Distributivgesetze gelten, zeigt man durch nachrechnen

In Ringen gelten die “üblichen” Rechenregeln, z.B.:

$$0 \cdot a = 0$$

Beweis:

$$\begin{aligned} 0 \cdot a &= 0 \cdot a + 0 \cdot a - 0 \cdot a = (0 + 0) \cdot a - 0 \cdot a \\ &= 0 \cdot a - 0 \cdot a = 0 \end{aligned}$$

oder auch

$$(-1) \cdot a = -a$$

Beweis:

$$\begin{aligned} a + (-1) \cdot a &= 0, \text{ denn} \\ a + (-1) \cdot a &= 1 \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0 \end{aligned}$$

analog zeigt man $a \cdot 0 = 0$ und $a \cdot (-1) = -a$

Bemerkung: Wenn in einem Ring die Gleichung $1 = 0$ gilt, folgt

$$a = a \cdot 1 = a \cdot 0 = 0$$

Somit muss R der Nullring sein, $R = \{0\}$.

Beispiel 2.17: Ring der Polynome

Sei $(R, +, \cdot)$ ein kommutativer Ring mit Eins. Ein Polynom mit Koeffizienten in R und der Unbekannten $t \in R$ (kurz Polynom über R) ist gegeben durch

$$p(t) = a_0 \cdot t^0 + a_1 \cdot t^1 + \dots + a_n \cdot t^n, \quad a_0, a_1, \dots, a_n \in R$$

Die Menge aller Polynome über R wird mit $P[t]$ bezeichnet.

Betrachte zwei Polynome $p, q \in P[t]$ mit

$$p(t) = a_0 + a_1 t + \dots + a_n t^n \text{ und } q(t) = b_0 + b_1 t + b_m t^m$$

mit $n \geq m$. Ist $n > m$, so setzen wir $b_j = 0$ für $j = m + 1, \dots, n$. $p(t)$ und $q(t)$ sind gleich, wenn $a_j = b_j$ für alle $j \in \{1, \dots, n\}$ gilt.

Aufgrund der Eigenschaften von R gilt

$$a_0 + a_1 t + \dots + a_{n-1} t^{n-1} + a_n t^n = a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$$

Der Grad eines Polynoms $p(t) \in P[t]$ ist definiert als der größte Index j für den $a_j \neq 0$ gilt. Gibt es keinen solchen Index, ist $p(t)$ das Nullpolynom, d.h. $p(t) = 0$ für alle $t \in R$ und man definiert den Grad von $p(t)$ als $-\infty$.

Sind zwei Polynome $p, q \in P[t]$ wie oben definiert, und setzen wir wieder $b_j = 0$ für alle $j \in \{m+1, \dots, n\}$, dann sind die Verknüpfungen “+” und “ \cdot ” wie folgt definiert:

$$p(t) + q(t) = (a_0 + b_0) + (a_1 + b_1)t + \dots + (a_n + b_n)t^n \text{ und}$$

$$p(t) \cdot q(t) = c_0 + c_1 t + \dots + c_{n+m} t^{n+m}, \quad c_k := \sum_{i+k=k} a_i \cdot b_j$$

Mit dem Nullpolynom definiert wie oben und dem Einspolynom definiert als $p(t) := 1$ kann man nachrechnen, dass $(P[t], +, \cdot)$ ein kommutativer Ring ist.

Definition 2.18: invertierbar

Es sei $(R, +, \cdot)$ ein Ring mit Eins und $a \in R$ gegeben. Ein Element $b \in R$ heißt **invers** (bezüglich \cdot) zu a , wenn gilt:

$$a \cdot b = b \cdot a = 1$$

Existiert zu $a \in R$ ein inverses Element, so heißt a **invertierbar**.

Satz 2.19: Es sei $(R, +, \cdot)$ ein Ring mit Eins. Dann gilt:

1. Existiert zu $a \in R$ ein inverses Element bezüglich \cdot , so ist dies eindeutig bestimmt. Dies wird mit a^{-1} gekennzeichnet.
2. Wenn $a, b \in R$ invertierbar sind, dann ist auch $a \cdot b$ invertierbar und es gilt:

$$(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$$

Beweis: Siehe oben im Abschnitt zu Abbildungen.

□

2.3. Körper

Eine knappe Definition eines Körpers:

Ein kommutativer Ring mit Eins heißt Körper, falls $0 \neq 1$ gilt (der Nullring wird ausgeschlossen) und jedes Element $a \in R \setminus \{0\}$ invertierbar ist.

Es folgt eine äquivalente und formale Definition:

Definition 2.20: Körper

Eine Menge K mit zwei Verknüpfungen

$$+ : K \times K \rightarrow K, (a, b) \mapsto a + b \quad \text{Addition}$$

$$\cdot : K \times K \rightarrow K, (a, b) \mapsto a \cdot b \quad \text{Multiplikation}$$

heißt **Körper**, wenn gilt:

- $(K, +)$ ist eine kommutative Gruppe
- $(K \setminus \{0\}, \cdot)$ ist auch eine kommutative Gruppe
- Es gelten die Distributivgesetze

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

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

Lemma 2.21: Sei $(K, +, \cdot)$ ein Körper. Gilt für $a, b \in K$, dass $a \cdot b = 0$, so ist mindestens eins davon die 0.

Beweis:

Fall 1: $a = b = 0$

Fall 2: o.B.d.A: $a \neq 0 \implies \exists a^{-1} : a \cdot a^{-1} = 1$

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

□

Diese Eigenschaft nennt man Nullteilerfreiheit.

Beispiel 2.22:

- $(\mathbb{R}, +, \cdot)$ ist ein Körper
- $(\mathbb{Z}, +, \cdot)$ ist kein Körper, da die multiplikativ inversen Elemente in \mathbb{Q} , aber nicht immer in \mathbb{Z} liegen

Beispiel 2.23: komplexe Zahlen

Die Menge der komplexen Zahlen ist definiert als:

$$\mathbb{C} := \{(x, y) \mid x, y \in \mathbb{R}\}$$

d.h. $\mathbb{C} = \mathbb{R} \times \mathbb{R}$. Die zwei Verknüpfungen Addition und Multiplikation werden wie folgt definiert:

$$+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, (a, b) + (c, d) = (a + c, b + d)$$

$$\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, (a, b) \cdot (c, d) = (a \cdot c - b \cdot d, a \cdot d + b \cdot c)$$

Wir verwenden implizit die Operationen auf den reellen Zahlen, $+$, $-$, \cdot . Dann sieht man:

- Das neutrale Element in \mathbb{C} bezüglich $+$ ist die $0_{\mathbb{C}} = (0, 0)$
- Das neutrale Element in \mathbb{C} bezüglich \cdot ist die $1_{\mathbb{C}} = (1, 0)$

Man rechnet nach, dass

- Das inverse Element bezüglich $+$ in \mathbb{C} definiert ist mit

$$-(x, y) = (-x, -y) \in \mathbb{C} \quad \forall (x, y) \in \mathbb{C}$$

- Das inverse Element bezüglich \cdot in \mathbb{C} definiert ist mit

$$(x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2} \right) \in \mathbb{C} \quad \forall (x, y) \in \mathbb{C} \setminus \{0_{\mathbb{C}}\}$$

Das Überprüfen der Rechengesetze zeigt, dass \mathbb{C} ein Körper ist.

Für die Teilmenge

$$M := \{(x, 0) \mid x \in \mathbb{R}\} \subset \mathbb{C}$$

kann man jedes Element der reellen Zahlen mit einem Element der Menge M mit der bijektiven Abbildung

$$\mathbb{R} \rightarrow M, x \mapsto (x, 0)$$

identifizieren. Mit $0_{\mathbb{R}} \mapsto (0, 0) = 0_{\mathbb{C}}$, $1_{\mathbb{R}} \mapsto (1, 0) = 1_{\mathbb{C}}$ kann man M als Teilkörper von \mathbb{C} auffassen. Es gilt jedoch auch $\mathbb{R} \not\subseteq \mathbb{C}$ (zumindest in LinA).

Eine besondere komplexe Zahl ist die imaginäre Einheit $(0, 1)$, für die gilt:

$$(0, 1) \cdot (0, 1) = (-1, 0) \hat{=} -1$$

Dabei wird $(-1, 0) \in \mathbb{C}$ mit $-1 \in \mathbb{R}$ über die oben genannte bijektive Abbildung identifiziert. Mit der Definition $i := (0, 1)$ folgt

$$i \cdot i = -1$$

Mit dieser Notation und Identifikation kann man eine komplexe Zahl $z \in \mathbb{C}$ beschreiben mit

$$\begin{aligned} z = (x, y) &= (x, 0) + (0, y) = (x, 0) + (0, 1) \cdot (y, 0) \\ &= x + iy \end{aligned}$$

Man schreibt $\operatorname{Re}(z) = x$ als **Realanteil** von z und $\operatorname{Im}(z) = y$ als **Imaginäranteil** von z .

Man definiert zu $(x, y) \in \mathbb{C}$ die **konjugiert komplexe Zahl** durch

$$\bar{z} = (x, -y) \in \mathbb{C}$$

Damit erhält man für ein $z = (x, y) \in \mathbb{C}$:

$$\begin{aligned} |z| &:= \sqrt{z \cdot \bar{z}} = \sqrt{(x + iy) \cdot (x - iy)} \\ &= \sqrt{x^2 - ixy + ixy - i^2 y^2} \\ &= \sqrt{x^2 + y^2} \end{aligned}$$

2.4. Vektorräume

Beispiel 2.24: Kräfteparallelogramm

Betrachten wir einige Gesetze aus der Mechanik:

Je zwei am selben Punkt angreifende Kräfte können durch eine einzige Kraft ersetzt werden. Diese resultierende Kraft (= Gesamtkraft) hat die gleiche Wirkung wie die Einzelkräfte.

$$F = F_1 + F_2 \quad \text{die Kräfte können als Vektoren betrachtet werden}$$

- Ein Vektor hat eine Länge und eine Richtung
- Vektoren kann man addieren
- Vektoren können mit einer reellen Zahl multipliziert werden

Beispiel 2.25: Interpolationsproblem

Gegeben sind reelle Zahlen $a, b, c \in \mathbb{R}$. Gesucht ist ein Polynom zweiten Grades $p(t) \in P[t]$ mit

$$p(1) = a \quad p(2) = b \quad p(3) = c$$

für ein $p(t) = a_0 + a_1 t + a_2 t^2$. D.h. es muss gelten:

$$p(1) = a_0 + a_1 \cdot 1 + a_2 \cdot 1 = a$$

$$p(2) = a_0 + a_1 \cdot 2 + a_2 \cdot 4 = b$$

$$p(3) = a_0 + a_1 \cdot 3 + a_2 \cdot 9 = c$$

Diese Gleichung hat genau eine Lösung.

$$p(t) = (3a - 3b + c) + \left(-5\frac{a}{2} + 4b - 3\frac{c}{2}\right)t + \left(\frac{a}{2} - b + \frac{c}{2}\right)t^2$$

Eine alternative Darstellung ist

$$p_1(t) = \frac{1}{2}(t-2)(t-3) \quad p_2(t) = -(t-1)(t-3) \quad p_3(t) = \frac{1}{2}(t-1)(t-2)$$

für die gilt:

$$p_{i(k)} = \begin{cases} 1 & \text{für } i = k \\ 0 & \text{sonst} \end{cases}$$

Dann ist $p(t)$ gegeben durch:

$$p(t) = ap_1(t) + bp_2(t) + cp_3(t)$$

Beobachtung: Die additive Verknüpfung zweier Elemente gleicher Art und Multiplikation mit einer reellen Zahl ($\hat{=}$ Skalar).

Solch eine algebraische Struktur wollen wir beschreiben:

Definition 2.26: Vektorraum

Sei K ein Körper. Ein Vektorraum über K , kurz K -Vektorraum, ist eine Menge V mit zwei Abbildungen:

- Addition

$$+ : V \times V \rightarrow V, (v, w) \mapsto v + w$$

- skalare Multiplikation

$$\cdot : K \times V \rightarrow V, (\lambda, v) \mapsto \lambda v$$

für die folgendes gilt:

- $(V, +)$ ist eine kommutative Gruppe
- Für alle $v, w \in V$ und $\lambda, \mu \in K$ gilt:

1. $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$
2. $1 \cdot v = v$
3. $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$
4. $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$

Ein Element $v \in V$ nennen wir **Vektor**, ein $\mu \in K$ nennen wir einen **Skalar**.

Beobachtung: Für einen Vektorraum sind die Operationen $+$ und die skalare Multiplikation \cdot abgeschlossen.

Beispiel 2.27: Für einen Körper K ist der Standardvektorraum gegeben durch die Menge $V = K^n$ für ein $n \in \mathbb{N}$. Die n -Tupel werden geschrieben als

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ mit } v_1, v_2, \dots, v_n \in K$$

Die Addition und die skalare Multiplikation ist komponentenweise definiert.

$$v + w = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

$$\lambda \cdot w = \lambda \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot w_1 \\ \lambda \cdot w_2 \\ \vdots \\ \lambda \cdot w_n \end{pmatrix}$$

Damit ist der Vektorraum $V = K^n$ ein K -Vektorraum. Der Nullvektor \vec{v}_0 ist definiert durch

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Das additiv inverse Element ist gegeben durch

$$-v = -\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix} \text{ für } v_1, v_2, \dots, v_n \in K$$

Da K ein Körper ist, ist die so definierte skalare Multiplikation assoziativ, distributiv und mit $1 \in K$ kompatibel ($1 \cdot v = v$).

Beispiel 2.28: Polynome

Die Menge $P[t]$ aller Polynome über einen Körper K mit der Unbekannten t bilden einen K -Vektorraum, wenn die Addition von Polynomen wie in Beispiel 2.17 definiert ist und die skalare Multiplikation für ein $p(t) = a_0 + a_1 t + \dots + a_n t^n \in P[t]$ definiert ist durch:

$$\cdot : K \times P[t] \rightarrow P[t]$$

$$\lambda \cdot p(t) = (\lambda a_0) + (\lambda a_1)t + \dots + (\lambda a_n)t^n$$

Beispiel 2.29: Abbildungen

Die Menge $V = \text{Abb}(\mathbb{R}, \mathbb{R})$ der Abbildungen $f : \mathbb{R} \rightarrow \mathbb{R}$ bilden einen Vektorraum über den Körper \mathbb{R} mit den Verknüpfungen

$$+ : V \times V \rightarrow V, (f, g) \mapsto f + g \quad (f + g)(x) := f(x) + g(x)$$

und

$$\cdot : \mathbb{R} \times V \rightarrow V, (\lambda, g) \mapsto \lambda \cdot g$$

Das Gleiche gilt für

$$V := \{\text{stetige Abbildungen } f : \mathbb{R} \rightarrow \mathbb{R}\}$$

$$V := \{\text{differenzierbare Abbildungen } f : \mathbb{R} \rightarrow \mathbb{R}\}$$

Lemma 2.30: Für den K -Vektorraum $(V, +, \cdot)$ mit dem Nullelement 0_K des Körpers und 0_V des Vektorraums. Dann gilt

1. $0_K \cdot v = 0_V$
2. $\lambda \cdot 0_V = 0_V$
3. $-(\lambda \cdot v) = (-\lambda) \cdot v = \lambda \cdot (-v) \quad \forall \lambda \in K, \forall v \in V$

Beweis:

zu 1) $\forall v \in V$ gilt

$$\begin{aligned} 0_K \cdot v &= (0_K + 0_K) \cdot v = 0_K \cdot v + 0_K \cdot v & | - (0_K \cdot v) \\ 0_V &= 0_K \cdot v + 0_V = 0_K \cdot v \\ 0_V &= 0_K \cdot v \end{aligned}$$

zu 2) $\forall \lambda \in K$ gilt

$$\begin{aligned} \lambda \cdot 0_V &= \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V & | - (\lambda \cdot 0_V) \\ 0_V &= \lambda \cdot 0_V \end{aligned}$$

zu 3) $\forall \lambda \in K, \forall v \in V$ gilt

$$\begin{aligned} \lambda \cdot v + ((-\lambda) \cdot v) &= (\lambda - \lambda) \cdot v = 0_K \cdot v = 0_V \quad \checkmark \\ \lambda \cdot v + (\lambda \cdot (-v)) &= (\lambda \cdot (v - v)) = \lambda \cdot 0_V = 0_V \quad \checkmark \end{aligned}$$

□

Definition 2.31: Untervektorraum

Sei $(V, +, \cdot)$ ein K -Vektorraum und sei $U \subseteq V$. Ist $(U, +, \cdot)$ ein K -Vektorraum, so heißt $(U, +, \cdot)$ ein **Untervektorraum**, kurz **Unterraum** von $(V, +, \cdot)$.
Wichtig! Die Abgeschlossenheit muss erhalten bleiben.

Lemma 2.32: Sei $(V, +, \cdot)$ ein K -Vektorraum und $U \subseteq V$. Dann ist $(U, +, \cdot)$ genau dann ein Unterraum von V , wenn gilt:

1. $u + w \in U \quad \forall u, w \in U$
2. $\lambda u \in U \quad \forall \lambda \in K, \forall u \in U$

Beweis: (Übung)

Ist U nicht abgeschlossen bezüglich der Addition und der skalaren Multiplikation, dann ist $(U, +)$ keine kommutative Gruppe und (U, \cdot) keine Halbgruppe. In beiden Fällen ist $(U, +, \cdot)$ dann kein Vektorraum und somit auch kein Untervektorraum von $(V, +, \cdot)$.

□

Beispiel 2.33:

- Jeder Vektorraum $(V, +, \cdot)$ hat die Vektorräume $(U = V, +, \cdot)$ und $(U = \{0_V\}, +, \cdot)$
- Für jedes $n \in \mathbb{N}_0$ ist die Menge aller Polynome mit dem Grad kleiner gleich n , d.h. die Menge $P[t]_{\leq n} = \{p(t) \in P[t] \mid \text{Grad}(p) \leq n\}$ mit den Verknüpfungen aus Beispiel 2.28 ein Unterraum von $(P[t], +, \cdot)$

Definition 2.34: Linearkombination

Seien $(V, +, \cdot)$ ein K -Vektorraum, $n \in \mathbb{N}$ und $v_1, \dots, v_n \in V$. Ein Vektor der Form

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \sum_{i=1}^n \lambda_i v_i = v \in V$$

heißt **Linearkombination** von $v_1, \dots, v_n \in V$ mit den Koeffizienten $\lambda_1, \dots, \lambda_n \in K$. Die **lineare Hülle** / Der **Spann** von $v_1, \dots, v_n \in V$ ist die Menge

$$\text{Span}\{v_1, \dots, v_n\} := \left\{ \sum_{i=1}^n \lambda_i v_i \mid \lambda_1, \dots, \lambda_n \in K \right\}$$

Lemma 2.35: Sei $(V, +, \cdot)$ ein K -Vektorraum und $v_1, \dots, v_n \in V$, dann ist $(\text{Span}\{v_1, \dots, v_n\}, +, \cdot)$ ein Unterraum von $(V, +, \cdot)$.

Beweis: Es gilt $\text{Span}\{v_1, \dots, v_n\} \subseteq V$. Wegen Lemma 2.32 reicht es zu zeigen, dass $U := \text{Span}\{v_1, \dots, v_n\}$ bezüglich $+$ und \cdot abgeschlossen ist. Dies gilt nach der Definition der linearen Hülle. □

Beispiel 2.36: Für $V = \mathbb{R}^3$, $K = \mathbb{R}$. Betrachte

$$M = \left\{ \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 9 \\ 0 \\ 3 \end{pmatrix} \right\}, \text{ ist } \begin{pmatrix} 12 \\ 0 \\ 4 \end{pmatrix} \in \text{Span}(M)?$$

Ja, denn es gilt:

$$1 \cdot v_1 + 1 \cdot v_2 = \begin{pmatrix} 3+9 \\ 0+0 \\ 1+3 \end{pmatrix} = \begin{pmatrix} 12 \\ 0 \\ 4 \end{pmatrix}$$

Des weiteren gilt:

$$\text{Span}(M) = \left\{ \lambda \cdot \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} = A$$

Beweis:

“ \subseteq ”:

$$A = \left\{ \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \cdot \lambda \mid \lambda \in \mathbb{R} \right\} \subseteq \left\{ \lambda \cdot \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 9 \\ 0 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \subseteq \text{Span}(M)$$

“ \supseteq ”:

$$\begin{aligned}x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \text{Span}(M) &\iff \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 9 \\ 0 \\ 3 \end{pmatrix} \\&= \lambda_1 \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + 3\lambda_2 \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \\&= (\lambda_1 + 3\lambda_2) \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \\&\implies x = (\lambda_1 + 3\lambda_2) \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \in A\end{aligned}$$

□

3. Basen und Dimensionen von Vektorräumen

Dieses Kapitel motiviert unter anderem die Frage, wie man Vektorräume effizient beschreiben kann.

3.1. Lineare Unabhängigkeit

Definition 3.1: lineare Unabhängigkeit

Sei V ein K -Vektorraum. Die Vektoren $v_1, \dots, v_n \in V$ heißen **linear unabhängig**, wenn aus

$$\sum_{i=1}^n \lambda_i v_i = 0 \quad \text{mit } \lambda_1, \dots, \lambda_n \in K$$

folgt, dass $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ gilt. Folgt dies nicht, d.h. gilt

$$\sum_{i=1}^n \lambda_i v_i = 0 \quad \text{mit } \lambda_1, \dots, \lambda_n \in K$$

die nicht alle gleich 0 sind, so heißen v_1, \dots, v_n **linear abhängig**.

- Die leere Menge ist linear unabhängig.
- Ist $M \neq \emptyset$ eine Menge und für jedes $m \in M$ ein Vektor $v_m \in V$ gegeben, so nennt man die Menge $\{v_m\}_{m \in M}$ linear unabhängig, wenn endlich viele Vektoren immer linear unabhängig sind. Gilt dies nicht, so ist die Menge $\{v_m\}_{m \in M}$ linear abhängig.

Bemerkung: Nach Definition sind die Vektoren v_1, \dots, v_n genau dann linear unabhängig, wenn sich der Nullvektor aus ihnen nur in der Form $0 = 0 \cdot v_1 + \dots + 0 \cdot v_n$ mit endlich vielen Vektoren darstellen lässt

Beispiel 3.2: Fortsetzung von Beispiel 2.36

Die Vektoren aus $M = \{(3, 0, 1), (9, 0, 3)\}$ sind linear abhängig, da

$$3 \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 9 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{mit } \lambda_1 = 3, \lambda_2 = -1$$

Der Vektor $v_1 = (3, 0, 1)$ dagegen ist linear unabhängig, da

$$0 \cdot \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{mit } \lambda_1 = 0$$

Beispiel 3.3: $V = \mathbb{R}^3, K = \mathbb{R}$

Betrachte $M = \{(3, 0, 1), (9, 0, 3), (0, 1, 1)\}$. Sind diese Vektoren linear unabhängig? Die allgemeine Vorgehensweise ist hier, ein lineares Gleichungssystem aufzustellen und zu lösen. Hat das Gleichungssystem eine Lösung $\neq 0$, dann sind die Vektoren linear abhängig.

Hier: \Rightarrow linear abhängig

Lemma 3.4: Sei V ein K -Vektorraum. Dann gilt:

1. Ein einzelner Vektor ist genau dann linear unabhängig, wenn $v \neq 0_V$ gilt
2. Sind $v_1, \dots, v_n \in V$ linear unabhängig und ist $\{u_1, \dots, u_m\} \subseteq \{v_1, \dots, v_n\}$, dann ist auch die Menge u_1, \dots, u_m linear unabhängig
3. Sind $v_1, \dots, v_n \in V$ linear abhängig und $u_1, \dots, u_m \in V$. Dann ist auch $v_1, \dots, v_n, u_1, \dots, u_m$ linear abhängig

Beweis:

zu 1: Sei $v \in V, v \neq 0_V$. Es soll gelten: $\lambda v = 0 \Rightarrow \lambda = 0$, da $v \neq 0_V$

zu 2: v_1, \dots, v_n sind linear unabhängig und es gilt $\{u_1, \dots, u_m\} \subseteq \{v_1, \dots, v_n\}$.

Damit die Vektoren $\{u_1, \dots, u_m\}$ linear unabhängig sind, muss gelten

$$\lambda_1 u_1 + \dots + \lambda_m u_m = 0 \Rightarrow u_1 = \dots = u_m = 0$$

Wegen der linearen Unabhängigkeit von $\{v_1, \dots, v_n\}$ gilt nach Umbenennung der Vektoren $\{u_1, \dots, u_m\}$:

$$\lambda_{i_1} v_{i_1} + \dots + \lambda_{i_m} v_{i_m} + \sum_{\substack{j=1 \\ j \notin \{i_1, \dots, i_m\}}}^n v_j = 0 \Rightarrow \lambda_{i_1} = \dots = \lambda_{i_m} = 0$$

zu 3: v_1, \dots, v_n sind linear abhängig $\Rightarrow \exists \lambda_i \in K$, so dass nicht alle λ_i gleich Null sind und $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = 0$ gilt.

$\Rightarrow \lambda_1 v_1 + \dots + \lambda_n v_n + \mu_1 u_1 + \dots + \mu_m u_m = 0 \Rightarrow \mu_1, \dots, \mu_m = 0 \Rightarrow$ Mit Koeffizienten $\lambda_1, \dots, \lambda_n, \mu_1, \mu_m$ lassen sich $v_1, \dots, v_n, u_1, \dots, u_m$ linear zu Null kombinieren ohne, dass alle Koeffizienten Null sind $\Rightarrow v_1, \dots, v_n, u_1, \dots, u_m$ sind linear abhängig

□

Eine alternative Definition der linearen Unabhängigkeit motiviert Satz 3.5:

Satz 3.5: Sei V ein K -Vektorraum. Eine Menge $M \subseteq V$ ist genau dann linear unabhängig, wenn kein Vektor $v \in M$ als Linearkombination der anderen Vektoren dargestellt werden kann.

Beweis:

“ \Rightarrow ”: Annahme: $M \subseteq V$ sind linear unabhängig und es existiert ein Vektor $v \in M$, der als Linearkombination von endlichen Vektoren aus $M \setminus \{v\}$ dargestellt werden kann. D.h. es existieren $\lambda_1, \dots, \lambda_n \in K \setminus \{0\}, n \geq 1$ und $v_1, \dots, v_n \in M \setminus \{v\}$ mit

$$\sum_{i=1}^n \lambda_i v_i = v \Rightarrow -v + \sum_{i=1}^n \lambda_i v_i = 0 \nmid$$

Dies ist ein Widerspruch zur Annahme, dass die Vektoren linear unabhängig sind. Es existiert also kein Vektor, welcher als Linearkombination ausgedrückt werden kann.

“ \Leftarrow ”: Angenommen M wäre linear abhängig. D.h. es existieren $n \in \mathbb{N}$ und $v_1, \dots, v_n \in M, \lambda_1, \dots, \lambda_n \in K$ mit $\lambda_1 \neq 0$ und $\sum_{k=1}^n \lambda_k v_k = 0 \Rightarrow$

$$\begin{aligned} \sum_{k=1}^n \frac{\lambda_k}{\lambda_1} v_k = 0 &\Rightarrow v_1 + \sum_{k=2}^n \frac{\lambda_k}{\lambda_1} v_k = 0 \\ &\Rightarrow v_1 = - \sum_{k=2}^n \frac{\lambda_k}{\lambda_1} v_k \quad \downarrow \end{aligned}$$

Dies ist ein Widerspruch dazu, dass man v nicht als Linearkombination darstellen kann. Die Vektoren sind also linear unabhängig. □

Definition 3.6: Span für unendliche Mengen

Sei K ein Körper, V ein K -Vektorraum, M eine Menge und $v_m \in M \forall m \in M$ gegeben. Dann ist der Spann der Familie $\{v_m\}_{m \in M}$ gegeben durch

$$\text{Span}\{v_m\}_{m \in M} := \left\{ v \in V \mid \exists n \in \mathbb{N} \text{ und endliche Teilmenge } J \subset M, \right. \\ \left. |J| = n \text{ mit } v \in \text{Span}\{v_j\}_{j \in J} \right\}$$

Beispiel 3.7: $M = \mathbb{N}, v_m := t^m, t \in K, \text{Span}\{v_m\}_{m \in M} = P[t]$

Satz 3.8: Sei K ein Körper, V ein K -Vektorraum und M eine Menge. Dann sind folgende Aussagen äquivalent:

1. $\{v_m\}_{m \in M}$ ist linear unabhängig
2. jeder Vektor $v \in \text{Span}\{v_m\}_{m \in M}$ hat eine eindeutige Darstellung als Linearkombination

Beweis:

$1 \Rightarrow 2$: Beweis per Kontraposition

Seien $I, J \subseteq M$ endlich. Sei $\lambda_k \in K, k \in I$ und $\mu_k \in K, k \in J$. Betrachte den Vektor $v \in V$. Für diesen gilt:

$$v_I = \sum_{k \in I} \lambda_k v_k \text{ und } v_J = \sum_{k \in J} \mu_k v_k$$

Überlegung: Wähle ein $k \in I \cup J$. Falls gilt $k \in I \setminus J$, setze $\lambda_k = 0$. Falls gilt $k \in J \setminus I$, setze $\mu_k = 0$. Es folgt

$$\Rightarrow 0 = \sum_{k \in I \cup J} (\lambda_k - \mu_k) v_k$$

Da die Darstellung v_I und v_J von v unterschiedlich sind, existiert ein $k \in I \cup J$ mit $\lambda_k - \mu_k \neq 0$.

$\Rightarrow \{v_m\}_{m \in M}$ ist linear abhängig

2 \Rightarrow 1: Beweis per Kontraposition

Angenommen $\{v_m\}_{m \in M}$ wäre linear abhängig. Dann existiert ein endliches $J \subseteq M$ und $\lambda_k \in K$ für $k \in J$ mit $0 = \sum_{k \in J} \lambda_k v_k$ und mindestens ein $\lambda_k \neq 0$.

Sei $v \in \text{Span}\{v_m\}_{m \in M}$, d.h. es existiert ein endliches $I \subseteq M$ mit $\mu_k \in K$ für alle $k \in I$ mit

$$v = \sum_{k \in I} \mu_k v_k \Rightarrow v + 0 = \sum_{k \in I} \mu_k v_k + \sum_{k \in J} \lambda_k v_k = \sum_{k \in I \cup J} (\mu_k + \lambda_k) v_k$$

Es gilt wieder $\mu_k = 0$ für $k \in J \setminus I$ und $\lambda_k = 0$ für $k \in I \setminus J$. Da für mindestens ein $k \in J \cup I$, $(\mu_k + \lambda_k) \neq \lambda_k$, ist dies eine zweite Darstellung von v .

□

Definition 3.9: Erzeugendensystem

Sei K ein Körper, V ein K -Vektorraum, M eine Menge und v_m für $m \in M$ Vektoren in V . Die Menge $\{v_m\}_{m \in M}$ heißt **Erzeugendensystem** von V , falls

$$\text{Span}\{v_m\}_{m \in M} = V$$

Beispiel 3.10: Sei K ein Körper, V ein Vektorraum mit $V = K^n$, $n \in \mathbb{N}$

Dann ist mit

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

die Menge $\{e_i\}_{i \in \{1, \dots, n\}}$ ein Erzeugendensystem von K^n .

Definition 3.11: Basis

Sei K ein Körper, V ein K -Vektorraum, M eine Menge und v_m für $m \in M$ Vektoren in V . Dann heißt $\{v_m\}_{m \in M}$ **Basis** von V , falls sie linear unabhängig und ein Erzeugendensystem von V ist.

Beispiel 3.12: Das Erzeugendensystem aus 3.10 ist eine Basis. Anmerkung: Basen sind nicht eindeutig. Für K^3 gilt etwa

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

ist ebenfalls eine Basis.

Beispiel 3.13: Die Familie $\{t^i\}_{i \in \mathbb{N}}$ ist ein Erzeugendensystem von $P[t]$, denn es gilt $\text{Span}\{t^i\}_{i \in \mathbb{N}} = P[t]$. Um zu prüfen, ob die Familie auch Basis von $P[t]$ muss die lineare Unabhängigkeit geprüft werden. Sei $n \in \mathbb{N}_0$, $a_0, a_1, \dots, a_n \in K$ und betrachte $p(t) = a_0 + a_1 t + \dots + a_n t^n$ mit $p(t) = 0$.

Falls $a_k \neq 0$ für ein $k \in \{0, \dots, n\}$ gilt, so hat $p(t)$ höchstens n Nullstellen in K . $0 \in P[t]$ hat aber unendlich viele Nullstellen. D.h. es existiert ein $t \in K$ mit $p(t) \neq 0$. Es folgt, dass die Familie $\{t^i\}_{i \in \mathbb{N}}$ linear unabhängig ist. Somit ist $\{t^i\}_{i \in \mathbb{N}}$ eine Basis von $P[t]$.

Satz 3.14: Sei K ein Körper, V ein Vektorraum und $B := \{v_1, \dots, v_n\} \subseteq V$. Dann ist äquivalent

1. B ist eine Basis
2. B ist ein unverkürzbares Erzeugendensystem, d.h. $\forall r \in \{1, \dots, n\}$ ist die Menge $B \setminus \{v_r\}$ kein Erzeugendensystem von V
3. Für alle $v \in V$ existieren eindeutige $\lambda_1, \dots, \lambda_n$ mit

$$v = \sum_{k=1}^n \lambda_k v_k$$

4. B ist unverlängerbar linear unabhängig. D.h. für alle $v \in V$ ist $B \cup \{v\}$ linear abhängig

Beweis:

$1 \implies 2$: Beweis durch Kontraposition

Angenommen B ist ein verkürzbares Erzeugendensystem. D.h. o.B.d.A $r = 1$, $B \setminus \{v_1\}$ ist auch ein Erzeugendensystem von V . Also existieren $\lambda_2, \dots, \lambda_n \in K$ mit

$$v_1 = \sum_{k=2}^n \lambda_k v_k$$

Mit Satz 3.5 folgt, dass B linear abhängig ist. Also ist B keine Basis.

$2 \implies 3$: Beweis durch Kontraposition

Angenommen, es existiert $v \in V$, $\lambda_k, \mu_k \in K$ mit

$$v = \sum_{k=1}^n \lambda_k v_k \text{ und } v = \sum_{k=1}^n \mu_k v_k$$

o.B.d.A gilt $\mu_1 \neq \lambda_1$. Dann folgt $0 = \sum_{k=1}^n (\lambda_k - \mu_k) v_k$.

Weiterhin gilt

$$0 = \sum_{k=1}^n \frac{\lambda_k - \mu_k}{\lambda_1 - \mu_1} v_k \implies v_1 = \sum_{k=2}^n -\frac{\lambda_k - \mu_k}{\lambda_1 - \mu_1} v_k$$

Sei nun $w \in V$, dann existiert ein $\alpha_k \in K$ mit

$$w = \sum_{k=1}^n \alpha_k v_k = \alpha_1 v_1 + \sum_{k=2}^n \alpha_k v_k = \star$$

Wir definieren

$$\beta_k = -\frac{\lambda_k - \mu_k}{\lambda_1 - \mu_1}$$

Dann gilt

$$\star = \sum_{k=2}^n (\alpha_1 \beta_k + \alpha_k) v_k$$

Also war B kürzbar.

3 \Rightarrow 4:

Satz 3.8 liefert, dass B linear unabhängig ist. Sei $v \in V \setminus B$. Dann existieren $\lambda_1, \dots, \lambda_k \in K$ mit

$$v = \sum_{k=1}^n \lambda_k v_k$$

Das heißt $B \cup \{v\}$ ist linear abhängig nach Satz 3.8.

4 \Rightarrow 1:

Sei B unverlängerbar linear abhängig. Sei $v \in V \setminus B$. Dann ist $B \cup \{v\}$ linear abhängig. Also existiert $\lambda_1, \dots, \lambda_n, \lambda \in K$ mit

$$0 = \sum_{k=1}^n \lambda_k v_k + \lambda v \text{ wobei nicht alle } \lambda_k, \lambda = 0$$

Da B linear unabhängig ist folgt $\lambda \neq 0$. Daraus folgt

$$v = \sum_{k=1}^n -\frac{\lambda_k}{\lambda} v_k$$

Also ist B ein Erzeugendensystem und somit eine Basis.

□

Definition 3.15: endlichdimensional, unendlichdimensional

Sei $(V, +, \cdot)$ ein K -Vektorraum für den eine endliche Menge $M = \{v_1, \dots, v_n\} \subset V$ existiert, so dass $\text{Span } M = V$. Dann nennt man V **endlich erzeugt** und sagt V ist **endlichdimensional**. Ist V nicht von endlich vielen Vektoren erzeugt, nennt man V **unendlichdimensional**.

Beispiel 3.16:

- Die Einheitsvektoren aus Beispiel 3.10 sind eine Basis des K^n für einen Körper K . Damit ist K^n endlich erzeugt.
- Der Vektorraum $P[t]$ der Polynome aus Beispiel 3.13 ist über dem Körper K mit der Basis $\{t^i\}_{i \in \mathbb{N}_0}$ ist nicht endlich erzeugt.
- Sei V der Vektorraum der stetigen reellwertigen Funktionen auf dem Intervall $[0, 1]$. Dann ist V unendlichdimensional, denn:

Sei für $n \in \mathbb{N}$ die Funktion $f_n \in V$ definiert durch

$$f_n : [0, 1] \rightarrow \mathbb{R}, x \mapsto \begin{cases} 0, & 0 \leq x \leq \frac{1}{n+1} \\ 2n(n+1)x - 2n, & \frac{1}{n+1} \leq x \leq \frac{1}{2} \left(\frac{1}{n} + \frac{1}{n+1} \right) \\ -2(n+1)x + 2n + 2, & \frac{1}{2} \left(\frac{1}{n} + \frac{1}{n+1} \right) \\ 0, & \frac{1}{n} \leq x \leq 1 \end{cases}$$

Es gilt für jede Linearkombination der f_n und $j, k \in \mathbb{N}$ mit $j \leq k$, dass

$$\sum_{j=1}^k \lambda_j f_j \left(\frac{1}{2} \left(\frac{1}{j} + \frac{1}{j+1} \right) \right) = \lambda_j \text{ bei } f_j = 1, \text{ sonst } 0$$

Damit ist:

$$\sum_{i=1}^k \lambda_i f_i(x) = 0_v \in V, \quad \forall x \in [0, 1]$$

nur erfüllt, wenn $\lambda_i = 0$ für alle $1 \leq i \leq k$. Damit sind die f_n linear unabhängig. Also ist V unendlichdimensional.

Frage: Hat jeder Vektorraum eine Basis?

Diese Frage ist relativ einfach im endlichdimensionalen Fall:

Lemma 3.17: Basisauswahlsatz

Ein K -Vektorraum $(V, +, \cdot)$ ist genau dann endlich, erzeugt, wenn er eine endliche Basis besitzt.

Beweis:

“ \Leftarrow ”: endliche Basis \implies endliches Erzeugendensystem \implies endlich erzeugt

“ \implies ”: Sei V endlich erzeugt $\implies \exists v_1, \dots, v_n : \text{Span}\{v_1, \dots, v_n\} = V$. Ist dieses Erzeugendensystem nicht minimal, d.h. linear abhängig, dann folgt mit Satz 3.5, dass ein $v_i, 1 \leq i \leq n$, als Linearkombination der anderen $v_j, i \neq j$ dargestellt werden kann. Entfernen des v_i liefert ein kleineres Erzeugendensystem. Wiederhole $n - 1$ -Mal, bis die verbleibende Menge linear unabhängig ist. Somit enthält jedes endliche Erzeugendensystem ein minimales Erzeugendensystem und somit eine Basis.

□

Für den unendlichdimensionalen Fall ist mehr Arbeit nötig:

Satz 3.18: Jeder K -Vektorraum $(V, +, \cdot)$ besitzt eine Basis (ein minimales Erzeugendensystem).

Beweis:

Idee: Wir wenden das zornsche Lemma auf $M = V$ und

$$\mathcal{S} = \left\{ A \subseteq V \mid \text{die Familie } \{v\}_{v \in A} \text{ ist linear unabhängig} \right\} \subseteq \mathcal{P}(M)$$

Dazu treffen wir die Annahme, dass die Voraussetzungen für das Zornsche Lemma gelten. Dann hat \mathcal{S} ein maximales Element bezüglich der Relation \leq .

Da maximal linear unabhängige Familien von Vektoren aus V Basen von V sind (Satz 3.14), ist damit die Behauptung gezeigt.

Jetzt müssen wir die Verwednung des Zornschen Lemmas rechtfertigen.

Für $\mathcal{K} \subseteq \mathcal{S}$, \mathcal{K} ist eine Kette, gilt $\bigcup_{A \in \mathcal{K}} A \in \mathcal{S}$. Sei $\mathcal{K} \subseteq \mathcal{S}$ seine Kette. zu zeigen:

$$B := \bigcup_{A \in \mathcal{K}} A \in \mathcal{S}$$

D.h. die Vektoren aus B sind eine Menge von linear unabhängigen Vektoren. Seien dazu endlich viele Vektoren $v_1, \dots, v_n \in B$ beliebig vorgegeben. Per Definition von B existiert in der gegebenen Kette \mathcal{K} für jeden Index $i \in \{1, \dots, n\}$ eine Menge $A_i \in \mathcal{K}$ mit $v_i \in A_i$.

Nach Lemma 1.26 über endliche Teilmengen von Ketten gibt es einen Index \tilde{i} mit $A_i \subseteq A_{\tilde{i}}$ für alle i . \implies alle Vektoren $v_1, \dots, v_n \in A_{\tilde{i}} \in \mathcal{S}$. Daraus folgt, dass $\{v_1, \dots, v_n\}$ linear unabhängig ist.

□

3.2. Basen

Man kann eine Basis als Koordinatensystem in einem Vektorraum auffassen. Wichtig ist, dass Basen nicht eindeutig sind (vergleiche Beispiel 3.12). Eine sehr wichtige Frage der linearen Algebra ist: Welche Basis wählt man?

Beispiel 3.19: Um aus einer im Verhältnis 1 : 1 in Wasser gelösten Substanz eine Lösung im Mischungsverhältnis $a : b$ zu bekommen, verdünnt man y Teile der Lösung mit x Teilen Wasser, so dass

$$x \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + y \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \begin{array}{l} x\text{-Koordinate ist Wasser, } y\text{-Koordinate ist Substanz} \end{array}$$

Eine andere Darstellung der Basis ist

$$\begin{aligned} \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} &= 1 \cdot e_1 + \frac{1}{2} \cdot e_2 = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \frac{1}{2}v_1 + \frac{1}{2}v_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

Man sieht:

- 1 Parameter in \mathbb{R}^1
- 2 Parameter in \mathbb{R}^2
- 3 Parameter in \mathbb{R}^3

Ziel: Alle Basen eines endlich erzeugten Vektorraums haben gleich viele Elemente.

Das ist durchaus nicht offensichtlich!!!

Beispiel 3.20: Es gibt eine bijektive Abbildung $f : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$. Dafür kann man z.B. das Diagonalargument von Georg Cantor verwenden.

Zum Beweis der Aussage sind noch Vorarbeiten notwendig.

Satz 3.21: Basisergänzungssatz

Sei $(V, +, \cdot)$ ein K -Vektorraum,

- $\{u_i\}_{i \in I} \subseteq V$ ein linear unabhängiges System
- $\{v_j\}_{j \in J} \subseteq V$ ein Erzeugendensystem von V

Dann gibt es eine Teilmenge $\tilde{J} \subseteq J$ mit der Eigenschaft, dass das System

$$B := \{w_k\}_{k \in I \cup \tilde{J}} \text{ mit } w_k := \begin{cases} u_i, & k = i \in I \\ v_j, & k = j \in \tilde{J} \end{cases}$$

eine Basis von V bildet.

Beweis: Sei $\tilde{J} \subseteq J$ eine bezüglich \subseteq maximal gewählte Teilmenge mit der Eigenschaft, dass B wie im Satz definiert, ein linear unabhängiges System ist. Für endliche Mengen J ist das klar (siehe Lemma 3.17). Für Mengen mit unendlich vielen Elementen folgt dies aus dem zornschen Lemma (Satz 3.18). Damit: B ist ein maximales linear unabhängiges System.

Wegen der Maximalität ist für jeden Index $j \in J \setminus \tilde{J}$ das System $B \cup \{v_j\}$ linear abhängig $\Rightarrow \exists \lambda, \lambda_i \in K, i \in I \cup \tilde{J}$:

$$\lambda \cdot v_j + \sum_{i \in I \cup \tilde{J}} \lambda_i \cdot v_i = 0$$

wegen der linearen Unabhängigkeit der $\{v_i\}_{i \in I \cup \tilde{J}}$ muss $\lambda \neq 0$ gelten

$$\Rightarrow v_j = - \sum_{i \in I \cup \tilde{J}} \frac{\lambda_i}{\lambda} \cdot w_i \in \text{Span } B$$

Dies gilt für alle $j \in J \setminus \tilde{J} \Rightarrow v_j \in \text{Span } B \ \forall j \in J \Rightarrow B$ Basis

□

Beispiel 3.22: In $V = \mathbb{R}^3$ bilden die Vektoren $\{e_1, e_2, e_3\}$ die Standardbasis. Des Weiteren sind die Vektoren $u_1 = (3, 1, 0)$ und $u_2 = (1, 3, 0)$ linear unabhängig. Satz 3.21 liefert, dass $\{u_1, u_2, e_3\}$ eine Basis ist.

Satz 3.23: Austauschsatz von Steinitz

(Ernst Steinitz, deutscher Mathematiker, 1871 - 1928)

Sei $(V, +, \cdot)$ ein K -Vektorraum,

- $B = \{v_1, \dots, v_n\}$ eine (endliche Basis)
- $C = \{u_1, \dots, u_m\}$ eine linear unabhängige Familie

Dann ist $m \leq n$ und nach geeigneten umnummerieren der Vektoren in B ist das durch austauschen der ersten m -Vektoren erhaltene System

$$\tilde{B} := \{u_1, \dots, u_m, v_{m+1}, \dots, v_n\} \text{ eine Basis von } V \text{ über } K$$

Beweis: Aus dem Basisergänzungssatz folgt, dass man das linear unabhängige System u_1, \dots, u_m zu einer Basis $u_1, \dots, u_m, v_{j_1}, \dots, v_{j_{k_0}}$ für ein $k_0 \geq 0$ und geeignete Indizes $j_1, \dots, j_{k_0} \in \{1, \dots, n\}$ erweitern kann.

Die Menge $\{u_1, \dots, u_{m-1}, v_{j_1}, \dots, v_{j_{k_0}}\}$ ist immer noch linear unabhängig aber keine Basis. Aus dem Basisergänzungssatz folgt wieder, dass man diese Menge zu einer Basis $\{u_1, \dots, u_{m-1}, v_{j_1}, \dots, v_{j_{k_0}}, v_{j_{k_0+1}}, \dots, v_{j_{k_1}}\}$ für ein $k_1 > k_0$ und weitere Indizes $j_{k_0+1}, \dots, j_{k_1} \in \{1, \dots, n\}$ setzt. Man dieses Verfahren induktiv v -mal fort, erhält man im r -ten Schritt eine Basis $\{u_1, \dots, u_{m-r}, v_{j_1}, v_{j_2}, \dots, v_{j_{k_{r-1}}}, v_{j_{k_{r-1}+1}}, \dots, v_{j_{k_r}}\}$ für ein $k_r > k_{r-1}$ und weitere Indizes $j_{k_{r-1}+1}, \dots, j_{k_r} \in \{1, \dots, n\}$. Nach m Schritten sind alle $u_i, 1 \leq i \leq m$ ersetzt.

Als neue Basis erhält man

$$\hat{B} := \{v_{j_1}, \dots, v_{j_m}\}$$

welche ausschließlich Vektoren aus B enthält. B war ein minimales Erzeugendensystem. Also muss die neue Menge \hat{B} bis auf die Umordnung mit der Menge B übereinstimmen. Für die Menge der Indizes gilt also:

$$\hat{B} = \{v_{j_1}, \dots, v_{j_m}\} = \{v_1, \dots, v_n\} = B$$

Also folgt: $k_m = n$

Wir haben jeweils mindestens einen Vektor ergänzt, d.h.

$$k_m > k_{m-1} > \dots > k_1 > 0 \implies k_m \geq m \implies n = k_m \geq m \implies n \geq m \implies \text{erste Aussage}$$

Dann folgt die zweite Aussage aus dem Basisergänzungssatz.

□

Lemma 3.24: Ist $(V, +, \cdot)$ ein von endlich vielen Vektoren erzeugter K -Vektorraum, so besitzt V eine Basis und je zwei Basen von V haben gleich viele Elemente.

Beweis: Sei $V = \text{Span}\{v_1, \dots, v_n\}$ mit $v_1 \neq 0$. Nach Satz 3.21 kann $\{v_1\}$ durch Hinzunahme von geeigneten Elementen aus $\{v_2, \dots, v_n\}$ zu einer Basis von V ergänzen. Also besitzt V eine Basis mit endlich vielen Elementen.

Seien $U = \{u_1, \dots, u_l\}$ und $W = \{w_1, \dots, w_k\}$ zwei solche Basen. Dann folgt aus dem Satz 3.23 aus Symmetrie, dass $k = l$.

□

Ausblick: Man kann mit Konzepten der Mengenlehre auch zeigen, dass es für unendlich erzeugte Vektorräume V gilt: Für je zwei Basen $\{u_i\}_{i \in I}$ und $\{w_j\}_{j \in J}$ von V existiert eine bijektive Abbildung $f : I \rightarrow J$.

Folgerung aus Satz 3.8 in Zusammenhang mit Lemma 3.24: Da für eine Basis $B := \{v_1, \dots, v_n\}$ eines K -Vektorraums V gilt, dass $\text{Span } B = V$, sind für $v \in V$ die Koeffizienten (= Koordinaten) $\lambda_1, \dots, \lambda_n$ zur Darstellung von v eindeutig.

Beispiel 3.25: für $V = \mathbb{R}^3$ sind

$$B_1 := \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \text{ und } B_2 := \left\{ \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} -5 \\ 0 \\ 2 \end{pmatrix} \right\}$$

zwei Basen. Der Vektor $v = (-5, 11, 2)$ besitzt bezüglich B_1 die Koordinaten $\lambda_1 = -5, \lambda_2 = 11, \lambda_3 = 2$. Was sind die Koordinaten bezüglich B_2 ? Es muss gelten:

$$a \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + b \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} + c \begin{pmatrix} -5 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -15 \\ 11 \\ 2 \end{pmatrix}$$

anders kann man dies notieren als

$$\begin{pmatrix} 2a - 2b - 5c \\ -a + 3b + 0c \\ 0a + 0b + 2c \end{pmatrix} = \begin{pmatrix} -15 \\ 11 \\ 2 \end{pmatrix}$$

Berechnung der a, b, c über ein lineares Gleichungssystem ergibt $a = 2, b = 3, c = 1$.

Beispiel 3.26: Für den \mathbb{R} -Vektorraum $P_{\leq 2}[t]$ der reellen Polynome vom Grad ≤ 2 sind:

$$B_1 := \{1, t, t^2\} \text{ und } B_2 := \{-t, t^2 - 2, t^2 + 2\}$$

zwei Basen. Sei $p(t) = a + bt + ct^2$ ein beliebiges Polynom aus $P_{\leq 2}[t]$. Dann sind die Koeffizienten für B_1 : $\lambda_1 = a, \lambda_2 = b, \lambda_3 = c$. Für B_2 gilt:

$$\text{Basiswechsel, siehe Kapitel 4} \quad \begin{cases} 1 = -\frac{1}{4}(t^2 - 2) + \frac{1}{4}(t^2 + 2) \\ t = (-1)(-t) \\ t^2 = \frac{1}{2}(t^2 + 2) + \frac{1}{2}(t^2 - 2) \end{cases}$$

Es gilt

$$\begin{aligned} a1 + bt + ct^2 &= a \left(-\frac{1}{4}(t^2 - 2) + \frac{1}{4}(t^2 + 2) \right) + b(-1)(-t) + c \left(\frac{1}{2}(t^2 + 2) + \frac{1}{2}(t^2 - 2) \right) \\ &= -b(-t) + \left(-\frac{a}{4} + \frac{c}{2} \right)(t^2 - 2) + \left(\frac{a}{4} + \frac{c}{2} \right)(t^2 + 2) \end{aligned}$$

3.3. Dimensionen

Definition 3.27: Dimension eines Vektorraums

Die Dimension eines Vektorraum $(V, +, \cdot)$ über K ist definiert als:

$$\dim_K(V) := \begin{cases} n & \text{falls } V \text{ eine Basis der Länge } n \text{ hat} \\ \infty & \text{sonst} \end{cases}$$

Wenn der Kontext klar ist schreibt man $\dim V$.

Beispiel 3.28: Sei K ein Körper. Es gilt

- $\dim_K(V) = 0$ genau dann, wenn $V = \{0\}$
- für $V = K^n$ folgt mit der Standardbasis, dass $\dim_K(V) = n$
- für die Dimension eines Vektorraums ist der jeweilige Grundkörper K entscheidend, z.B. \mathbb{C} und $K = \mathbb{C}$ gilt $\dim_{\mathbb{C}} V = 1$ für $K = \mathbb{R}$ aber $\dim_{\mathbb{R}} V = 2$.
- der K -Vektorraum $P[t]$ ist nicht endlich erzeugt, also $\dim_K P[t] = \infty$

Beispiel 3.29: Sei $V = K^n$ für einen Körper K . Um zu prüfen, dass n Vektoren aus V eine Basis werden, muss nur deren lineare Unabhängigkeit geprüft werden. Seien z.B. B in $V = \mathbb{R}^3$ die Vektoren

$$v_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} -5 \\ 0 \\ 2 \end{pmatrix}$$

gegeben. Sind diese linear unabhängig?

$$\lambda_1 \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -5 \\ 0 \\ 2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Dazu wird ein lineares Gleichungssystem aufgestellt

$$\begin{pmatrix} 2a - 2b - 5c \\ -a + 3b + 0c \\ 0a + 0b + 2c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow a = 0, b = 0, c = 0.$$

Somit sind die Vektoren linear unabhängig.

Lemma 3.30: Sei $(V, +, \cdot)$ ein K -Vektorraum mit $n := \dim_K V < \infty$ und $B = \{v_1, \dots, v_n\} \subset V$ eine Familie von genau n Vektoren. Dann sind folgende Aussagen äquivalent:

1. B ist eine Basis
2. B ist linear unabhängig
3. B ist ein Erzeugendensystem

Beweis: (Übungsaufgabe, Blatt 8, Aufgabe 1)

$1 \Rightarrow 2$ und $1 \Rightarrow 3$: Folgt direkt aus der Definition einer Basis.

$2 \Rightarrow 1$: Angenommen B ist keine Basis von V . Da B linear unabhängig ist, kann B kein Erzeugendensystem von V sein. Mit dem Basisergänzungssatz (Satz 3.21) lässt sich das linear unabhängige System B mit Vektoren eines Erzeugendensystems von V , wir wählen V selbst, zu einer Basis B' ergänzen. Definiere die B' wie folgt:

$$B' = B \cup \{v_{n+1}, \dots, v_{n+k}\}, \quad v_{n+1}, \dots, v_{n+k} \in V \setminus B$$

Dann hat B' mit Sicherheit mehr als n Elemente. Dies ist ein Widerspruch dazu, dass die Dimension des Vektorraums n ist (Definition 3.27 und Lemma 3.24). Die Annahme muss also falsch sein. $\Rightarrow B$ ist eine Basis von V .

$3 \Rightarrow 1$: Angenommen B ist keine Basis von V . Da B ein Erzeugendensystem ist, kann B nicht linear unabhängig sein. Daraus folgt, dass B zu einem minimalen Erzeugendensystem, also einer Basis B' verkürzbar ist. Definiere B' wie folgt:

$$B' = B \setminus \{v_{n_0}, \dots, v_{n_k}\}, \quad v_{n_0}, \dots, v_{n_k} \in B$$

Dann hat B' mit Sicherheit weniger als n Elemente. Dies ist ein Widerspruch dazu, dass die Dimension des Vektorraums n ist. Die Annahme muss also falsch sein. $\Rightarrow B$ ist eine Basis von V .

□

Lemma 3.31: Sei $(V, +, \cdot)$ ein endlich erzeugter K -Vektorraum. Jeder Untervektorraum $W \subseteq V$ ist dann ebenfalls endlichdimensional und es gilt:

$$\dim W \leq \dim V$$

mit Gleichheit genau dann, wenn $W = V$.

Beweis: jede linear unabhängige Familie in W ist auch linear unabhängig in V . Damit besteht sie nach dem Austauschsatz von Steinitz aus höchstens $\dim_K V$ Elementen. Daraus folgt, dass W endlich erzeugt ist mit $\dim_K W \leq \dim_K V$. Im Fall von $\dim_K W = \dim_K V$ folgt mit Lemma 3.30, dass jede Basis von W auch eine Basis von V ist. $\Rightarrow V = W$

□

Achtung: Die letzte Aussage (d.h. $V = W$) gilt nicht für unendlich erzeugte Vektorräume. Denn der K -Vektorraum $P[t]$ aller Polynome hat die Basis der Monome $\{t^n\}_{n \in \mathbb{N}_0}$. Die Menge aller Polynome aus $P[t]$ mit $a_0 = 0$ ist ein Unterraum mit der Basis $\{t^n\}_{n \in \mathbb{N}}$ und wird mit W bezeichnet. Dann gilt

$$\dim_K P[t] = \infty = \dim_K W, \text{ aber } P[t] \neq W$$

3.4. Direkte Summen

Aus der Definition von Mengenoperationen aus dem ersten Kapitel folgt: sind U_1 und U_2 zwei Unterräume des K -Vektorraums $(V, +, \cdot)$, so gilt für ihren Durchschnitt:

$$U_1 \cap U_2 = \{u \in V \mid u \in U_1 \wedge u \in U_2\}$$

Definition 3.32: Summe von Mengen

Sei $(V, +, \cdot)$ ein K -Vektorraum für die Unterräume $U_1, \dots, U_r \subseteq V$ definiert man ihre Summe als die Teilmenge

$$U_1 + U_2 + \dots + U_r := \{u_1 + u_2 + \dots + u_r \mid u_i \in U_i \text{ für } 1 \leq i \leq r\} \subseteq V$$

Für den Durchschnitt und die Summe von Untervektorräumen gelten folgende Regeln:

Lemma 3.33: Sind U_1, U_2, U_3 Unterräume des K -Vektorraums $(V, +, \cdot)$, dann gilt:

1. $U_1 \cap U_2$ und $U_1 + U_2$ sind Unterräume von V
2. $U_1 + (U_2 + U_3) = (U_1 + U_2) + U_3$ und $U_1 + U_2 = U_2 + U_1$
3. $U_1 + \{0\} = U_1$ und $U_1 + U_1 = U_1$
4. $U_1 \subseteq U_1 + U_2$ mit Gleichheit, d.h. $U_1 = U_1 + U_2$, wenn $U_2 \subseteq U_1$

Beweis: (Übungsaufgabe)

□

Beispiel 3.34: Sei $V := \mathbb{R}^3$, $U_1 := \text{Span}\{(1, 0, 0), (0, 1, 0)\}$, $U_2 := \{(0, 1, 0), (0, 0, 1)\}$. Dann gilt für $v \in V$, $v = (v_1, v_2, v_3)$

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

Also gilt: $U_1 + U_2 = V$, insbesondere gilt auch $\dim(U_1 + U_2) = 3$, $\dim U_1 = 2 = \dim U_2$. Weiterhin gilt

$$v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in U_1 \text{ und } v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in U_2$$

Also ist die Darstellung von v als Summe **nicht** eindeutig. Insbesondere ist $\dim(U_1 \cap U_2) = 1$.

Lemma 3.35: Sei V ein K -VR und $r \in \mathbb{N}$ und U_1, \dots, U_r Untervektorräume von V . Dann sind folgende Aussagen äquivalent.

1. Für $u \in \sum_{i=1}^r U_i$ existieren eindeutige $u_i \in U_i, i \in \{1, \dots, r\}$ mit $u = \sum_{i=1}^r u_i$
2. Für $u_i \in U_i, i \in \{1, \dots, r\}$ mit $0 = \sum_{i=1}^r u_i$ gilt

$$u_i = 0 \quad i \in \{1, \dots, r\}$$

3. Für $i \in \{1, \dots, r\}$ gilt

$$U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^r U_j = \{0_V\}$$

Beweis: (Übungsaufgabe, Blatt 9, Aufgabe 1)

1 \Rightarrow 2:

Eine Darstellung des Nullvektors ist $0 = 0_1 + \dots + 0_r$ mit $0_i \in U_i$. Da jeder Vektor eine eindeutige Darstellung besitzt, folgt 2).

2 \Rightarrow 3: Beweis durch Kontraposition

Angenommen es gilt $U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^r U_j \neq \{0_V\}$. Dann existiert ein Vektor v mit $v \in U_i$ und $v \in U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_r$. Da U_i und $U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_r$ Untervektorräume von V bilden, gilt auch $-v \in U_i$ und $-v \in U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_r$. Daraus folgt

$$\sum_{j=1}^r u_j = v + \sum_{\substack{j=1 \\ j \neq i}}^r u_i = v + (-v) = 0$$

Somit gilt

$$\sum u_i = 0 \text{ aber } u_j \neq 0, j \in \{1, \dots, r\}$$

3 \Rightarrow 1: Beweis durch Kontraposition

Angenommen ein Vektor $v \in V$ besitzt keine eindeutige Darstellung. Dann gilt für ein $u \in U_i$:

$$u = \sum_{j=1}^r u_j \text{ mit } u_j = 0, j \neq i \text{ und } u_j = u, j = i$$

Dann gilt $u \in U_i$ und insbesondere $u \in U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_r$. Somit ist die Schnittmenge $U_i \cap U_1 + \dots + U_{i-1} + U_{i+1} + \dots + U_r \neq \{0_V\}$.

□

Definition 3.36: Direkte Summe

Sei V ein K -VR und $r \in \mathbb{N}$, U_1, \dots, U_r Untervektorräume von V . Dann heißt die Summe $\sum U_i$ **direkt**, falls eine der Bedingungen aus Lemma 3.35 zutrifft.

Wir schreiben dann

$$U_1 \oplus U_2 \oplus \dots \oplus U_r$$

Beispiel 3.37: Seien V, U_1, U_2 wie im Beispiel 3.34. Dann gilt $V = U_1 + U_2$, aber nicht $U_1 \oplus U_2$. Sei weiter $U_3 := \text{Span}\{(0, 0, 1)\}$. Dann gilt $V = U_1 + U_3$ und $U_1 \oplus U_3$.

Lemma (ohne Nummer)

Sei V ein K -VR. Seien U_1, U_2 UVRs von V . Dann gilt

$$\dim(U_1 + U_2) \leq \dim U_1 + \dim U_2 \quad (*)$$

Falls $U_1 \oplus U_2$, gilt sogar Gleichheit.

Dabei sei $\infty + \infty = \infty$, $\infty + n = \infty$, $n \leq \infty$ für $n \in \mathbb{N}$ und es gilt $\infty \leq \infty$.

Beweis:

Fall 1: $\dim U_1 = \infty$ oder $\dim U_2 = \infty$. Dann gilt $(*)$ nach den Rechenregeln der erweiterten Arithmetik.

Andernfalls existieren $m, l \in \mathbb{N}$ mit $\dim U_1 = m$ und $\dim U_2 = l$.

Sei u_1, \dots, u_m eine Basis von U_1 und u_{m+1}, \dots, u_{m+l} eine Basis von U_2 . Sei weiter $u \in U_1 + U_2$, dann existieren $v \in U_1$ und $w \in U_2$, sodass $u = v + w$. Zu v und w existieren $\lambda_1, \dots, \lambda_m \in K$ bzw. $\lambda_{m+1}, \dots, \lambda_{m+l} \in K$ mit

$$v = \sum_{i=1}^m \lambda_i u_i \quad \text{und} \quad w = \sum_{i=m+1}^{m+l} \lambda_i u_i$$

Also $u = v + w$

$$u = \sum_{i=1}^m \lambda_i u_i + \sum_{i=m+1}^{m+l} \lambda_i u_i = \sum_{i=1}^{m+l} \lambda_i u_i$$

Also ist u_1, \dots, u_{m+l} ein Erzeugendensystem von $U_1 + U_2$. Es folgt $\dim(U_1 + U_2) \leq m + l = \dim U_1 + \dim U_2$.

Sei nun $U_1 \oplus U_2$. Falls $\dim(U_1 + U_2) = \infty$ gilt $(*)$ mit Gleichheit. Andernfalls existieren $n \in \mathbb{N}$ mit $\dim(U_1 + U_2) = n$. Da U_1 und U_2 Untervektorräume von $U_1 + U_2$ sind, existieren $m, l \in \mathbb{N}$ mit $\dim U_1 = m \leq n$ und $\dim U_2 = l \leq n$. Sei wieder u_1, \dots, u_m eine Basis von U_1 und u_{m+1}, \dots, u_{m+l} eine Basis von U_2 . Seien $\lambda_1, \dots, \lambda_{m+l} \in K$ mit

$$0 = \sum_{i=1}^{m+l} \lambda_i u_i = \underbrace{\sum_{i=1}^m \lambda_i u_i}_{=v \in U_1} + \underbrace{\sum_{i=m+1}^{m+l} \lambda_i u_i}_{=w \in U_2}$$

Da $U_1 \oplus U_2$ folgt $v = 0 = w$. Da $u_i, i \in \{1, \dots, m\}$ eine Basis von U_1 ist, folgt $0 = \lambda_i u_i \in \{1, \dots, m\}$. Analog folgt dies für $\lambda_{m+1}, \dots, \lambda_{m+l}$. Also ist u_1, \dots, u_{m+l} linear unabhängig.

$$\Rightarrow \dim U_1 + \dim U_2 = m + l \leq \dim(U_1 + U_2)$$

□

Satz 3.38: Sei V ein K -VR und U ein UVR von V . Dann existiert ein Untervektorraum $U^\top \subseteq V$ mit $V = U \oplus U^\top$ (heißt $V = U + U^\top$ und $U \oplus U^\top$). Insbesondere gilt dann

$$\dim V = \dim U + \dim U^\top$$

Beweis: Sei $(u_i)_{i \in I}$ eine Basis von U . Nach Satz 3.21 existiert eine Menge J und Vektoren $w_j, j \in J$ mit

$$I \cap J = \emptyset \text{ und } v_k := \begin{cases} u_k & k \in I \\ w_k & k \in J \end{cases} \quad k \in I \cup J$$

ist eine Basis von V .

Mit $U^\top = \text{Span}\{w_j\}_{j \in J}$ gilt dann $V = U + U^\top$. Sei $v \in V$, dann existieren eindeutige $\{\lambda_i\}_{i \in I} \subseteq K$ mit

$$v = \sum_{k \in I \cup J} \lambda_k v_k = \underbrace{\sum_{k \in I} \lambda_k v_k}_{\in U} + \underbrace{\sum_{k \in J} \lambda_k v_k}_{\in U^\top}$$

die Eindeutigkeit der $\lambda_k, k \in I \cup J$ garantiert die Eindeutigkeit von u und w . Also $U \oplus U^\top$.

□

Ein durch Satz 3.38 aus U und V erhaltener Untervektorraum U^\top heißt **Komplement** von U in V .

Beispiel 3.39: Seien V, U_1 und U_3 wie in Beispiel 3.37. Dann gilt $V = U_1 \oplus U_3$ d.h. U_3 ist ein Komplement von U_1 in V . Sei weiter $\tilde{U}_3 := \text{Span}\{(1, 0, 1)\}$. Dann gilt auch $V = U_1 \oplus \tilde{U}_3$.

Insbesondere sind die Komplemente aus Satz 3.38 nicht eindeutig bestimmt.

Satz 3.40: Sei V ein endlich erzeugter K -VR. Seien U_1, U_2 UVRs von V . Dann gilt

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

Beweis: Sei $U := U_1 \cap U_2$ und für $i \in \{1, 2\}$ sei W_i das Komplement von U in U_i . Es gilt

$$U_i = U \oplus W_i, \quad i \in \{1, 2\}$$

Dann gilt $U_1 + U_2 = U + W_1 + U + W_2 = U + W_1 + W_2$

$$\begin{aligned}
 W_1 \cap (U + W_2) &= W_1 \cap W_2 \\
 &= W_1 \cap U_1 \cap U_2 \text{ (da } W_1 \subseteq U_1) \\
 &= W_1 \cap U \\
 &= \{0_V\}
 \end{aligned}$$

Analog folgt $W_2 \cap (U + W_1) = \{0_V\}$. Sei $u \in U, w_1 \in W_1, w_2 \in W_2$ mit $0 = u + w_1 + w_2$, dann gilt

$$\begin{aligned}
 w_1 &= -(u + w_2) \\
 w_2 &= -(u + w_1)
 \end{aligned}$$

Also $w_1 = 0$ und $w_2 = 0$ und damit auch $u = 0$. Also $U \oplus W_1 \oplus W_2$. Es folgt

$$\begin{aligned}
 \dim(U_1 + U_2) &= \dim(U + W_1) + \dim W_2 \\
 &= \dim U + \dim W_1 + \dim W_2
 \end{aligned}$$

Aus der Wahl von W_i folgt

$$\begin{aligned}
 \dim U_1 &= \dim U + \dim W_1 \\
 \dim U_2 &= \dim U + \dim W_2
 \end{aligned}$$

durch Einsetzen erhält man

$$\dim U + \dim(U_1 + U_2) = \dim U_1 + \dim U_2$$

oder (nur für endliche wegen $-\infty$)

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim U$$

□

4. Lineare Abbildungen

Nun behandeln wir Abbildungen, die zur Vektorstruktur “passen”. Diese heißen lineare Abbildungen. Unser Ziel ist es, die Eigenschaften von linearen Abbildungen zu analysieren.

4.1. Definitionen und grundlegende Eigenschaften

Definition 4.1: Lineare Abbildung

Seien V und W zwei K -Vektorräume. Eine Abbildung $f : V \rightarrow W$ heißt **lineare Abbildung**, wenn gilt

1. $\underbrace{f(\lambda \cdot v)}_{\text{Skalarmultiplikation in } V} = \underbrace{\lambda \cdot f(v)}_{\text{Skalarmultiplikation in } W} \quad \forall v \in V, \forall \lambda \in K$
2. $\underbrace{f(v + w)}_{\text{Addition in } V} = \underbrace{f(v) + f(w)}_{\text{Addition in } W} \quad \forall v, w \in V$

Die Menge aller linearen Abbildungen von V nach W bezeichnet man mit $L(V, W)$. Eine lineare Abbildung $f : V \rightarrow W$ wird auch **lineare Transformation** oder **(Vektorraum-) Homomorphismus** genannt.

Eine bijektive lineare Abbildung nennt man **(Vektorraum-) Isomorphismus**. Gibt es für zwei K -Vektorräume V und W einen Isomorphismus, so heißen die Räume V und W isomorph, geschrieben

$$V \cong W$$

Eine lineare Abbildung $f : V \rightarrow V$ heißt **Endomorphismus** und ein bijektiver Endomorphismus heißt **Automorphismus**.

Bemerkung: Als Übungsaufgabe:

$$\text{Definition 4.1, 1) + 2) } \iff f(\lambda v + \mu w) = \lambda f(v) + \mu f(w) \quad \forall v, w \in V, \forall \lambda, \mu \in K$$

Beispiel 4.2: Für $a \in \mathbb{R}$ ist $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = ax$ eine lineare Abbildung. Ihr Graph ist eine Gerade durch den Ursprung.

Betrachte eine Gerade $f(x) = ax + b$ und betrachte

$$\begin{aligned} f(x + y) &= a(x + y) + b \\ f(x) + f(y) &= a(x) + b + a(y) + b = a(x + y) + 2b \end{aligned}$$

f ist also nur eine lineare Abbildung, wenn $b = 0$ gilt. Streng genommen sind Geraden der Form $f(x) = mx + n$ keine linearen Abbildung. Korrekt ist, sie als affine Abbildungen zu bezeichnen.

Beispiel 4.3: Für $a, b, c, d \in \mathbb{R}$ ist $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} \in \mathbb{R}^2$$

eine lineare Abbildung, denn

1)

$$\begin{aligned} f\left(\lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) &= \begin{pmatrix} a\lambda x_1 + b\lambda x_2 \\ c\lambda x_1 + d\lambda x_2 \end{pmatrix} = \begin{pmatrix} \lambda(ax_1 + bx_2) \\ \lambda(cx_1 + dx_2) \end{pmatrix} \\ &= \lambda \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} = \lambda f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) \quad \checkmark \end{aligned}$$

2)

$$\begin{aligned} f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) &= f\left(\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}\right) = \begin{pmatrix} a(x_1 + y_1) + b(x_2 + y_2) \\ c(x_1 + y_1) + d(x_2 + y_2) \end{pmatrix} \\ &= \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} + \begin{pmatrix} ay_1 + by_2 \\ cy_1 + dy_2 \end{pmatrix} = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) + f\left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \end{aligned}$$

Damit haben wir auch bewiesen, dass für $\varphi \in \mathbb{R}$ die Abbildung

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} \cos(\varphi)x_1 - \sin(\varphi)x_2 \\ \sin(\varphi)x_1 + \cos(\varphi)x_2 \end{pmatrix}$$

linear ist. Dies ist eine Drehung um den Ursprung mit dem Drehwinkel φ . Für $\varphi = 45^\circ$

$$f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = (\cos(45^\circ), \sin(45^\circ)) = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Beispiel 4.4: Sei $V = C^\infty(\mathbb{R})$ der reelle Vektorraum aller unendlich oft differenzierbaren Funktionen $g : \mathbb{R} \rightarrow \mathbb{R}$ mit punktweisen Addition und skalaren Multiplikation. Dann ist

$$\frac{d}{dx} : V \rightarrow V, f \mapsto f' = \frac{d}{dx} f$$

eine lineare Abbildung, denn für alle $f, g \in V$ und $a, b \in \mathbb{R}$ gilt:

$$\begin{aligned} (f + g)'(x) &= f'(x) + g'(x) \quad \forall x \in \mathbb{R} \\ (a \cdot f)'(x) &= af'(x) \quad \forall x \in \mathbb{R} \end{aligned}$$

Lemma 4.5: Seien V, W K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung, 0_V der Nullvektor in V und 0_W der Nullvektor in W . Dann gilt

1. $f(0_V) = 0_W$
2. $f(-x) = -f(x) \quad \forall x \in V$

Beweis: Es folgt aus Definition 4.1, 1), dass

$$\begin{aligned} f(0_V) &= f(0_K \cdot 0_V) = 0_K \cdot f(0_V) = 0_W \\ f(-x) &= f((-1) \cdot x) = (-1) \cdot f(x) = -f(x) \end{aligned}$$

□

Lemma 4.6: Seien V, W K -Vektorräume. Für $f, g \in L(V, W)$ und $\lambda \in K$ seien $f + g$ und $\lambda \cdot f$ definiert durch

$$\begin{aligned} (f + g)(v) &= f(v) + g(v) \quad \forall v \in V \\ (\lambda \cdot f)(v) &= \lambda f(v) \quad \forall v \in V, \forall \lambda \in K \end{aligned}$$

Dann ist $(L(V, W), +, \cdot)$ ein K -Vektorraum.

Beweis: (Übungsaufgabe)

□

Lemma 4.7: Sei V ein K -Vektorraum und $B := \{v_1, \dots, v_n\} \subset V$ eine endliche Familie von Vektoren. Dann ist:

$$\Phi_B : K^n \rightarrow V, (a_i)_{1 \leq i \leq n} \mapsto \sum_{i=1}^n a_i v_i$$

ein Homomorphismus von K -Vektorräumen.

Beweis: Seien $\lambda \in K$ und zwei Tupel $a = (a_1, \dots, a_n) \in K^n$, $b = (b_1, \dots, b_n) \in K^n$ gegeben. Mit der Definition der direkten Summe ist.

$$a + \lambda b = (a_i + \lambda b_i)_{1 \leq i \leq n} \in K^n$$

und deswegen gilt

$$\begin{aligned} \Phi_B(a + \lambda b) &= \Phi_B(a_i + \lambda b_i)_{1 \leq i \leq n} = \sum_{i=1}^n \underbrace{(a_i + \lambda b_i)}_{\in K} \underbrace{v_i}_{\in V} \\ &= \sum_{i=1}^n a_i v_i + \lambda \sum_{i=1}^n b_i v_i = \Phi_B(a) + \lambda \Phi_B(b) = \text{Def 4.1} \end{aligned}$$

□

Wichtig: $B = \{v_1, \dots, v_n\}$ ist endlich! Für B mit unendlich vielen Elementen bräuchte man eine äußere direkte Summe.

Aus abstrakter Sicht kennen wir jetzt endlichdimensionale Vektorräume, denn: Jeder endlichdimensionale Vektorraum ist isomorph zu einer direkten Summe von Kopien des Grundkörpers.

Satz 4.8: Struktursatz für Vektorräume

Sei V ein K -Vektorraum und $B = \{v_1, \dots, v_n\} \subset V$ eine Basis von V . Dann ist die Abbildung

$$\Phi_B : K^n \rightarrow V, (\lambda_i)_{1 \leq i \leq n} \mapsto \sum_{i=1}^n \lambda_i v_i$$

ein Isomorphismus, d.h. $K^n \cong V$.

Beweis: Nach Lemma 4.7 ist Φ_B eine lineare Abbildung. Zu zeigen: Φ_B ist bijektiv

$$B \text{ Basis} \implies \text{Span } B = V \implies \Phi_B \text{ surjektiv}$$

$$B \text{ Basis} \implies \text{jedes } v \in V \text{ besitzt eine eindeutige Darstellung } v = \sum_{i=1}^n \lambda_i v_i$$

□

Für eine gegebene Basis B nenn man Φ_B auch Koordinatenabbildung.

$$\Phi \text{ bijektiv} \implies \exists \Phi_B^{-1} \text{ als inverse Abbildung}$$

Beispiel 4.9: Sei $K^{n+1} = \mathbb{R}^{n+1}$ und $P_{\leq n}[t]$ der Raum der Polynome von Grad kleiner gleich n für $n \in \mathbb{N}$. Eine Basis von $P_{\leq n}[t]$ ist gegeben durch $B = \{1, t, t^2, \dots, t^n\}$, vergleiche Beispiel 3.13.

Dann ist:

$$\Phi_B : K^{n+1} \rightarrow P_{\leq n}[t], (a_i)_{0 \leq i \leq n} \mapsto \sum_{i=0}^n a_i t^i$$

ein Isomorphismus. Weiterhin ist

$$\Phi_B^{-1} = P_{\leq n}[t] \rightarrow K^{n+1}, p(t) = \sum_{i=0}^n a_i t^i \mapsto (a_i)_{0 \leq i \leq n} \in K^{n+1}$$

die inverse Koordinatenabbildung.

Um eine lineare Abbildung zu definieren, reicht es ihre Werte auf einer beliebigen Basis anzugeben.

Lemma 4.10: Sei V ein endlichdimensionaler Vektorraum über K mit einer Basis $B := \{v_1, \dots, v_n\}$, W ein beliebiger K -Vektorraum und $C = \{w_1, \dots, w_n\}$ eine Familie von Vektoren in W . Dann gibt es genau eine lineare Abbildung f von V nach W mit

$$f(v_i) = w_i \text{ für } 1 \leq i \leq n$$

Beweis: zu zeigen: Existenz und Eindeutigkeit

Eindeutigkeit: Seien $f, g \in L(V, W)$ mit $f(v_i) = g(v_i) = w_i, \quad \forall i \in \{1, \dots, n\}$. Sei $v \in \sum \lambda_i v_i$: Dann gilt

$$\begin{aligned}
f(v) &= f\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i f(v_i) \\
&= \sum_{i=1}^n \lambda_i w_i = \sum_{i=1}^n \lambda_i g(v_i) = g(v) \\
&\implies f(x) = g(x) \quad \forall v \in V \\
&\implies f \text{ ist eindeutig bestimmt}
\end{aligned}$$

Existenz: B Basis $\implies \text{Span } B = V \implies$ zu $v \in V$ existieren nach Satz 3.8 eindeutige Koordinaten

$$\lambda_i^v, 1 \leq i \leq n \text{ mit } v = \sum_{i=1}^n \lambda_i^v v_i$$

Definiere $f : V \rightarrow W$, $f(v) = \sum_{i=1}^n \lambda_i^v w_i \in W \implies f(v_i) = w_i$

Ist f linear?

Für jedes $\mu \in K$ gilt:

$$\begin{aligned}
\mu v &= \sum_{i=1}^n (\mu \lambda_i^v v_i) \implies f(\mu v) = f\left(\sum_{i=1}^n (\mu \lambda_i^v v_i)\right) \stackrel{\text{Def}}{=} \sum_{i=1}^n (\mu \lambda_i^v w_i) \\
&= \mu \sum_{i=1}^n \lambda_i^v w_i = \mu f(v) \quad \text{erste Eigenschaft } \checkmark
\end{aligned}$$

Sei $u = \sum_{i=1}^n \lambda_i^u v_i \implies$

$$\begin{aligned}
f(v+u) &= f\left(\sum_{i=1}^n \lambda_i^v v_i + \sum_{i=1}^n \lambda_i^u v_i\right) = f\left(\sum_{i=1}^n (\lambda_i^v + \lambda_i^u) v_i\right) \\
&= \sum_{i=1}^n (\lambda_i^v + \lambda_i^u) w_i = \sum_{i=1}^n \lambda_i^v w_i + \sum_{i=1}^n \lambda_i^u w_i \\
&= f(v) + f(u) \quad \text{zweite Eigenschaft } \checkmark \\
&\implies f \text{ linear}
\end{aligned}$$

□

4.2. Kern, Bild und Rang von linearen Abbildungen

Kern und Bild wurden bereits im Kapitel 2 behandelt. Hier behandeln wir die Konzepte im Kontext linearer Abbildungen.

Definition 4.11: Kern, Bild

Es seien V und W K -Vektorräume sowie $f : V \rightarrow W$ eine lineare Abbildung. Die Menge

$$\ker(f) := \{v \in V \mid f(v) = 0\}$$

heißt **Kern** der linearen Abbildung f .

Die Menge

$$\operatorname{im}(f) := f(V) = \{w \in W \mid \exists v \in V : w = f(v)\}$$

heißt **Bild** der linearen Abbildung f .

Beispiel 4.12:

1. Für den Isomorphismus aus Beispiel 4.9 gilt:

$$\Phi_B : K^{n+1} \rightarrow P_{\leq n}[t]$$

mit

$$\ker(\Phi_B) = \{0_{K^{n+1}}\}$$

denn nur $0 \in K^{n+1}$ wird auf das Nullpolynom abgebildet.

2. Fortsetzung von Beispiel 4.3 (Drehungsmatrizen)

Für $a = 0 = b, c, d \in \mathbb{R}, c \neq 0$ und damit

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} = \begin{pmatrix} 0 \\ cx_1 + dx_2 \end{pmatrix}$$

gilt:

$$\ker(f) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 = -\frac{dx_2}{c} \right\},$$

$$\operatorname{im}(f) = \left\{ \begin{pmatrix} 0 \\ x \end{pmatrix} \mid x \in \mathbb{R} \right\} \subset \mathbb{R}^2$$

weil $x = cx_1 + dx_2$ für jedes $x \in \mathbb{R}$ lösbar ist ($c \neq 0$).

Lemma 4.13: Es seien V und W zwei K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Dann gilt :

1. $\ker(f) \subseteq V$ ist ein Untervektorraum von V
2. $\operatorname{im}(f) \subseteq W$ ist ein Untervektorraum von W

Beweis:

zu 1) $0_V \in \ker(f) \implies \ker(f) \neq \emptyset$. Weiterhin ist $\ker(f)$ abgeschlossen bezüglich der Addition und Multiplikation, denn sei $x, \tilde{x} \in \ker(f), \lambda, \mu \in K$.

Zu zeigen: $\lambda x + \mu \tilde{x} \in \ker(f)$

$$f(\lambda x + \mu \tilde{x}) = \lambda f(x) + \mu f(\tilde{x}) = 0_W \\ \implies \text{Abgeschlossenheit in } V$$

Daraus folgt auch mit $\tilde{x} = 0_V$ die Abgeschlossenheit bezüglich der Multiplikation.

$\implies \ker(f)$ Unterraum ✓

zu 2) $f(0_V) = 0_W \in \text{im}(f) \neq \emptyset$

zu zeigen: $\text{im}(f)$ abgeschlossen bezüglich $+$ und \cdot . Seien $y, \tilde{y} \in \text{im}(f)$, $\lambda, \mu \in K$

$\implies x, \tilde{x} \in V : f(x) = y \wedge f(\tilde{x}) = \tilde{y}$. Dann gilt für $\lambda x + \mu \tilde{x}$, dass

$$f(\lambda x + \mu \tilde{x}) = \lambda f(x) + \mu f(\tilde{x}) = \lambda y + \mu \tilde{y} \\ \implies \lambda x + \mu \tilde{x} \text{ ist das Urbild von } \lambda y + \mu \tilde{y} \\ \implies \lambda y + \mu \tilde{y} \in \text{im}(f)$$

□

Definition 4.14: Rang einer Abbildung

Es seien V und W zwei K -Vektorräume und $f : V \rightarrow W$ eine lineare Abbildung. Der **Rang** der linearen Abbildung f ist definiert als

$$\text{rg}(f) := \dim_K(\text{im}(f))$$

Lemma 4.15: Es seien V und W zwei K -Vektorräume und $f \in L(V, W)$, W sei endlichdimensional. Dann gilt

1. f surjektiv $\iff \text{rg}(f) = \dim W$
2. f injektiv $\iff \dim(\ker(f)) = 0$

Beweis:

zu 1) “ \implies ”: Sei f surjektiv. Dann gilt $W = f(V) = \text{im}(f)$ und damit auch

$$\text{rg}(f) = \dim(\text{im}(f)) = \dim W$$

“ \impliedby ”: Es gilt $\text{rg}(f) = \dim(\text{im}(f)) = W$. Nach Lemma 4.13 ist $\text{im}(f)$ ein Vektorraum. Da W endlichdimensional ist, folgt mit Lemma 3.31

$$\implies \text{im}(f) = W, \quad f \text{ surjektiv}$$

zu 2) “ \implies ”: Sei f injektiv. Wir wissen $f(0_V) = 0_W$.

$$\implies \ker(f) = \{v \in V \mid f(v) = 0\} = \{0_V\} \implies \dim(\ker(f)) = 0$$

“ \impliedby ”: Es gilt $\dim(\ker(f)) = 0 \implies \ker(f) = \{0_V\}$. Damit erhält man für $x, \tilde{x} \in V$ wegen der Linearität in f :

$$f(x) = f(\tilde{x}) \iff f(x) - f(\tilde{x}) = 0 \iff f(x - \tilde{x}) = 0 \in \ker(f) \\ \iff x - \tilde{x} \in \ker(f) \iff x - \tilde{x} = 0 \iff x = \tilde{x} \quad f \text{ injektiv}$$



Satz 4.16: Dimensionsformel

Seien V und W zwei endlichdimensionale K -Vektorräume und $f \in L(V, W)$. Dann gilt

$$\dim(\ker(f)) + \dim(\operatorname{im}(f)) = \dim V$$

Beweis: Mit $r := \dim(\ker(f))$, $n := \dim V$ gilt

$$r < \infty \text{ und } n < \infty \text{ sowie } r \leq n$$

Sei $\tilde{B} := \{v_1, \dots, v_r\}$ eine Basis von $\ker(f)$. Mithilfe des Austauschsatzes von Steinitz (3.23) kann \tilde{B} zu einer Basis $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ von V ergänzt werden.

zu zeigen: $C = \{f(v_{r+1}), \dots, f(v_n)\}$ ist eine Basis von $\operatorname{im}(f)$.

1) Erzeugendensystem

f linear $\implies \operatorname{Span} C \subset \operatorname{im}(f)$. Sei $y \in \operatorname{im}(f) \implies \exists x \in V : f(x) = y$.

B ist eine Basis von $V \implies$ es existiert eine eindeutige Darstellung

$$x = \sum_{i=1}^n \lambda_i v_i$$

Mit der Linearität von f folgt:

$$f(x) = f\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i f(v_i) =$$

$$v_1, \dots, v_r \in \ker(f) \implies y = f(x) = \sum_{i=1}^n \lambda_i f(v_i)$$

$$\implies \operatorname{im}(f) \subseteq \operatorname{Span} C$$

$$\implies \operatorname{im}(f) = \operatorname{Span} C$$

✓

lineare Unabhängigkeit: Sei

$$\sum_{i=r+1}^n \lambda_i f(v_i) = 0$$

zu zeigen: $\lambda_i = 0$ für $r+1 \leq i \leq n$

$$f \text{ linear} \implies 0 = \sum_{i=r+1}^n \lambda_i f(v_i) = f\left(\sum_{i=r+1}^n \lambda_i v_i\right)$$

$$\implies \sum_{i=r+1}^n \lambda_i v_i \in \ker(f)$$

\implies es existieren eindeutige Koeffizienten $\mu_i \in K$, $1 \leq i \leq r$

$$\sum_{i=r+1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i \implies \sum_{i=r+1}^n \lambda_i v_i - \sum_{i=1}^n \mu_i v_i = 0$$

$\{v_1, \dots, v_n\}$ sind Basis von V . $\implies \mu_i = \dots = \mu_r = \lambda_{r+1} = \dots = \lambda_n = 0$. $\implies f(v_{n+1}), \dots, f(v_n)$ sind linear unabhängig $\implies C$ ist Basis von $\text{im}(f)$

$$\implies \dim(\text{im}(f)) = n - r$$

□

Beispiel 4.17: Fortsetzung von Beispiel 4.12.

zu 1)

$$\Phi_B : K^{n+1} \rightarrow P_{\leq n}[t]$$

$$\text{Es gilt } \ker(\Phi_B) = \{0\} \implies \dim(\ker(f)) = 0$$

$$\implies \dim(P_{\leq n}[t]) = n + 1$$

zu 2)

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad a = b = 0, c \neq 0, c, d \in \mathbb{R}$$

Es gilt

$$\dim(\ker(f)) = \dim\left(\left\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 = \frac{dx_2}{c}\right\}\right) = 1$$

$$\dim(\text{im}(f)) = \dim\left(\left\{\begin{pmatrix} 0 \\ x \end{pmatrix} \mid x \in \mathbb{R}\right\}\right) = 1$$

$$\dim(\ker(f)) + \dim(\text{im}(f)) = \dim(\mathbb{R}^2) = 2$$

Beispiel 4.18: Sei $V = \mathbb{R}^3$, $W = \mathbb{R}^3$ und $f : V \rightarrow W$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 - x_2 + 2x_3 \\ x_1 + x_2 + x_3 \\ 0x_1 + 3x_2 + 0x_3 \end{pmatrix}$$

gegeben. Damit ist f linear. Nach Lemma 4.10 wird das Bild $\text{im}(f)$ durch die Bilder $f(e_1)$, $f(e_2)$ und $f(e_3)$ erzeugt.

$$\implies \text{im}(f) = \text{Span}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}\right\}$$

$$\implies \text{rg}(f) = \dim(\text{im}(f)) = 2$$

Aus der Dimensionsformel folgt

$$\dim(\ker(f)) = \dim(V) - \dim(\text{im}(f)) = 3 - 2 = 1$$

$$\implies \{0_{\mathbb{R}^3}\} \subset \ker(f)$$

Lemma 4.19: Zwei endlichdimensionale K -Vektorräume V und W sind genau dann isomorph, wenn

$$\dim(V) = \dim(W)$$

Beweis:

“ \Rightarrow ”: Gilt $V \cong W$, so existiert ein Isomorphismus $f \in L(V, W)$.

Lemma 4.15:

$$\begin{aligned} \ker(f) &= \{0\} & (f \text{ injektiv}) \\ \operatorname{im}(f) &= W & (f \text{ surjektiv}) \end{aligned}$$

Satz 4.16:

$$\begin{aligned} \dim(V) &= \dim(\ker(f)) + \dim(\operatorname{im}(f)) \\ &= 0 + \dim(W) \\ &= \dim(W) \quad \checkmark \end{aligned}$$

“ \Leftarrow ”: $\dim(V) = \dim(W) < \infty$

Seien $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_n\}$ Basen von V bzw. W . Nach Lemma 4.10 gibt es genau eine Abbildung $f \in L(V, W)$ mit

$$f(v_i) = w_i, \quad 1 \leq i \leq n$$

Ist $v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \in \ker(f)$. Dann gilt

$$\begin{aligned} 0 &= f(v) = f(\lambda_1 v_1 + \dots + \lambda_n v_n) \\ &\stackrel{f \in L(V, W)}{=} \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) \\ &= \lambda_1 w_1 + \dots + \lambda_n w_n \end{aligned}$$

w_1, \dots, w_n linear unabhängig

$\Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0 \Rightarrow v = 0 \Rightarrow \ker(f) = \{0_V\} \stackrel{4.15}{\Rightarrow} f \text{ injektiv.}$

Mit der Dimensionsformel folgt:

$$\begin{aligned} \dim(W) &= \dim(V) = 0 + \dim(\operatorname{im}(f)) \\ \Rightarrow \text{mit } \operatorname{im}(f) &\subseteq W \Rightarrow \operatorname{im}(f) = W \Rightarrow f \text{ surjektiv} \end{aligned}$$

□

Was passiert bei Verknüpfungen von linearen Abbildungen?

Satz 4.20: Seien V, W und X drei endlichdimensionale K -Vektorräume sowie $f \in L(V, W)$ und $g \in L(W, X)$. Dann gilt

$$\begin{aligned} g \circ f &\in L(V, X) \text{ und} \\ \dim(\operatorname{im}(g \circ f)) &= \dim(\operatorname{im}(f)) - \dim(\operatorname{im}(f) \cap \ker(g)) \end{aligned}$$

Beweis: 1) $g \circ f \in L(V, X)$: Für $u, v \in V$ und $\lambda, \mu \in K$ erhält man:

$$\begin{aligned}
(g \circ f)(\lambda u + \mu v) &= g(f(\lambda u + \mu v)) \\
&= g(\lambda f(u) + \mu f(v)) = \lambda g(f(u)) + \mu g(f(v)) \\
&= \lambda(g \circ f)(u) + \mu(g \circ f)(v) \quad \checkmark
\end{aligned}$$

2) Betrachte $\tilde{g} = g|_{\text{im}(f)}$. Die Dimensionsformel liefert

$$\dim(\text{im}(f)) = \dim(\text{im}(\tilde{g})) + \dim(\ker(\tilde{g}))$$

Des Weiteren gilt: $\text{im}(\tilde{g}) := \{g(v) \in X \mid v \in \text{im}(f)\} = \text{im}(g \circ f)$

$$\ker(\tilde{g}) = \{v \in \text{im}(f) \mid \tilde{g}(v) = 0\} = \text{im}(f) \cap \ker(g)$$

□

Lemma 4.21: Sei K ein Körper. Die linearen Abbildungen $f : K^n \rightarrow K^m$ sind genau die Abbildungen der Form:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \rightarrow \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

mit Koeffizienten $a_{ij} \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$.

Beweis: “ \Leftarrow ”: Die Dimensionen passen aufgrund der Definitionen. Die Linearität f folgt aus den Rechengesetzen im Körper.

“ \Rightarrow ”: Sei $f \in L(K^n, K^m)$. Zu zeigen: f ist in angegebener Form darzustellen.

Beobachtung: Wenn f so darstellbar ist, haben alle Bilder der Standardbasis e_1, \dots, e_n folgende Form:

$$f(e_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \in K^m$$

Deswegen definieren wir

$$\begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} := f(e_i) \in K^m$$

Jetzt definieren wir $g \in L(K^n, K^m)$ durch

$$g\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}$$

Per Konstruktion gilt $f(e_i) = g(e_i)$.

Mit Lemma 4.10 folgt $f = g$.

□

Dieses Resultat motiviert das nächste Kapitel:

5. Matrizen

James Sylvester (brit. Mathematiker, 1814 - 1897) erfand den Begriff der Matrix im Jahr 1850. Die im Folgenden definierte Matrixoperationen führte Arthur Cayley (brit. Mathematiker, 1821 - 1895) im Jahr 1858 ein.

5.1. Definitionen und Basisoperationen

Wir nehmen für dieses Kapitel an: R ist ein Ring mit $1 \neq 0$.

Definition 5.1: Matrix

Sei $(R, +, \cdot)$ ein kommutativer Ring mit Eins und seien $n, m \in \mathbb{N}_0$. Ein rechteckiges Schema der Form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in R$ für $1 \leq i \leq m$ und $1 \leq j \leq n$ heißt **$(m \times n)$ Matrix** mit den Einträgen in R . Die Einträge nennt man auch Koeffizienten. Die Menge aller $(m \times n)$ Matrizen nennt man $R^{m \times n}$.

Bemerkungen:

- Rein formal erlaubt diese Definition $n = 0$ oder $m = 0$. Dann erhält man Matrizen der Form $0 \times n$, $m \times 0$ oder 0×0 .

Diese leeren Matrizen werden in manchen Beweisen aus technischen Gründen benötigt. In der Regel gilt aber $n, m \geq 1$.

- Die Nullmatrix in $R^{m \times n}$ ist die Matrix, bei der alle Einträge gleich 0 sind. Sie wird mit $0^{m \times n}$ bezeichnet
- Ist $m = n$, so nennt man $A \in R^{m \times n}$ quadratisch bzw. eine quadratische Matrix
- Ist $A \in R^{n \times n}$ heißen die Einträge a_{jj} für $1 \leq j \leq n$ **Diagonaleinträge** von

$$A = \begin{pmatrix} a_{11} & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & a_{nn} \end{pmatrix}$$

- Die Kronecker-Delta Funktion δ_{ij} für $i \in I$ und $j \in J$, I und J sind Indexmengen, ist benannt nach Leopold Kronecker (1823 - 1891) und gegeben durch

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases}$$

Damit definiert man die Einheitsmatrix $I_n \in R^{n \times n}$ durch

$$I_n := [\delta_{ij}] = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

Man verwendet I , wenn n aus dem Kontext klar ist.

- Die i -te Zeile von $A \in R^{m \times n}$ ist

$$(a_{i1}, \dots, a_{in}) \in R^{1 \times n} \text{ für } i = 1, \dots, m \text{ ist ein Zeilenvektor}$$

Die j -te Spalte von $A \in R^{m \times n}$ ist

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in R^{m \times 1} \text{ für } j = 1, \dots, n \text{ ist ein Spaltenvektor}$$

Diese sind selbst wieder Matrizen.

- Sind $m_1, m_2, n_1, n_2 \in \mathbb{N}_0$ und $A_{ij} \in R^{m_i, n_j}$ für $i, j = 1, 2$ gegeben, definieren diese eine sogenannte **Blockmatrix** der Form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in R^{m_1+m_2, n_1+n_2}$$

Beispiel 5.2: Für

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & 6 \end{pmatrix} \in \mathbb{Z}^{2,3} \text{ und } B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix} \in \mathbb{Z}^{3,2}$$

ist $a_{23} = 6$, $(1, -2, 3) \in \mathbb{Z}^{1,3}$ die erste Zeile von A und $b_{22} = 1$ und

$$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{Z}^{3,1}$$

die zweite Spalte von B . Keine dieser Matrizen ist quadratisch.

Definition 5.3: Addition von Matrizen

Seien $A, B \in R^{m,n}$ zwei Matrizen. Dann ist $C = A + B \in R^{m,n}$ definiert durch

$$C := A + B = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Die Addition in $R^{m,n}$ erfolgt also komponentenweise basierend auf der Addition in R .

Achtung: Die Addition ist nur für Matrizen gleicher Größe / Dimension definiert.

Definition 5.4: Multiplikation einer Matrix mit einem Skalar

Sei $A \in R^{m,n}$ eine Matrix und $\lambda \in R$. Dann ist $C = \lambda \cdot A \in R^{m,n}$ definiert durch

$$C = \lambda A = \lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$$

Die Multiplikation einer Matrix mit einem Skalar aus R erfolgt komponentenweise basierend auf der Multiplikation in R .

Definition 5.5: Multiplikation von Matrizen

Seien $A \in R^{m,n}$ und $B \in R^{n,l}$ zwei Matrizen. Dann ist $C = A \cdot B \in R^{m,l}$ definiert durch

$$C := A \cdot B = \begin{pmatrix} c_{11} & \cdots & c_{1l} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{ml} \end{pmatrix} \quad \text{mit} \quad c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj}$$

für $i = 1, \dots, m$ und $j = 1, \dots, l$

Bemerkung: Um das Produkt $A \cdot B$ berechnen zu können **muss** die Anzahl der Spalten von A gleich der Anzahl von Zeilen in B sein.

Merkregel:

$$c_{ij} = i\text{-te Zeile von } A \text{ mal } j\text{-te Spalte von } B$$

Oder Zeile “mal” Spalte. (“mal” $\hat{=}$ Skalarprodukt)

Beispiel 5.6: Für die Matrizen

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & 6 \end{pmatrix} \in \mathbb{Z}^{2,3}, \quad B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix} \in \mathbb{Z}^{3,2} \quad \text{und} \quad C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \in \mathbb{Z}^{3,2}$$

gilt:

$A + B$ geht nicht

$$B + C = \begin{pmatrix} 0 & 2 \\ 3 & 5 \\ 6 & 5 \end{pmatrix} \in \mathbb{Z}^{3,2} \quad 3 \cdot A = \begin{pmatrix} 3 & -6 & 9 \\ -12 & 15 & 18 \end{pmatrix} \in \mathbb{Z}^{2,3}$$

$$A \cdot B = \begin{pmatrix} 2 & -5 \\ 10 & -1 \end{pmatrix} \in \mathbb{Z}^{2,2} \quad B \cdot A = \begin{pmatrix} -1 & 2 & -3 \\ -4 & 5 & 6 \\ 5 & -7 & -3 \end{pmatrix} \in \mathbb{Z}^{3,3}$$

Also $A \cdot B$ und $B \cdot A$ sind für dieses Beispiele definiert, aber $A \cdot B \neq B \cdot A$. Damit ist die Multiplikation von Matrizen nicht kommutativ. Dies gilt auch für gleichgroße Matrizen:

$$D = \begin{pmatrix} 1 & -2 \\ -4 & 5 \end{pmatrix} \in \mathbb{Z}^{2,2}, \quad E = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{Z}^{2,2} \implies$$

$$D \cdot E = \begin{pmatrix} -1 & -2 \\ 4 & 5 \end{pmatrix} \neq \begin{pmatrix} -1 & 2 \\ -4 & 5 \end{pmatrix} = E \cdot D$$

Lemma 5.7: Für $A, \tilde{A} \in R^{m,m}$, $B, \tilde{B} \in R^{m,l}$, $C \in R^{l,k}$ sowie $\lambda, \mu \in R$ gelten

- Assoziativgesetze, d.h.

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C \quad \text{und} \quad (\lambda\mu) \cdot A = \lambda \cdot (\mu \cdot A)$$

- Distributivitätsgesetze, d.h.

$$(A + \tilde{A}) \cdot B = A \cdot B + \tilde{A} \cdot B$$

$$A \cdot (B + \tilde{B}) = A \cdot B + A \cdot \tilde{B}$$

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

$$\lambda \cdot (A + \tilde{A}) = \lambda \cdot A + \lambda \cdot \tilde{A}$$

$$\lambda \cdot (A \cdot B) = A \cdot (\lambda \cdot B)$$

- und mit

$$I_n = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} \in R^{n,n} \quad \text{bzw.} \quad I_m \in R^{m,m}$$

gilt

$$I_n \cdot A = A = A \cdot I_m$$

Beweis: Nachrechnen

□

Definition 5.8: Transposition von Matrizen

Sei

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

eine Matrix. Dann ist die zu A **transponierte** Matrix A^\top definiert durch

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

Beispiel 5.9: Die zu

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & 6 \end{pmatrix} \in \mathbb{Z}^{2,3}$$

transponierte Matrix ist

$$A^\top = \begin{pmatrix} 1 & -4 \\ -2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathbb{Z}^{3,2}$$

Lemma 5.10: Für $A, \tilde{A} \in R^{m,n}$, $B \in R^{n,l}$ und $\lambda \in R$ gilt

1. $(A^\top)^\top = A$
2. $(A + \tilde{A})^\top = A^\top + \tilde{A}^\top$
3. $(\lambda \cdot A)^\top = \lambda \cdot A^\top$
4. $\underbrace{(A \cdot B)^\top}_{\substack{\in R^{m,l} \\ \in R^{l,m}}} = \underbrace{B^\top}_{\in R^{l,n}} \cdot \underbrace{A^\top}_{\in R^{n,m}}$

Beweis: Nachrechnen.

□

Zur Notation:

$$R^n, v \in R^n, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in R^{n,1}$$

d.h. v wird immer als Spaltenvektor interpretiert.

Für Zeilenvektor:

$$w \in R^{1,n} \text{ gilt } w^\top \in R^{n,1} \cong R^n$$

Beobachtung: Alle Operationen, die wir für Matrizen definiert haben, sind konsistent mit den Vektoroperationen, wenn diese im obigen Sinn als Matrizen interpretiert werden.

5.2. Matrizengruppen und -ringe

Lemma 5.11: Mit den angegebenen Rechenregeln gilt

1. $(R^{m,n}, +)$ ist eine kommutative Gruppe mit dem neutralen Element $0 \in R^{m,n}$, d.h. der Nullmatrix und dem zu $A \in R^{m,n}$ inversen Element $-A = (-a_{ij}) \in R^{m,n}$. Man schreibt statt $A + (-B) = A - B$.
2. Ist R ein Körper, so ist $R^{m,n}$ ein R -Vektorraum der Dimension $m \cdot n$.

Beweis:

1) Folgt durch nachrechnen unter Ausnutzung der Eigenschaften von R .

2) Aufgrund der Rechenregeln ist $R^{m,n}$ abgeschlossen bezüglich

$$+ : R^{m,n} \times R^{m,n} \rightarrow R^{m,n} \quad \text{und} \quad \cdot : R \times R^{m,n} \rightarrow R^{m,n}$$

Zur Dimension: $A \in R^{m,n}$ kann dargestellt werden durch

$$a_{11} \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} + a_{12} \cdot \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} + \dots + a_{mn} \cdot \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

\Rightarrow Erzeugendensystem + linear unabhängig \Rightarrow Basis

□

Aufgrund der Eigenschaft 2 aus Lemma 5.10 ist die Transposition von Matrizen ein Homomorphismus (vgl. Definition 2.11) der Gruppen $(R^{m,n}, +)$ und $(R^{n,m}, +)$

Lemma 5.12: Sei $n \in \mathbb{N}$. Die Menge der quadratischen Matrizen $R^{n,n}$, d.h. $(R^{n,n}, +, \cdot)$, ist ein Ring mit Eins, welche durch die Einheitsmatrix I_n gegeben ist. Dieser Ring ist nur für $n = 1$ kommutativ.

Beweis: Lemma 5.11: $(R^{n,n}, +)$ ist eine kommutative Gruppe mit neutralem Element $0 \in R^{n,n}$. Lemma 5.7: Assoziativität, Distributivitätsgesetze, Einselement I_n . \Rightarrow Ring mit 1 ✓

$n = 1$: Kommutativität folgt aus Eigenschaften von R . Für $n = 2$ gilt mit

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \text{ dass}$$

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

□

Aus dem Beispiel im letzten Satz folgt auch, dass $A, B \in R^{n,n}$ mit

- $A \neq 0 \in R^{n,n}, B \neq 0 \in R^{n,n}$
- $A \cdot B = 0$

existiert. Damit besitzt $R^{n,n}$ sogenannte nichttriviale Nullteiler. Mit $R = \mathbb{R}$ gilt dies auch wenn R ein Körper ist.

Weitere wichtige Eigenschaft:

Invertierbarkeit bezüglich der Multiplikation!

Frage: Gibt es für jede Matrix $A \in R^{n,n}$ eine Matrix A^{-1} , so dass $A \cdot A^{-1} = I_n = A^{-1} \cdot A$? Wenn dies gilt, dann müsste A^{-1} existieren, so dass für

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ gilt } A \cdot A^{-1} = I = A^{-1} \cdot A$$

$$A \cdot A^{-1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Für das erste Element der Matrix folgt

$$0 \cdot a + 0 \cdot c = \underbrace{0 \neq 1}_{\text{Ring mit } 0 \neq 1} \quad \Downarrow$$

Folgerung: Nicht alle quadratischen Matrizen sind invertierbar.

Beispiel 5.13: Die Matrix

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \in \mathbb{Z}^{2,2}$$

ist damit über $R = \mathbb{Z}$ nicht invertierbar. Es gilt

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} \end{pmatrix}}_{A^{-1} \in \mathbb{Q}^{2,2}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

Ist $A \in \mathbb{Q}^{2,2}$, dann ist A invertierbar. Also: Invertierbarkeit hängt vom Ring R ab!

Lemma 5.14: Sind $A, B \in R^{n,n}$ zwei invertierbare Matrizen. D.h. es existieren $A^{-1}, B^{-1} \in R^{n,n}$ mit $A \cdot A^{-1} = I = A^{-1} \cdot A$ und $B \cdot B^{-1} = I = B^{-1} \cdot B$, so gilt:

- $A \cdot B$ ist invertierbar und es ist

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

- A^\top ist invertierbar und es ist $(A^\top)^{-1} = (A^{-1})^\top = A^{-\top}$

Beweis:

1. Aussage: Folgt aus der allgemeinen Aussage für Ringe mit Eins im Satz 2.13.
2. Aussage: Regel 4 aus Lemma 5.10:

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

□

Lemma 5.15: Die Menge $GL_n(R) := \{A \in R^{n,n} \mid A \text{ invertierbar}\}$ bilden mit der Matrixmultiplikation eine Gruppe.

Beweis: Lemma 5.14, 1. Aussage: Abgeschlossenheit bezüglich \cdot . Lemma 5.7: Assoziativität + neutrale Element I_n . Mit $(A^{-1})^{-1} = A$ ist auch A^{-1} invertierbar $\implies A^{-1} \in GL_n(R)$.

□

Hinweis: $GL_n(R)$ = General linear Group

Matrixmultiplikation: $A \in R^{m,n}, v \in R^{n,1} = R^n$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix} = A \cdot v$$

definiert damit auch die Matrix-Vektor-Multiplikation.

Diese Beobachtung motiviert:

5.3. Matrizen und lineare Abbildungen

Beispiel 5.16: Fortsetzung von Beispiel 4.3.

Die lineare Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$,

$$y = f(x) = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}$$

mit $a, b, c, d \in \mathbb{R}$ wird beschrieben durch

$$y = f(x) = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{\in R^{2,2}} \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\in R^{2,1}} = A \cdot x \quad \text{mit } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Die Kombination mit Lemma 4.21 liefert:

Satz 5.17: Sei K ein Körper. Zu jeder linearen Abbildung $f : K^n \rightarrow K^m$, $y = f(x)$, existiert eine Matrix $A \in K^{m,n}$, so dass gilt

$$y = f(x) = A \cdot x$$

Beweis: Nach Lemma 4.21 besitzt jede lineare Abbildung $f : K^n \rightarrow K^m$ die Form

$$y = f(x) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

mit Koeffizienten $a_{ij} \in K, i = 1, \dots, m, j = 1, \dots, n$. Mit den Rechenregeln für Matrizen inklusive dem Spezialfall der Vektoren folgt

$$y = f(x) = A \cdot x \quad \text{mit} \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

□

Was gilt für allgemeine Basen?

Satz 5.18: Seien V und W zwei K -Vektorräume mit den Basen $B = \{v_1, \dots, v_n\}$ von V und $C = \{w_1, \dots, w_m\}$ von W und $f : V \rightarrow W$ eine lineare Abbildung. Dann gibt es eine eindeutig bestimmte Matrix $A_f^{B,C} = (a_{ij}) \in K^{m,n}$, so dass

$$W \ni f(v_j) = \sum_{i=1}^m a_{ij} w_i \quad j = 1, \dots, n$$

gilt. Die Abbildung $F : L(V, W) \rightarrow K^{m,n}$ mit $F(f) = A_f^{B,C}$ ist ein Isomorphismus zwischen zwei K -Vektorräumen.

Beweis: Da C eine Basis von W ist, besitzt jedes $w \in W$ eine eindeutige Darstellung als Linearkombination der Vektoren $\{w_1, \dots, w_m\}$. Damit besitzt auch für jedes $f \in L(V, W)$ die Vektoren $f(v_1), \dots, f(v_n)$ eine eindeutige Darstellung. Die Koeffizienten dieser Linearkombinationen bestimmen eindeutig die Matrix $A_f^{B,C}$.

zu zeigen: F ist ein Isomorphismus.

F ist linear: Seien $f, g \in L(V, W)$ mit den zugehörigen Matrizen $A_f^{B,C}$ bzw. $A_g^{B,C}$ mit

$$A_f^{B,C} = (a_{ij}), A_g^{B,C} = (b_{ij}), A_f^{B,C} \in K^{m,n}, A_g^{B,C} \in K^{m,n}$$

Für $\lambda, \mu \in K$ gilt

$$\begin{aligned} (\lambda f + \mu g)(v_j) &= \lambda f(v_j) + \mu g(v_j) = \lambda \sum_{i=1}^m a_{ij} w_i + \mu \sum_{i=1}^m b_{ij} w_i \\ &= \sum_{i=1}^m \underbrace{(\lambda a_{ij} + \mu b_{ij})}_{\text{Matrix für } \lambda f + \mu g} w_i \end{aligned}$$

Also ist

$$F(\lambda f + \mu g) = \lambda F(f) + \mu F(g)$$

$\Rightarrow F$ linear

noch zu zeigen: F ist bijektiv

Sei $f \in \ker(F)$, dann $F(f) = 0 \in K^{m,n}$. Also ist $f(v_j) = 0$ für $j = 1, \dots, n$. Da $\{v_1, \dots, v_n\}$ eine Basis von V ist, folgt für alle $v \in V$:

$$f(v) = f\left(\sum_{j=1}^n \lambda_j v_j\right) \stackrel{f \text{ linear}}{=} \sum_{j=1}^n \lambda_j \underbrace{f(v_j)}_{0 \in W} = 0 \in W$$

$$\Rightarrow f = 0 \in L(V, W) \Rightarrow \ker(F) = \{0\} \stackrel{\text{Lemma 4.15}}{\Rightarrow} F \text{ injektiv}$$

Sei $A \in K^{m,n}$ beliebig gewählt. zu zeigen: $\exists f \in L(V, W) : F(f) = A$.

Lemma 4.10: Es existiert genau eine lineare Abbildung $f \in L(V, W)$ mit

Zwischenüberlegung

$$\{\tilde{w}_1, \dots, \tilde{w}_n\} \exists! f : f(v_j) = \tilde{w}_j \quad j = 1, \dots, n$$

$$\Rightarrow \tilde{w}_j = \sum_{i=1}^m a_{ij} w_i, \quad j = 1, \dots, n \quad \text{mit} \quad A = (a_{ij}) \in K^{m,n}$$

$$f(v_j) = \sum_{i=1}^m a_{ij} w_i \quad j = 1, \dots, n$$

Für diese f gilt dann $F(f) = A$

$$\Rightarrow F \text{ surjektiv}$$

□

Bemerkungen:

- Im letzten Satz kann man **nicht**

$$f(v_j) = w_j$$

fordern, da $n \neq m$ gelten kann.

- Wir hatten schon: V K -Vektorraum, $\dim(V) = n \in \mathbb{N}$

$$\text{dann } V \cong K^n$$

Mit Satz 5.18 folgt jetzt:

$$L(V, W) \cong K^{n,m}$$

- Aus Satz 4.19 und Lemma 5.11 mit dem Körper K folgt

$$\dim(L(V, W)) = \dim(K^{m,n})$$

Definition 5.19: Matrixdarstellung

Seien V und W zwei K -Vektorräume mit den Basen $B = \{v_1, \dots, v_n\}$ von V und $C = \{w_1, \dots, w_m\}$ von W sowie $f \in L(V, W)$. Die durch Satz 5.18 eindeutig bestimmte Matrix $A_f^{B,C} \in K^{m,n}$ heißt **Matrixdarstellung** oder die **darstellende Matrix** von $f \in L(V, W)$ bezüglich der Basen B und C .

Beispiel 5.20: Fortsetzung von Beispiel 4.3. Die lineare Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ einer Drehung im \mathbb{R}^2 für einen gegebenen Winkel $\varphi \in \mathbb{R}$ bezüglich der Standardbasis $S = \{e_1, e_2\}$ ist eindeutig durch die Darstellungsmatrix

$$A_f^{S,S} = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

bestimmt.

Beispiel 5.21: Fortsetzung Beispiel 4.18. Die lineare Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ besitzt für die Standardbasis die Darstellungsmatrix

$$A_f^{S,S} = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 1 & 2 \\ 0 & 3 & 0 \end{pmatrix} \rightarrow \text{rg}(A) = 2$$

für $S = \{e_1, e_2, e_3\}$. Beispiel 4.18: $\text{rg}(f) = \dim(\text{im}(f))$.

Definition 5.22: Spaltenrang, Zeilenrang

Sei $A \in K^{m,n}$, K ein Körper, gegeben. Die Maximalzahl linear unabhängiger Spaltenvektoren von A heißt **Spaltenrang** von A und wird mit $\text{rg}_S(A)$ bezeichnet. Die Maximalzahl der linear unabhängigen Zeilenvektoren von A heißt **Zeilenrang** von A und wird mit $\text{rg}_Z(A)$ bezeichnet.

$$\Rightarrow A_f^{S,S} = \begin{pmatrix} 1 & -1 & 2 \\ 1 & 1 & 2 \\ 0 & 3 & 0 \end{pmatrix} \quad \text{rg}_S(A) = 2$$

Satz 5.23: Seien V, W zwei K -Vektorräume mit den Basen $B = \{v_1, \dots, v_n\}$ von V und $C = \{w_1, \dots, w_m\}$ von W sowie $f \in L(V, W)$ mit der Darstellungsmatrix $A_f^{B,C} = (a_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in K^{m,n}$. Dann gilt die Gleichheit

$$\text{rg}(f) = \text{rg}_S(A_f^{B,C})$$

Beweis: Es gilt

$$f(v_j) = \sum_{i=1}^m a_{ij} w_i \quad j = 1, \dots, n$$

Damit sind die Spaltenvektoren

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m \quad j = 1, \dots, n$$

die Koordinaten von $f(v_j)$ bezüglich der Basis C . Aufgrund der Isomorphie von W und K^m können damit diese Spaltenvektoren als Erzeugendensystem von $\text{im}(f)$ interpretiert werden. Damit folgt

$$\operatorname{rg}(f) = \dim(\operatorname{im}(f)) = \dim \left(\operatorname{Span} \left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right) \right) = \operatorname{rg}_S(A)$$

□

Beispiel 5.24: Fortsetzung von Beispiel 5.21

Für die Darstellungsmatrix

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

gilt $\operatorname{rg}(f) = 2 = \operatorname{rg}_S(A_f^{S,S})$.

Graphische Darstellung der Ergebnisse: V, W sind K -Vektorräume, $B = \{v_1, \dots, v_n\}$ ist Basis von V , $C = \{w_1, \dots, w_m\}$ ist Basis von W . Für $f \in L(V, W)$ gilt

Mit Satz 5.18 und dem Struktursatz 4.8 können die Ergebnisse mit folgenden **kommutativen Diagramme** zusammengefasst werden:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \Phi_B \uparrow & & \uparrow \Phi_C \\ K^n & \xrightarrow[A_f^{B,C}]{} & K^m \end{array}$$

Solche kommutativen Diagramme sind sehr wichtige Hilfsmittel! Man erhält daraus folgenden Aussagen:

- Für die gegebenen Basen B und C gibt es zwei Wege von V nach K^m :

$$\Phi_C \circ f = A_f^{B,C} \circ \Phi_B$$

D.h.:

$$\Phi_C(f(v)) = A_f^{B,C} \cdot \Phi_B(v) \quad \forall v \in V$$

- Mithilfe der inversen Abbildungen kann man auch andere Wege gehen, z.B.:

$$f = \Phi_C^{-1} \circ A_f^{B,C} \cdot \Phi_B$$

Was passiert bei der Verknüpfung von linearen Abbildungen mit den Darstellungsmatrizen?

Satz 5.25: Seien V, W und X drei endlichdimensionale K -Vektorräume mit den Basen B, C und D . Für gegebene $f \in L(V, W)$ und $g \in L(W, X)$ gilt

$$A_{g \circ f}^{B,D} = A_g^{C,D} \cdot A_f^{B,C}$$

Beweis: Satz 4.20: $g \circ f \in L(V, X)$. Seien $B = \{v_1, \dots, v_n\}$, $C = \{w_1, \dots, w_m\}$, $D = \{x_1, \dots, x_s\}$ und die Darstellungsmatrizen gegeben durch

$$A_f^{B,C} = (a_{ij})_{\substack{j=1,\dots,n \text{ Spalten} \\ i=1,\dots,m \text{ Zeilen}}} \text{ sowie } A_g^{C,D} = (b_{ij})_{\substack{j=1,\dots,m \\ i=1,\dots,s}}$$

Dann folgt für $v_j, j = 1, \dots, n$

$$\begin{aligned} (g \circ f)(v_j) &= g\left(\underbrace{\sum_{k=1}^m a_{kj} w_k}_{= f(v_j)}\right) = \sum_{k=1}^m a_{kj} g(w_k) \\ &= \sum_{k=1}^m a_{kj} \sum_{i=1}^s b_{ik} x_i = \sum_{i=1}^s \sum_{k=1}^m a_{kj} b_{ik} x_i \\ &= \sum_{i=1}^s \left(\sum_{k=1}^m b_{ik} a_{kj} \right) x_i \end{aligned}$$

D.h. für die Darstellungsmatrix

$$A_{g \circ f}^{B,C} = (c_{ij})_{\substack{j=1,\dots,n \\ i=1,\dots,s}}$$

$$\text{gilt } A_{g \circ f}^{B,D} = A_f^{C,D} \cdot A_f^{B,C}.$$

□

5.4. Basiswechsel

Seien V und W zwei K -Vektorräume mit $\dim(V) = n \in \mathbb{N}$ und $\dim(W) = m \in \mathbb{N}$ mit $B = \{v_1, \dots, v_n\}$ als Basis von V . Nach dem Struktursatz für Vektorräume Satz 4.8 gilt

$$\Phi_B : K^n \rightarrow V, (x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i v_i$$

Die Kunst der linearen Algebra besteht darin, zu einem gegebenen Problem, welches durch eine lineare Abbildung beschreiben werden kann, geschickt Basen zu wählen, so dass die Darstellung möglichst einfach ist.

Beispiel 5.26: Sei $V = \mathbb{R}^2$. Für die Standardbasis $C = \{e_1, e_2\}$ erhält man

$$\begin{aligned} \Phi_C = \mathbb{R}^2 \rightarrow V, (x_1, x_2) &\mapsto \sum_{i=1}^2 x_i e_i \\ &= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

Die Basis $B = \{e_1, e_1 + e_2\}$ liefert

$$\begin{aligned}\Phi_B : \mathbb{R}^2 &\rightarrow V, (x_1, x_2) \mapsto \sum_{i=1}^2 x_i v_i \\ &= x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\end{aligned}$$

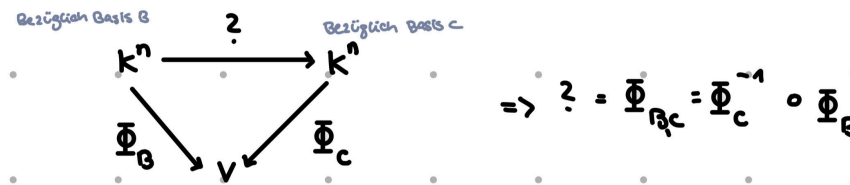
V = Ebene die durch \mathbb{R}^2 beschrieben wird.

Frage: Darstellung von F in den unterschiedlichen Basen.

Lemma 5.27: Der Basiswechsel zwischen zwei Basen B und C des K -Vektorraums V mit $\dim(V) = n \in \mathbb{N}$ ist der Isomorphismus

$$\Phi_{B,C} = \Phi_C^{-1} \circ \Phi_B$$

Beweis: Basierend auf dem Diagramm



Es fehlt noch die Isomorphieeigenschaft. Nach dem Struktursatz 4.8 gilt: Φ_B, Φ_C sind Isomorphismen. Dann ist auch

- Φ_C^{-1} ein Isomorphismus
- $\Phi_C^{-1} \circ \Phi_B$ ein Isomorphismus

siehe Satz 1.38.

□

Beobachtung: Sei V ein K -Vektorraum mit Basis $B = \{v_1, \dots, v_n\}$ und $C = \{e_1, \dots, e_n\}$ die Standardbasis. Dann ist

- $\Phi_C : K^n \rightarrow V$ die Identität, d.h. $A_{\Phi_C}^{C,C} = I_n$ (siehe oben)
- $\Phi_{B,C} = \Phi_C^{-1} \circ \Phi_B = \Phi_B \implies \Phi_{B,C}(e_i) = \Phi_B(e_i) = v_i$

Die zum Basiswechsel gehörende Matrix $A_{\Phi_{B,C}}^{B,C}$ von einer gegebenen Basis $B = \{v_1, \dots, v_n\}$ zur Standardbasis ist die Matrix, deren Spalten den Basisvektoren von B entsprechen, d.h.

$$A_{\Phi_{B,C}}^{B,C} = (v_1 \ v_2 \ \dots \ v_n) \quad v_i \text{ ist immer ein Spaltenvektor}$$

Beispiel 5.28: Sei $V = \mathbb{R}^2$ und $B = \{e_1, e_1 + e_2\}$ sowie $C = \{e_1, e_2\}$. Damit ist

$$A_{\Phi_{B,C}}^{B,C} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \text{Gl}_2(\mathbb{R}) \quad \text{mit}$$

$$\left(A_{\Phi_{B,C}}^{B,C}\right)^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Man erhält z.B.:

$$\Phi_{B,C} \left(\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \right) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$$

Jetzt für lineare Abbildungen:

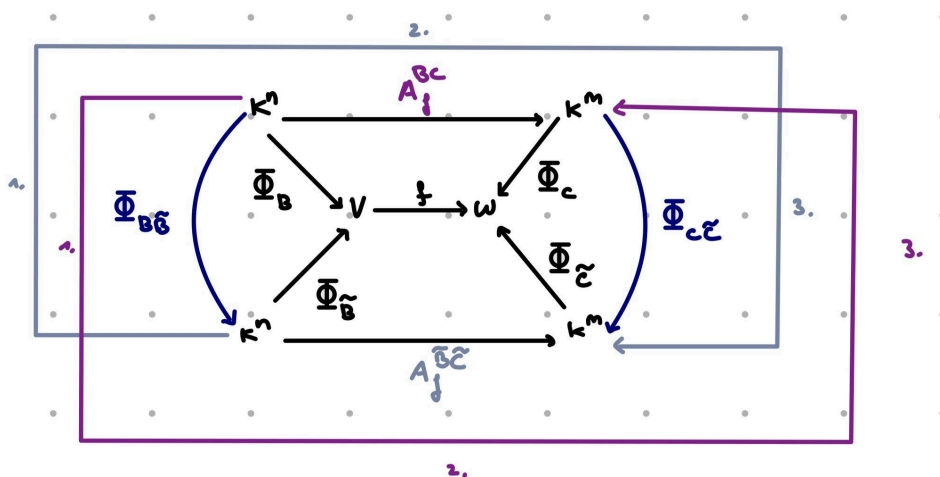
Lemma 5.29: Seien V und W K -Vektorräume mit $\dim(V) = n, \dim(W) = m$ sowie $f \in L(V, W)$. Weiter seien B, \tilde{B} zwei Basen von V und C, \tilde{C} zwei Basen von W . Dann gilt für die Darstellungsmatrizen, dass

$$A_f^{\tilde{B}, \tilde{C}} = \Phi_{C, \tilde{C}} \cdot A_f^{B, C} \cdot \Phi_{B, \tilde{B}}^{-1}$$

bzw.

$$A_f^{B, C} = \Phi_{C, \tilde{C}}^{-1} \cdot A_f^{\tilde{B}, \tilde{C}} \cdot \Phi_{B, \tilde{B}}$$

Beweis: Dies folgt aus dem kommutativen Diagramm.



Beispiel 5.30: Sei die Darstellungsmatrix der linearen Abbildung $f : V \rightarrow W$ mit $V = \mathbb{R}^3$ und $W = \mathbb{R}^2$ bezüglich der Standardbasen gegeben durch

$$A_f^{B,C} = \begin{pmatrix} 1 & 1 & -2 \\ -6 & 3 & 3 \end{pmatrix}$$

Wir suchen jetzt die Darstellung bezüglich der Basen

$$\tilde{B} = \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \quad \tilde{C} = \left\{ \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$$

Mit $v_1 = -1 \cdot e_3$, $v_2 = e_2$, $v_3 = e_1 + e_2 + e_3$ folgt aus der Linearität von f , dass

$$f(v_1) = f(-1 \cdot e_3) = -f(e_3) = w_1$$

$$f(v_2) = f(e_2) = w_2$$

$$f(v_3) = f(e_1) + f(e_2) + f(e_3) = 0$$

Also wird f bezüglich der Basen \tilde{B} und \tilde{C} durch

$$A_f^{\tilde{B}, \tilde{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

beschrieben.

Fazit: Der richtige Basiswechsel bringt die Abbildung f in sehr einfache Gestalt. Man kann zeigen: Für V, W endlichdimensional und $f \in L(V, W)$ mit $\dim(\text{im}(f)) = r \in \mathbb{N}$, $c = \dim(W) - r$, $a = \dim(V) - r$ existieren Basen B von V und C von W , so dass

$$A_f^{B,C} = \begin{pmatrix} I_r & 0_{r \times b} \\ 0_{a \times r} & 0_{a \times b} \end{pmatrix} \quad \text{z.B. Mehrmann Satz 10.24}$$

5.5. Elementarumformung von Matrizen

Für eine gegebene Matrix $A = (a_{ij})_{j=1, \dots, m}^{i=1, \dots, n} \in K^{m,n}$ bezeichnen wir die Spaltenvektoren mit

$$a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \text{mit } j = 1, \dots, n \implies a_j \in K^{m,1} \cong K^m$$

und die Zeilenvektoren mit

$$a^i = (a_{i1}, a_{i2}, \dots, a_{in}) \quad \text{mit } i = 1, \dots, m \implies a^i \in K^{1,n}$$

Die drei elementaren Spaltenumformungen sind gegeben durch

- (S1) Vertauschen zweier verschiedener Spalten a_j und a_k für $j, k \in \{1, \dots, n\}$, $j \neq k$
- (S2) Multiplikation einer Spalte a_j , $j \in \{1, \dots, n\}$ mit Skalar $\lambda \in K$, $\lambda \neq 0$
- (S3) Addition einer Spalte a_j zu einer anderen Spalte a_k für $j, k \in \{1, \dots, n\}$, $j \neq k$

Die drei elementaren Zeilenumformungen sind gegeben durch.

- (Z1) Vertauschen zweier Zeilen a^i und a^k für $i, k \in \{1, \dots, m\}$, $i \neq k$
- (Z2) Multiplikation einer Zeile a^i , $i \in \{1, \dots, m\}$ mit einem Skalar $\lambda \in K$, $\lambda \neq 0$
- (Z3) Addition einer Zeile a^i zu einer anderen Zeile a^k für $i, k \in \{1, \dots, m\}$, $i \neq k$

Lemma 5.31: Es bezeichne $A = (a_{ij})_{j=1, \dots, m}^{i=1, \dots, n} \in K^{m,n}$. Dann ändern sich der Spaltenrang $\text{rg}_S(A)$ und der Zeilenrang $\text{rg}_Z(A)$ von A bei elementaren Umformungen (S1) - (S3) bzw. (Z1) - (Z3) nicht.

Beweis: Hier: Spaltenrang

Nach Definition gilt:

$$\text{rg}_S(A) = \dim(\text{Span}(a_1, \dots, a_m))$$

Dann folgt aus

(S1)

$$\text{Span}(a_1, \dots, a_j, \dots, a_k, \dots, a_n) = \text{Span}(a_1, \dots, a_k, \dots, a_j, \dots, a_n) \\ j, k \in \{1, \dots, n\}, j \neq k, \text{ o.B.d.A. } j < k$$

(S2)

$$\text{Span}(a_1, \dots, a_j, \dots, a_n) = \text{Span}(a_1, \dots, \lambda a_j, \dots, a_n) \\ j \in \{1, \dots, n\}, \lambda \neq 0$$

(S3)

$$\text{Span}(a_1, \dots, a_j, \dots, a_k, \dots, a_n) = \text{Span}(a_1, \dots, a_j, \dots, a_k + a_j, \dots, a_n) \\ j, k \in \{1, \dots, n\}, j \neq k, \text{ o.B.d.A. } j < k$$

die Invarianz von $\text{rg}_S(A)$ für (S1) - (S3)

2. Zeilenumformungen

Wir betrachten die Koordinaten der Spaltenvektoren a_1, \dots, a_n bezüglich der Standardbasis $B = \{e_1, \dots, e_m\}$ bezüglich der Abbildung in den K^m welche durch A gegeben ist. Durch Vertauschen der Einheitsvektoren $e_i, e_k, i, k \in \{1, \dots, n\}, i \neq k$ erhalten wir eine neue Basis $\tilde{B} = \{e_1, \dots, e_k, \dots, e_i, \dots, e_m\}$, o.B.d.A. $i < k$ für den K^m und die neuen Koordinaten

$$\tilde{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{k1} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \tilde{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{kn} \\ \vdots \\ a_{in} \\ \vdots \\ a_{mn} \end{pmatrix}$$

aus den ursprünglichen Koordinaten für B

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{k1} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, a_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{kn} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Dies entspricht (Z1), damit folgt

$$\text{rg}_S(A) = \dim(\text{Span}(a_1, \dots, a_n)) = \dim(\text{Span}(\tilde{a}_1, \dots, \tilde{a}_n))$$

Jetzt (Z2). Dies entspricht der Multiplikation des i -ten Einheitsvektor mit λ^{-1} . D.h. man erhält die neue Basis $\hat{B} = \{e_1, \dots, \lambda^{-1}e_i, \dots, e_m\}$ und die zugehörigen Koordinaten

$$\hat{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ \lambda a_{i1} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \hat{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ \lambda a_{in} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Es folgt

$$\operatorname{rg}_S(A) = \dim(\operatorname{Span}(a_1, \dots, a_n)) = \dim(\operatorname{Span}(\hat{a}_1, \dots, \hat{a}_n))$$

(Z3)

Ersetzt man den k -ten Einheitsvektor durch $e_k - e_i$, $i, k \in \{1, \dots, m\}$, $i < k$ o.B.d.A. Man erhält die neue Basis $\check{B} = \{e_1, \dots, e_i, \dots, e_k - e_i, \dots, e_m\}$ sowie die neuen Koordinaten

$$\check{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{k1} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \check{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{kn} + a_{in} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Es folgt daraus wieder, dass

$$\operatorname{rg}_S(A) = \dim(\operatorname{Span}(a_1, \dots, a_m)) = \dim(\operatorname{Span}(\check{a}_1, \dots, \check{a}_n))$$

□

Beispiel 5.32: $K = \mathbb{R}$ und

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{pmatrix} \longrightarrow A^{(1)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix}$$

mit (Z1): $i = 1$, $\lambda = -1$, (Z3): $k = 2$, $i = 1$, (Z1): $i = 1$, $\lambda = -1$

$$\longrightarrow A^{(2)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\implies \operatorname{rg}_S(A^{(2)}) = 2 \quad \text{und} \quad \operatorname{rg}_Z(A^{(2)}) = 2$$

6. Lineare Gleichungssysteme

Ein System von Gleichungen der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{12}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

$$\iff Ax = b$$

für eine gegebene Matrix $A \in K^{m,n}$, einem gegebenen Vektor $b \in K^m$ und einem unbekannten Vektor $x \in K^n$ heißt **lineares Gleichungssystem (LGS)**. Die Lösung von linearen Gleichungssystemen ist ein zentrales Problem der linearen Algebra. Wichtige Fragen:

1. Existenz einer Lösung? Unter welchen Bedingungen besitzt (LGS) überhaupt Lösungen?
2. Lösungsmannigfaltigkeiten: Wann (LGS) mindestens eine Lösung besitzt: Unter welchen Bedingungen gibt es genau eine Lösung? Wenn es mehrere Lösungen gibt: Wie ist die Struktur der Lösungsmenge?
3. Lösungsverfahren: Wie kann man Lösungen von (LGS) praktisch berechnen?

Hier: Fragen 1 + 2), etwas zu 3). Beantwortung von 3): Numerische Lineare Algebra

6.1. Existenz von Lösungen und Lösungsmengen

Definition 6.1: Lineares Gleichungssystem

Für das lineare Gleichungssystem

$$Ax = b \quad \text{mit } A \in K^{m,n}, b \in K^m$$

nennt man A die **Koeffizientenmatrix** und b die **rechte Seite**. Die Matrix

$$A_{\text{erw}} = (A \mid b) \in K^{m,n+1}$$

nennt man **erweiterte Koeffizientenmatrix**.

Ist $b = 0 \in K^m$, so heißt (LGS) **homogen**, ansonsten **inhomogen**. Gilt für $\bar{x} \in K^n$, dass $A\bar{x} = b$, so nennt man \bar{x} eine **Lösung** von (LGS). Die Menge aller Lösungen heißt **Lösungsmenge** und wird mit

$$\mathcal{L}(A, b)$$

bezeichnet.

Bemerkung: Im normalen Fall, d.h. sei $b = 0 \in K^m$ gilt offensichtlich immer $\text{rg}_S(A) = \text{rg}_S(A_{\text{erw}})$. (Später: $\text{rg}(A) = \text{rg}(A_{\text{erw}})$). Außerdem existiert im homogenen Fall immer eine Lösung, nämlich $x^* = 0 \in K^n$. Denn es gilt $Ax^* = A0_m = 0 = b$.

Beantwortung Frage 1:

Satz 6.2: Das LGS ist genau dann lösbar, d.h. $\mathcal{L}(A, b) \neq \emptyset$, wenn

$$\text{rg}_S(A) = \text{rg}_S(A_{\text{erw}})$$

gilt.

Beweis: Seien $a_i, 1 \leq i \leq n$, die Spalten von A und $f : K^n \rightarrow K^m$ die zu A gehörende lineare Abbildung. Dann gilt

$$\begin{aligned} \text{LGS lösbar} &\iff \exists x \in K^n : Ax = b \iff \exists x \in K^n : f(x) = b \iff b \in \text{im}(f) \\ &\iff b \in \text{Span}\{f(e_1), \dots, f(e_n)\} \\ &\iff b \in \text{Span}\{a_1, \dots, a_n\} \\ &\iff \text{Span}\{a_1, \dots, a_n\} = \text{Span}\{a_1, \dots, a_n, b\} \\ &\iff \text{rg}_S(A) = \text{rg}_S(A_{\text{erw}}) \end{aligned}$$

□

Folgerung: Ist $m \leq n$ und gilt, dass $\text{rg}_S(A) = m$, so besitzt (LGS) mindestens eine Lösung.

Denn: Wegen $m \leq n$ spannen die Spaltenvektoren $a_1, \dots, a_m, a_{m+1}, \dots, a_n \in K^m$ von A den K^m auf.

$$\implies \text{rg}_S(A) = m = \text{rg}_{S(A_{\text{erw}})}$$

Satz 6.3: Sei $L_0 = \mathcal{L}(A, 0)$, d.h. die Menge aller Lösungen des homogenen linearen Gleichungssystems (LGS). Dann ist $L_0 \subseteq K^n$ ein Unterraum der Dimension

$$\dim L_0 = n - \underbrace{\text{rg}_S(A)}_{\leq n} \geq 0$$

Beweis: Die Matrix $A \in K^{m,n}$ definiert eine lineare Abbildung $f : K^n \rightarrow K^m, f(x) = Ax$. Also gilt

$$L_0 = \{x \in K^n \mid Ax = 0\} = \{x \in K^n \mid f(x) = 0\} = \ker(f)$$

Lemma 4.13: L_0 ist Unterraum von K^n .

Dimensionssatz Satz 4.16: Satz 5.23.

$$\dim L_0 = \dim(\ker(f)) \stackrel{4.16}{=} \dim(K^n) - \text{rg}(f) \stackrel{5.23}{=} n - \text{rg}_S(A)$$

□

Folgerungen:

- Ist $m \geq n$ und gilt $\text{rg}_S(A) = n$, so besitzt (LGS) für $b = 0$ nur die triviale Lösung $x^* = 0$, denn aus dem letzten Satz folgt

$$\dim L_0 = n - \text{rg}_S(A) = n - n = 0 \implies L_0 = \{0\}$$

- Besitzt umgekehrt (LGS) für $b = 0$ nur $x^* = 0 \in K^n$ als Lösung, d.h.

$$L_0 = \{0\} \implies 0 = n - \text{rg}_S(A) \implies \text{rg}_S(A) = n$$

- Im Fall $m = n$, dann gilt für das homogene Gleichungssystem $Ax = b$

$$L_0 = \{0\} \iff A \text{ invertierbar, d.h. } A \in \text{GL}_n(K)$$

Alternativer Beweis:

$$A \in \text{GL}_n(K) \iff \text{Es existiert die inverse Matrix } A^{-1} \in \text{GL}_n(K)$$

$$Ax = 0$$

$$A^{-1}Ax = A^{-1}0$$

$$x = 0$$

Lösungsmenge für den allgemeinen Fall:

Satz 6.4: Es gelte für die Lösungsmenge von (LGS), dass

$$\mathcal{L}(A, b) \neq \emptyset$$

D.h. die Menge aller Lösungen von (LGS) ist nicht leer. Dann gilt: Es gibt $x^* \in \mathcal{L}(A, b)$ und

$$\mathcal{L}(A, b) = \{x \in K^n \mid x = x^* + y, y \in \mathcal{L}(A, 0)\}$$

D.h. die Menge aller Lösungen von (LGS) erhält man aus einer speziellen Lösung des inhomogenen linearen Gleichungssystems und Addition sämtlicher Lösungen des zugehörigen homogenen LGS.

Beweis: Nach Voraussetzung gilt $Ax^* = b$. Damit folgt für $y \in \mathcal{L}(A, 0)$, d.h. $Ay = 0$, dass

$$A(x^* + y) = Ax^* + Ay = b + 0 = b$$

$$\implies \{x \in K^n \mid x = x^* + y, y \in \mathcal{L}(A, 0)\} \subseteq \mathcal{L}(A, b)$$

Sei $x \in \mathcal{L}(A, b)$. Dann erhält man für $y = x^* - x$, dass

$$Ay = A(x^* - x) = Ax^* - Ax = b - b = 0 \implies y \in \mathcal{L}(A, 0)$$

$$\implies Ay = 0 \implies -Ay = 0 \implies A(-y) = 0 \implies -y \in \mathcal{L}(A, 0)$$

$$\mathcal{L}(A, b) \subseteq \{x \in K^n \mid x = x^* + y, y \in \mathcal{L}(A, 0)\}$$

□

Was ist die "Größe" des Lösungsraumes?

Problem: $\mathcal{L}(A, b)$ ist ja kein Unterraum, da für $b \neq 0, x = 0 \notin \mathcal{L}(A, b)$ gilt. Deswegen: Übertragung des Dimensionsbegriffs auf Mengen dieser Struktur.

Definition 6.5: Affiner Unterraum

Sei V ein K -Vektorraum. Eine Teilmenge $U \subset V$ heißt **affiner Unterraum** falls es ein $v \in V$ und einen Untervektorraum W gibt, so dass

$$U = v + W := \{u \in V \mid \text{es gibt ein } w \in W \text{ mit } u = v + w\}$$

Die leere Menge bildet ebenfalls einen affinen Unterraum.

Affine Unterräume kann man als "Parallelverschiebung" eines Unterraums interpretieren.

Frage:

$$\{x \in K^n \mid x = x^* + y, y \in \mathcal{L}(A, 0)\} = \{x \in K^n \mid x = \tilde{x}^* + y, y \in \mathcal{L}(A, 0)\} \\ \text{für } x^*, \tilde{x}^* \in \mathcal{L}(A, b)?$$

Satz 6.6: Sei V ein K -Vektorraum und $U \subset V$ ein affiner Untervektorraum. Der Aufhängepunkt $v \in V$ kann beliebig gewählt werden. D.h. für beliebiges $\tilde{v} \in U$ ist $U = \tilde{v} + W$. Weiter ist zu jedem affinen Unterraum $U = v + W$ der Untervektorraum W eindeutig bestimmt. D.h. ist $\tilde{v} \in V$ und $\tilde{W} \subseteq V$ ein Unterraum mit

$$v + W = \tilde{v} + \tilde{W}$$

so folgt $W = \tilde{W}$ und $\tilde{v} - v \in W$

Beweis: $U = v + W, \tilde{v} \in U \implies \exists \bar{w} \in W : \tilde{v} = v + \bar{w} \implies v = \tilde{v} - \bar{w}$

$$1) U \subseteq \tilde{v} + W$$

$$u \in U \implies u = v + w \text{ mit } w \in W \\ \implies u = \tilde{v} + \underbrace{w - \bar{w}}_{\in W}$$

$$\implies U \subseteq \tilde{v} + W$$

$$2) \tilde{v} + W \subseteq U$$

$$U = \tilde{v} + W \stackrel{*}{=} v + \underbrace{\bar{w} + w}_{\in W} \in U$$

$$\implies U = \tilde{v} + W$$

Sei nun $U = v + W = \tilde{v} + \tilde{W}$. Mit $U - U = \{u - \tilde{u} \mid u, \tilde{u} \in U\}$ als Menge der Differenzen folgt, dass $U - U \stackrel{(1)}{=} W$ und $U - U \stackrel{(2)}{=} \tilde{W}$.

$$\implies W = \tilde{W}$$

wir wissen: $v + W = \tilde{v} + W$

$$\implies \exists w \in W : \tilde{v} = v + w \implies \tilde{v} - v = w \in W$$

□

Definition 6.7: Dimension affiner Unterräume

Sei V ein K -Vektorraum und $U = v + W \subset V$ ein affiner Unterraum. Dann ist

$$\dim U = \dim W$$

die Dimension von U . Dies ist wegen Satz 6.6 eine sinnvolle Definition. Satz 6.4 + Satz 6.3 liefern, dass

$$\dim \mathcal{L}(A, b) = \dim \mathcal{L}(A, 0) = n - \operatorname{rg}_S(A)$$

Damit folgt

Korollar 6.8: Für $A \in K^{n,n}$ und $b \in K^n$ sind folgende Bedingungen äquivalent:

1. Das LGS $Ax = b$ ist eindeutig lösbar

Überlegung:

$$\mathcal{L}(A, b) = x^* + \underbrace{\mathcal{L}(A, 0)}_{=\{0\}}$$

2. $\operatorname{rg}_S(A) = \operatorname{rg}_S(A_{\text{erw}}) = n$

Lemma 6.9: Sind $A \in K^{m,n}$, $b \in K^m$ und $S \in K^{m,m}$, so gilt

$$\mathcal{L}(A, b) \subseteq \mathcal{L}(SA, Sb)$$

Ist S invertierbar, so gilt sogar

$$\mathcal{L}(A, b) = \mathcal{L}(SA, Sb)$$

Beweis: Übungsaufgabe

□

6.2. Der Gauß-Algorithmus

Bemerkung:

- benannt nach Carl-Friedrich-Gauß (1777 - 1855)
- Ein ähnliches Verfahren findet man in 'Neun Bücher zu arithmetische Technik' (ca. 200 v. Chr. in China)
- siehe LR-(bzw LK) Zerlegung in der numerischen linearen Algebra

Wann ist ein LGS einfach lösbar? wenn

$$\left(\begin{array}{cccc|ccc} 1 & * & 0 & * & 0 & & \\ & & 1 & * & 0 & * & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & 0 & \\ & & & & & & 1 \end{array} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Diese Gestalt der Matrix A heißt **Treppennormalform** (wird auch als Echelon-Form oder normierte Zeilenstufenform bezeichnet).

Warum einfach?

Spaltenvertauschungen ändern den Rang nicht und können als Multiplikation mit einer invertierbaren Matrix interpretiert werden.

\Rightarrow Lösungsraum bleibt gleich

Dies liefert

$$\left(\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & a_{1,r+1} & \cdots & a_{1,n} \\ 0 & 1 & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & a_{r,r+1} & \cdots & a_{r,n} \\ 0 & 0 & \cdots & 0 & \boxed{0} & \cdots & 0 \end{array} \right) \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

1. Es existiert genau eine Lösung, wenn

$$\text{rg}_S(A) = \text{rg}_S(A_{\text{erw}}) \iff b_{r+1} = \dots = b_m = 0$$

2. Mit $\text{rg}_S(A) = \text{rg}_S(A_{\text{erw}})$, dann ist eine spezielle Lösung gegeben durch

$$x^* \in K^n \text{ mit} \\ x_1^* = b_1, \dots, x_r^* = b_r, x_{r+1}^* = \dots = x_n^* = 0$$

3. Für $b = 0 \in K^m$ erhält man das zugehörige homogene Gleichungssystem

$$\mathcal{L}(A, b) = x^* + \mathcal{L}(A, 0)$$

$$\text{Basis von } \mathcal{L}(A, 0)? \quad \dim \mathcal{L}(A, 0) = n - r = n - \text{rg}_S(A)$$

Basisvektoren der Form:

$$\begin{pmatrix} -a_{1r+i} \\ -a_{2r+i} \\ \vdots \\ -a_{r+i} \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad i = 1, \dots, n-r, \quad 1 \text{ bei der } r+i\text{-ten Zeile}$$

Deswegen: Gegeben ist ein (LGS)

Ziel: Umformung von (LGS), sodass die resultierende Matrix normalisierte Treppenstufen hat und sich die Lösungsmenge nicht ändert.

Dazu: Elementarmatrizen (geht auch für $R =$ kommutativer Ring mit 1)

hier: $A \in K^{m,n}$, K Körper

- $m \in \mathbb{N}$, $i, j \in \{1, \dots, m\}$, $I_m \in K^{m,m}$ als Einheitsmatrix $e_i \hat{=} i\text{-ter Einheitsvektor}$
 $\Rightarrow I_m = (e_1, e_2, \dots, e_m)$
 - $E_{ij} = e_i e_j^\top = (0_m \dots 0_m \ e_i \ 0_m \dots 0_m) \in K^{m,m}$
- D.h. der Eintrag (i, j) der Matrix E_{ij} ist 1, alle anderen Einträge sind 0.
- für $m \geq 2$, $i, j \in \{1, \dots, m\}$ und $i < j$ die Matrizen

$$P_{ij} = (e_1 \dots e_{i-1} \ e_j \ e_{i+1} \dots e_{j-1} \ e_i \ e_{j+1} \dots e_m) \in K^{m,m}$$

Die Matrix P_{ij} ist eine Permutationsmatrix. D.h. multipliziert man $A \in K^{m,n}$ mit P_{ij} , so werden die Zeilen i und j von A vertauscht.

ohne Beweis, dafür ein Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} P_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow P_{12} \cdot A = \begin{pmatrix} 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

$\hat{=} \text{Zeilenumformung (Z1). } P_{ij} \in \text{GL}_m(K)$

Analog: $\tilde{e}_i \in K^n$ Einheitsvektoren aus K^n

$\tilde{P}_{ij} \in K^{n,n} \hat{=} \text{Spaltenumformung (S1). } \tilde{P}_{ij} \in \text{GL}_n(K)$

- für $m \geq 2$, $i, j \in \{1, \dots, m\}$ und $i < j$ sind die Matrizen

$$G_{ij}(\lambda) = I_m + \lambda E_{ij} = (e_1 \dots e_{j-1} \ e_j + \lambda e_i \ e_{j+1} \dots e_m) \in K^{m,m}$$

D.h. die j -te Spalte von $G_{ij}(\lambda)$ ist $e_j + \lambda e_i$. Multipliziert man $A \in K^{m,n}$, so wird das λ -fache der i -ten Zeile von A zur j -ten Zeile von A addiert.

Das Produkt $G_{ij}(\lambda)^T \cdot A$ bewirkt, dass das λ -fache der j -ten Zeile zur i -ten Zeile von A addiert wird.

- $\lambda \in K, i \in \{1, \dots, m\}, M_i(\lambda) := (e_1, \dots, e_{j-1}, \lambda e_i, e_{j+1}, \dots, e_m) \in K^{m,m} \Rightarrow$ Multipliziert man A von links mit $M_i(\lambda)$, so wird die i -te Zeile von A mit λ multipliziert \cong (Z2) für $\lambda \neq 0$

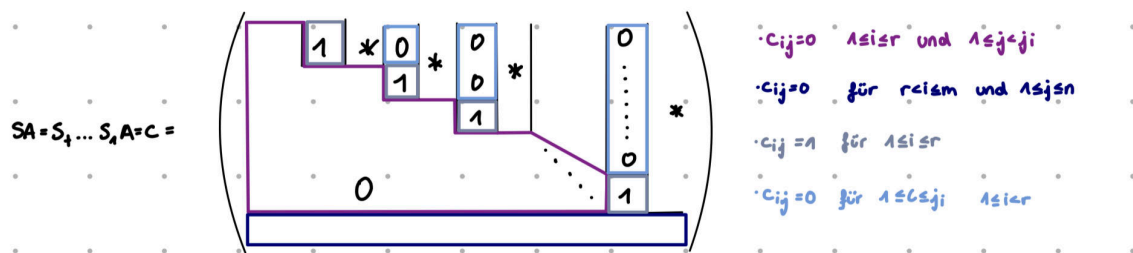
Beobachtung: $P_{ij}, G_{ij}(\lambda), M_i(\lambda)$ sind für $\lambda \neq 0$ invertierbar!

Damit: Beweis, dass $Ax = b$ in einfachere Form gebracht werden kann, **ohne** die Lösungsmenge zu ändern.

Satz 6.10: Sei K ein Körper und $A \in K^{m,n}$. Dann existieren invertierbare Matrizen $S_1, \dots, S_t \in K^{m,m}$ als Produkte der oben eingeführten Elementarmatrizen, so dass

$$C := S_t \cdots S_2 S_1 A$$

in **Treppennormalform** ist, d.h.



Genauer: $C = (c_{ij})$ ist entweder die Nullmatrix oder es gibt eine Folge von natürlichen Zahlen j_1, \dots, j_r als Stufen der Treppennormalform mit $1 \leq j_1 < j_2 < \dots < j_r \leq n$ und $1 \leq r \leq \min\{m, n\}$, so dass

- $c_{ij} = 0$ für $1 \leq i \leq r$ und $1 \leq j < j_i$ $1 \leq i \leq r$
- $c_{ij} = 0$ für $1 \leq i \leq m$ und $1 \leq j \leq n$
- $c_{ij_i} = 1$ für $1 \leq i \leq r$
- $c_{lj_i} = 0$ für $1 \leq l < j_i$ $1 \leq i \leq r$

Bemerkungen:

- Ist $n = m$, so ist $A \in K^{n,n}$ genau dann invertierbar, wenn $C = I_n$, denn

“ \Leftarrow ”:

$$I_n = C = \underbrace{S_t \cdots S_1 A}_{=: A^{-1}}$$

“ \Rightarrow ”: A invertierbar und S_1, \dots, S_t invertierbar. Das Produkt von invertierbaren Matrizen ist invertierbar $\Rightarrow C$ invertierbar $\Rightarrow C$ kann keine Nullzeilen bzw. Spalten enthalten $\Rightarrow C = I_n$

- Da Spaltenvertauschungen den Spaltenrang einer Matrix nicht ändern (Lemma 5.31) ist nun folgende Definition sinnvoll:

Der **Rang einer Matrix** $\text{rg}(A)$ einer Matrix $A \in K^{m,n}$ ist definiert als

$$\text{rg}(A) = \text{rg}_S(A) \stackrel{6.10}{=} \text{rg}_Z(A)$$

Außerdem folgt

$$\text{rg}(A) = \text{rg}(A^T)$$

- Es gibt eine Verallgemeinerung der Treppennormalform für Matrizen über Ringen. Diese Hermit-Normalform spielt eine Rolle in der Zahlentheorie.

Beweis: Ist $A = 0 \in K^{m,n}$, dann folgt mit $S_1 = I_m, t = 1$

$$S_1 A = 0 \quad \checkmark$$

Für $A \neq 0$ sei j_1 der Index der ersten Spalte von

$$A^{(1)} = (a_{ij}^{(1)}) := A$$

die nicht aus lauter Nullen besteht.

$$A = \left(\begin{array}{c|c} \begin{smallmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{smallmatrix} & \begin{smallmatrix} \cdot \\ \cdot \\ \cdot \end{smallmatrix} \end{array} \right) \begin{array}{l} \downarrow j_1 \text{ Spalte} \\ \begin{smallmatrix} 0 \\ 0 \\ \square \\ * \\ * \end{smallmatrix} \end{array} \leftarrow i_1$$

Sei $a_{i_1 j_1}^{(1)}$ das erste Element in dieser Spalte welches nicht Null ist.

$$A = \left(\begin{array}{c|c} \begin{smallmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{smallmatrix} & \begin{smallmatrix} \cdot \\ \cdot \\ \cdot \end{smallmatrix} \end{array} \right) \begin{array}{l} \begin{smallmatrix} 0 \\ \vdots \\ 0 \\ a_{i_1 j_1}^{(1)} \\ * \end{smallmatrix} \end{array}$$

Jetzt: Vertauschen der Zeilen 1 und i_1 falls $i_1 > 1$ und normieren der ersten Zeile durch Multiplikation mit $(a_{i_1 j_1}^{(1)})^{-1}$ liefert eine 1 an der Stelle $(1, j_1)$

$$M_{i_1, 1} (a_{i_1 j_1}^{(1)})^{-1} \cdot P_{1, i_1} \cdot A^{(1)} = \left(\begin{array}{c|c} \begin{smallmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{smallmatrix} & \begin{smallmatrix} \cdot \\ \cdot \\ \cdot \end{smallmatrix} \end{array} \right) \begin{array}{l} \begin{smallmatrix} 1 \\ * \\ \vdots \\ * \end{smallmatrix} \end{array}$$

Anschließend werden alle Einträge unterhalb des Eintrags $(1, j_1)$ mit Wert 1 eliminiert und man erhält

$$\underbrace{G_{m,1}(-\tilde{a}_{m,j_1}^{(1)}) \cdot \dots \cdot G_{3,1}(-\tilde{a}_{3,j_1}^{(1)}) G_{2,1}(-\tilde{a}_{2,j_1}^{(1)}) \cdot M_1(a_{i_1,j_1}^{(1)})^{-1} \cdot p_{A_{i_1}}^{(1)}}_{=: S_1 \in GL_m(K)} = \left(\begin{array}{c|c|c} 0 & \begin{smallmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{smallmatrix} & * \end{array} \right)$$

$$= \left(\begin{array}{c|c} 1 & * \\ \hline 0 & A^{(1)} \end{array} \right)$$

mit $A^{(2)} \in K^{m-1, n-j_1}$. Ist $A^{(2)} = ()$, d.h. $m-1=0, n-j_1=0$ oder $A^{(2)} = 0 \in K^{m-1, n-j_1}$, sind wir fertig, da $C := S_1 A$ in Treppenform ist und $t=1$.

Ist mindestens ein Eintrag von $A_{(2)}$ ungleich Null, so führt man die obigen Schritte für die Matrix $A^{(2)}$ aus.

Für $k=2, 3, \dots$ werden die Matrizen S_k rekursiv durch

$$S_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & \tilde{S}_k \end{pmatrix} \quad \text{und}$$

$$\tilde{S}_k \cdot A^{(k)} = \left(\begin{array}{c|c|c} 0 & \begin{smallmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{smallmatrix} & \begin{smallmatrix} * \\ \hline A^{(k-1)} \end{smallmatrix} \end{array} \right)$$

wobei die Matrix \tilde{S}_k analog zu S_1 konstruiert wird.

S_k ist wieder ein Produkt von Elementarmatrizen. Dieses Verfahren bricht nach $r \leq \min\{m, n\}$ Schritten ab. Wenn entweder $A^{(r+1)} = 0$ oder $A^{(r+1)} = ()$. \Rightarrow Nach r Schritten erhält man damit die Zeilenstufenform.

$$\hat{A} = \left(\begin{array}{c|c|c|c|c|c|c} 0 & \begin{smallmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{smallmatrix} & \begin{smallmatrix} * \\ \hline 0 \end{smallmatrix} & \begin{smallmatrix} * \\ \hline 0 \end{smallmatrix} & \begin{smallmatrix} * \\ \hline 0 \end{smallmatrix} & \dots & \begin{smallmatrix} * \\ \hline 1 \\ \hline 0 \end{smallmatrix} & \begin{smallmatrix} * \\ \hline 0 \end{smallmatrix} \end{array} \right)$$

$$SA = C = \left(\begin{array}{ccc|ccc|c} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right)$$

$$S = S_t \cdots S_1$$

Beweis:

$$S_k = \left(\begin{array}{c|c} I_{k-1} & 0 \\ \hline 0 & \tilde{S}_k \end{array} \right)$$

$$S_k = G_{mk} (-\tilde{a}_{mj_k}^{(k)}) \dots G_{k-1,k} (-\tilde{a}_{k-1,j_k}^{(k)}) U_k ((a_{kj_k}^{(k)})^{-1}) P_{kj_k}$$

Nach r Schritten

$$S_r \dots S_1 A = \left(\begin{array}{ccc|ccc|c} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right) = R^{(1)}$$

j_r-Spalte

← Zeile r

Mit Einsen an Stellen $(1, j_1), (2, j_2), \dots, (r, j_r)$. Gilt $r = 1$, so ist $S_1 A$ in normalisierter Trepp-
penform. ✓

Jetzt $r > 1$

⇒ Einträge über den Einsen müssen den Wert 0 bekommen

Dazu berechnet man rekursiv für $k = 2, \dots, r$

7. Determinanten von Matrizen

Sei K ein Körper, $n \in \mathbb{N}$. Die Determinante

- ist eine Abbildung, die jeder quadratischen Matrix $A \in K^{n,n}$ ein Element aus K zuordnet

$$\det : K^{n,n} \rightarrow K$$

- liefert notwendige Bedingungen dafür, dass eine Matrix $A \in K^{n,n}$ invertierbar ist.
- bildet die Grundlage für die Definition des charakteristischen Polynoms einer Matrix $A \in K^{n,n}$ (siehe LinA II)
- trägt in der analytischen Geometrie zur Berechnung von Volumen einfacher Mengen bei.
- spielt eine wichtige Rolle in Transformationsformeln von Integralen in der Analysis mehrerer Veränderlicher.

Gottfried Wilhelm Leibnitz (1646 - 1716) gab 1690 eine Formel zur Berechnung von Determinanten an. Karl Weierstraß (1815 - 1897) führte die Determinanten über axiomatische Eigenschaften ein.

7.1. Definitionen und grundlegende Eigenschaften

Zwei motivierende Beispiele:

Lösung linearer Gleichungssysteme

- Gegeben sei ein lineares Gleichungssystem

$$a_{11}x_1 + a_{12}x_2 = b_1 \quad I$$

$$a_{21}x_1 + a_{22}x_2 = b_2 \quad II$$

Die Lösung für beliebige Werte der Koeffizienten?

$$a_{22}I - a_{12}II \Rightarrow$$

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{11}b_2$$

Als Ausdruck für x_1 und

$$a_{11}II - a_{21}I \Rightarrow$$

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1$$

Mit der Abbildung $\det : K^{2,2} \rightarrow K$

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

erhält man die Lösungsformeln

$$x_1 = \det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{12} \end{pmatrix} \cdot \left(\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right)^{-1}$$

$$x_2 = \det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix} \cdot \left(\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right)^{-1}$$

Dies ist die sogenannte Cramersche Regel für $n = 2$. Sie kann für $n \geq 2$ verallgemeinert werden. ACHTUNG: Man sollte die Cramersche Regel wirklich nie zur Lösung von linearen Gleichungssystem verwenden!

$$x_1 = \det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix} \cdot (\det(A))^{-1} \quad x_2 = \det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix} \cdot (\det(A))^{-1}$$

Warum ist die Cramersche Regel trotzdem interessant?

x_i berechenbar, wenn $\det(A) \neq 0 \in K$

$$\det(A) = 0 \iff a_{11}a_{22} - a_{12}a_{21} = 0$$

$$\iff a_{22} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} - a_{12} \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

D.h. die Spalten von A sind linear abhängig $\implies \text{rg}(A) < 2 \xrightarrow{\text{Korollar 6.8}} \implies$ es existiert keine eindeutige Lösung.

- Flächenberechnung für Parallelogramme:

Die Fläche eines Dreiecks ist gegeben durch

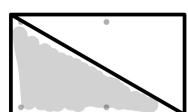
$$\frac{1}{2} \cdot \text{Grundfläche} \cdot \text{Höhe}$$

Beweis z.B. durch das Cavalierische Prinzip, wobei

1. das Dreieck zu einem Parallelogramm ergänzt wird



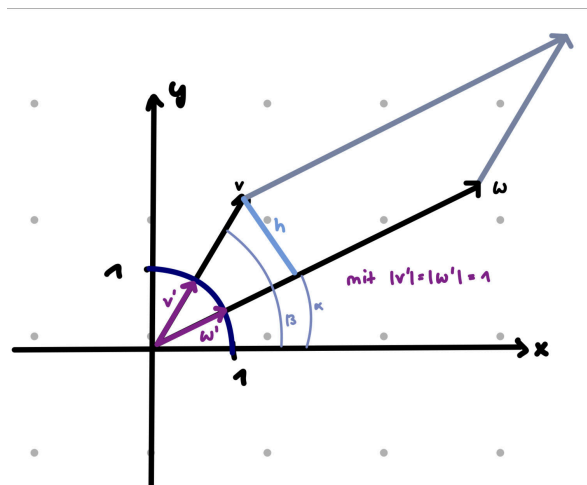
2. "Verschieben" des Parallelogramms zu einem Rechteck



Zur Beschreibung dieser Fläche in der Ebene nehmen wir an, dass sie durch zwei Vektoren

$$v = \begin{pmatrix} a \\ b \end{pmatrix} \quad w = \begin{pmatrix} c \\ d \end{pmatrix}, \quad v, w \in \mathbb{R}^2$$

gegeben ist.



$$v = \rho v' = \rho \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix} \quad w = \sigma w' = \sigma \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$$

mit $\rho, \sigma > 0$. Bezeichnet F bzw. F' die Fläche des von v und w bzw. v' und w' aufgespannten Parallelogramms, so erhält man für $0 \leq \beta - \alpha \leq \pi$, dass

$$\begin{aligned} h' &= \sin(\beta - \alpha) = \cos \alpha \sin \beta - \cos \beta \sin \alpha \\ &= \det \begin{pmatrix} \cos \alpha & \sin \alpha \\ \cos \beta & \sin \beta \end{pmatrix} \end{aligned}$$

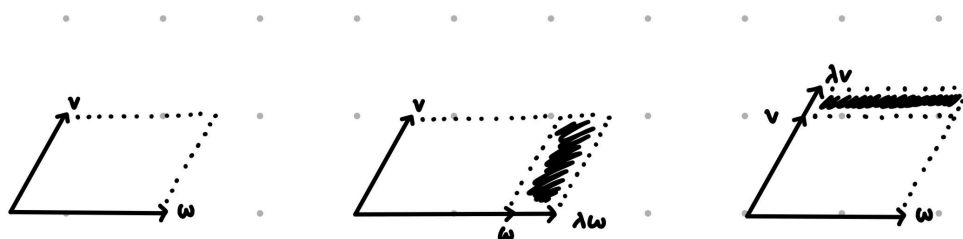
und damit auch

$$F = \rho \sigma F' = \rho \sigma h' = \rho \sigma \det \begin{pmatrix} \cos \alpha & \sin \alpha \\ \cos \beta & \sin \beta \end{pmatrix} \geq 0$$

Im Fall $\sin(\beta - \alpha) \geq 0$ entspricht die Fläche des Parallelogramms damit der skalierten Determinante.

Beobachtungen:

1.



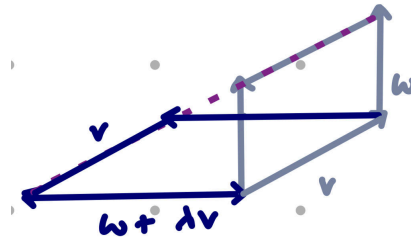
Auswirkungen auf die Fläche

D.h. für $\lambda, \mu \in \mathbb{R}$ sollte gelten

$$\det \begin{pmatrix} \mu v \\ w \end{pmatrix} = \mu \det \begin{pmatrix} v \\ w \end{pmatrix} \quad \text{sowie} \quad \det \begin{pmatrix} v \\ \lambda w \end{pmatrix} = \lambda \det \begin{pmatrix} v \\ w \end{pmatrix}$$

mit $v = \begin{pmatrix} a \\ b \end{pmatrix}$ und $w = \begin{pmatrix} c \\ d \end{pmatrix}$.

2. Nach dem Cavalierischen Prinzip ist die Fläche unabhängig vom Scherungswinkel:

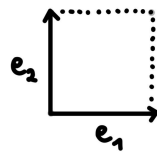


D.h. für $\lambda \in \mathbb{R}$ sollte gelten

$$\det \begin{pmatrix} v \\ w \end{pmatrix} = \det \begin{pmatrix} v \\ w + \lambda v \end{pmatrix}$$

für $v = \begin{pmatrix} a \\ b \end{pmatrix}$ und $w = \begin{pmatrix} c \\ d \end{pmatrix}$.

3.



Fläche 1 \Rightarrow

$$\det \begin{pmatrix} e_1^\top \\ e_2^\top \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det I_2 = 1$$

4. $\xrightarrow{w} \xrightarrow{v}$ oder $\xleftarrow{w} \xrightarrow{v} \Rightarrow \det \begin{pmatrix} v \\ w \end{pmatrix} = 0$

Basierend auf diesen Beobachtungen:

Definition 7.1: Determinante

Eine Abbildung

$$\det : K^{n,n} \rightarrow K, A \mapsto \det(A)$$

heißt **Determinante** falls gilt:

D1. \det ist linear in jeder Zeile, d.h. für jeden Index $i \in \{1, \dots, n\}$ gilt

1. Ist $a_i = a'_i + a''_i$ so ist

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_i \\ \vdots \end{pmatrix}$$

2. Ist $a_i = \lambda a'_i$ für $\lambda \in K$, so ist

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} = \lambda \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix}$$

D2. \det ist **alternierend**, d.h. hat A zwei gleiche Zeilen, so gilt $\det(A) = 0$

D3. \det ist normiert, d.h. $\det(I_n) = 1$

Bemerkung:

- Definition ist sehr einfach, aber sehr mächtig, siehe D4 - D13
- Existenz und Eindeutigkeit einer solchen Abbildung sind noch zu zeigen.

Satz 7.2: Eine Determinante

$$\det : K^{n,n} \mapsto K$$

hat die folgenden weiteren Eigenschaften:

D4. Für jedes $\lambda \in K$ gilt $\det(\lambda A) = \lambda^n \det(A)$

D5. Ist eine Zeile von A gleich 0, so gilt $\det(A) = 0$

D6. Ist B eine Matrix die aus A durch Zeilenvertauschungen entsteht (Z1), so gilt $\det(B) = -\det(A)$

D7. Ist $\lambda \in K$ und entsteht B aus A durch Addition der λ -fachen j -ten Zeile zur i -ten Zeile ($i \neq j$) (Z2+Z3) so ist $\det B = \det A$

D8. Ist A eine obere Dreiecksmatrix

$$A = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda \end{pmatrix}$$

so ist $\det A = \lambda_1 \cdot \dots \cdot \lambda_n$!

D9. Ist $n \geq 2$ und $A \in K^{n,n}$ von der Gestalt

$$A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$$

wobei A_1 und A_2 quadratisch sind, d.h. $A_1 \in K^{n_1, n_1}$ und $A_2 \in K^{n_2, n_2}$ mit $n_1 + n_2 = n$, so gilt $\det A = \det A_1 \cdot \det A_2$

D10. $\det A = 0$ ist äquivalent zu $\operatorname{rg} A < n$

D11. Für Matrizen $A, B \in K^{n, n}$ gilt der Determinanten-Multiplikationssatz, d.h.

$$\det(A \cdot B) = \det A \cdot \det B$$

Insbesondere gilt damit für $A \in \operatorname{GL}_n(K)$, dass $\det A^{-1} = (\det A)^{-1}$.

D12. $\det A^\top = \det A$. D.h. die obigen Aussagen für Zeilen gelten analog für Spalten.

D13. Achtung: Für Matrizen $A, B \in K^{n, n}$ ist im Allgemeinen $\det(A + B) \neq \det A + \det B$

Beweis: D4 und D5 folgen direkt aus D1.

D6. O.B.d.A $i < j$. Dann gilt wegen D1.1 und D2

$$\begin{aligned} \det A + \det B &= \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} \\ &= \underbrace{\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix}}_{=0} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} - \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} - \underbrace{\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}}_{=0} \\ &\stackrel{\text{D1.1}}{=} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} \stackrel{\text{D1.1}}{=} \det \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} \stackrel{\text{D2}}{=} 0 \end{aligned}$$

D7. O.B.d.A $i < j$. Wegen D1 und D2 gilt

$$\det B = \det \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \stackrel{\text{D1.1}}{=} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det A + \lambda \underbrace{\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}}_{=0} \stackrel{\text{D2}}{=} \det A$$

D8. Sind alle $\lambda_i \neq 0$, so folgt durch wiederholte Anwendung von D7, dass

$$\det A = \det \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} \stackrel{\text{D1.2}}{=} \lambda_1 \cdot \dots \cdot \lambda_n \cdot \underbrace{\det I_n}_{=1 \text{ (D3)}}$$

Gibt es ein $\lambda_i = 0$, so wählen wir i maximal, d.h. $\lambda_{i+1} \neq 0, \dots, \lambda_n \neq 0$. Dann kann der Rest der i -ten Zeile, also a_{ii+1}, \dots, a_{in} auf den Wert Null gebracht werden. Mit D7 folgt, dass dies die Determinante nicht ändert. Damit erhalten wir eine Nullzeile und aus D5 folgt dann $\det A = 0$

D9. $A_1 \in K^{n_1, n_1}$. Mit den Zeilenumformungen (Z1), (Z2)+(Z3) kann A_1 zu einer oberen Dreiecksmatrix C_1 umgeformt werden. Es gilt mit D6 und D7 dass $\det A_1 = (-1)^k \det C_1$

$$\begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix} \rightarrow \begin{pmatrix} C_1 & \tilde{B} \\ 0 & A_2 \end{pmatrix}$$

Also ist A_2 unverändert.

Dann erzeugt man wieder mit (Z1), (Z2)+(Z3) aus A_2 eine obere Dreiecksmatrix C_2 , dabei bleiben C_1 und \tilde{B} unverändert und es gilt

$$\det A_2 = (-1)^l \det C_2$$

Für

$$C = \begin{pmatrix} C_1 & \tilde{B} \\ 0 & C_2 \end{pmatrix} = \left(\begin{array}{ccc|ccc} \lambda_1 & & * & & & \\ & \ddots & & & & \\ 0 & & \lambda_n & & & \\ \hline & & & & \tilde{B} & \\ & & & & & \\ 0 & & & & & \\ \hline & & & & \lambda_{n+1} & * \\ & & & & & \ddots \\ & & & & 0 & \lambda_{n_2} \end{array} \right)$$

mit C_1, C_2 obere Dreiecksmatrizen gilt

$$\begin{aligned} \det C &= \det C_1 \cdot \det C_2 \\ \det A &= (-1)^{k+l} \det C = (-1)^{k+l} \det(C_1) \cdot \det(C_2) \\ &= (-1)^{k+l} (-1)^k \det(A_1) (-1)^l \det(A_2) = (-1)^{2k+2l} \det(A_1) \cdot \det(A_2) \\ &= \det(A_1) \cdot \det(A_2) \end{aligned}$$

D10. Durch Zeilenumformungen (Z1), (Z2)+(Z3) bringen wir A auf Zeilenstufenform

$$B = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix}$$

Wegen D6+D7 gilt $\det(A) = \pm \det(B)$. Weiter ist $\text{rg}(A) = \text{rg}(B)$ (Lemma 5.31). Aus D8

$$\text{rg}(B) = n \iff \det(B) = \lambda_1 \cdot \dots \cdot \lambda_n \neq 0$$

D11. Ist $\text{rg}(A) < n$, so ist $\text{rg}(A \cdot B) < n$ (siehe Übungsaufgabe). Damit erhält man aus D10

$$0 = 0$$

Ist $\text{rg}(A) = n \implies A \in \text{GL}_n(K)$. Dann gibt es nach Satz 6.10 invertierbare Matrizen

Satz 6.10: $C = S_t \cdots S_1 A$ in Treppennormalform

$$\implies C = I_n$$

$$\implies A = S_1^{-1} \cdots S_{t-1}^{-1} S_t^{-1} I_n \quad (*)$$

$$\stackrel{\text{D3}}{\implies} \det I_n = 1$$

$S_i \hat{=}$ Produkt von $G_{ij}(\lambda), M_i(\lambda), P_{ij} \implies S_i^{-1} \hat{=}$ Produkt von $G_{ij}^{-1}(\lambda), M_i^{-1}(\lambda), P_{ij}^{-1}$

Sei $D \in K^{n,n}$, dann gilt

$$\det \left(\underbrace{G_{ij}^{-1}(\lambda) D}_* \right) = ?$$

*: Subtrahieren eines Vielfachen einer Zeile von einer anderen

$$\det(G_{ij}^{-1}(\lambda)) \stackrel{\text{D7}}{=} \det I_n = 1$$

$$\tilde{D} = G_{ij}^{-1}(\lambda) D \implies \det(\tilde{D}) \stackrel{\text{D7}}{=} \det D$$

$$\implies \det(G_{ij}^{-1}(\lambda) D) = \det(G_{ij}^{-1}(\lambda)) \cdot \det(D) = \det(D)$$

Analog folgt:

$$\det(P_{ij}^{-1}(\lambda) D) = \underbrace{\det(P_{ij}^{-1}(\lambda))}_{-1} \cdot \det(D)$$

$$\det(M_i^{-1}(\lambda) D) \stackrel{\text{D1.2}}{=} \det(M_i^{-1}(\lambda)) \cdot \det(D)$$

$$\text{induktiv folgt: } \det(S_i^{-1} D) = \det(S_i^{-1}) \cdot \det(D) \quad (\star)$$

$$\implies \det(A) = \det(S_1^{-1}) \cdot \dots \cdot \det(S_t^{-1}) \cdot 1$$

$$\implies \det(A \cdot B) \stackrel{(*)}{=} \det(S_1^{-1} \cdot \dots \cdot S_t^{-1} B)$$

$$\stackrel{(*)}{=} \det(S_1^{-1}) \cdot \dots \cdot \det(S_t^{-1}) \cdot \det(B)$$

$$= \det(A) \cdot \det(B)$$

$$A \in \text{GL}_n(K) \implies \exists A^{-1} : A \cdot A^{-1} = I_n$$

$$1 \stackrel{\text{D3}}{=} \det(I_n) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$$

$$\implies \det(A^{-1}) = (\det(A))^{-1}$$

D12. $\text{rg}(A) < n$:

Mit Lemma 5.31: $\text{rg}(A^\top) < n$ (Zeilenrang = Spaltenrang)

Mit D10 folgt: $\det(A) = 0 = \det(A^\top)$

Sei $\text{rg}(A) = n$ und sei A wie in D11:

$$A = S_1^{-1} \cdot \dots \cdot S_t^{-1}$$

Dann gilt

$$\det(A^\top) = \det(S_t^{-\top} \cdot S_{t-1}^{-\top} \cdot \dots \cdot S_1^{-1}) \stackrel{\text{D11}}{=} \det(S_t^{-\top}) \cdot \det(S_{t-1}^{-\top}) \cdot \dots \cdot \det(S_1^{-1})$$

induktiv und analog zu oben folgt:

$$\begin{aligned} \det(S_t^{-\top}) &= \det(S_t^{-1}) \\ \implies \det(A^\top) &= \det(S_t^{-\top}) \cdot \dots \cdot \det(S_1^{-1}) \\ &\stackrel{\text{D11}}{=} \det(S_t^{-1} \cdot \dots \cdot S_1^{-1}) = \det(A) \end{aligned}$$

D13. Für $K = \mathbb{R}$, $A = I_2$, $B = I_2$

d.h. $A, B \in \mathbb{R}^{2,2}$. Dann gilt:

$$\begin{aligned} \det(A + B) &= \det(0_{2 \times 2}) \stackrel{\text{D8}}{=} 0 \\ \det(A) + \det(B) &= 1 + 1 = 2 \neq 0 \end{aligned}$$

□

Beispiel 7.3: Berechnung einer Determinante

Für $A \in \mathbb{R}^{n,n}$ kann man eine Determinante berechnen, indem man A auf obere Dreiecksgestalt bringt.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 4 & 1 \\ 1 & -2 & 2 \end{pmatrix} \rightarrow A_1 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & -4 & 2 \end{pmatrix} = G_{31}(-1)G_{21}A$$

(\rightarrow von Zeile 3 wird Zeile 1 mit -1 abgezogen)

$$\implies \det(A_1) \stackrel{\text{D8}}{=} \underbrace{\det(G_{31}(-1))}_{=1} \cdot \underbrace{\det(G_{21}(-1))}_{=1} \cdot \det(A)$$

$$A_2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{pmatrix} = G_{32}(2) \cdot A_1$$

$$8 = \det(A_2) = \det(G_{32}(2)) \cdot \det(A_1) = \underbrace{\det(G_{32}(2))}_{=1} \cdot \det(A)$$

7.2. Existenz und Eindeutigkeit

Frage: Ist die Berechnung der Determinante eindeutig? Oder hängt sie von der Reihenfolge der Zeilenumformungen ab?

Definition 7.4: Permutation

Sei $n \in \mathbb{N}$. Eine bijektive Abbildung

$$\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}, \quad j \mapsto \sigma(j)$$

heißt **Permutation** der Zahlen $\{1, \dots, n\}$. Die Menge aller dieser Abbildungen nennt man S_n und man schreibt eine Permutation $\sigma \in S_n$ in der Form

$$[\sigma(1) \quad \sigma(2) \quad \dots \quad \sigma(n)]$$

Frage: Gilt für ein gegebenes $\sigma \in S_n$ und

$$I_\sigma = \begin{pmatrix} e_{\sigma(1)}^\top \\ \vdots \\ e_{\sigma(n)}^\top \end{pmatrix}$$

$$\det(I_\sigma) = \pm 1 ?$$

Definition 7.5: Transposition

Seien $n \in \mathbb{N}$, $n \geq 2$. Eine Permutation $\mathcal{T} \in S_n$ heißt **Transposition**, wenn \mathcal{T} nur zwei Elemente aus $\{1, 2, \dots, n\}$ vertauscht und alle andere unverändert bleiben. D.h. es existieren $k, l \in \{1, \dots, n\}$ so, dass

$$\mathcal{T}(K) = C, \mathcal{T}(C) = K, \mathcal{T}(i) = i \quad i \in \{1, \dots, n\} \setminus \{k, l\}$$

Lemma 7.6: Ist $n \geq 2$, so existieren zu jedem $\sigma \in S_n$ Transpositionen $\mathcal{T}_1, \dots, \mathcal{T}_k \in S_n$ mit

$$\sigma = \mathcal{T}_1 \circ \mathcal{T}_2 \circ \dots \circ \mathcal{T}_k$$

Beweis: Ist $\sigma = \text{id}$ und \mathcal{T} eine Transposition, dann gilt

$$\text{id} = \sigma = \mathcal{T} \circ \mathcal{T}$$

Ist $\sigma \neq \text{id}$, so existiert ein $i_1 \in \{1, \dots, n\}$ mit

$$\sigma(i) = i \quad \text{für } i = 1, \dots, i_1 - 1 \quad \text{und} \quad \sigma(i_1) > i_1$$

Sei \mathcal{T}_1 die Transposition, die i_1 mit $\sigma(i_1)$ vertauscht und $\sigma_1 := \mathcal{T}_1 \circ \sigma$

Dann gilt

$$\sigma_1(i) = i \quad \text{für } i = 1, \dots, i_1$$

Ist $\sigma_1 = \text{id}$ so gilt $\sigma = \mathcal{T}_1^{-1} = \mathcal{T}_1$, sonst existiert ein i_2 mit $i_2 > i_1$ und

$$\sigma_1(i) = i \quad \text{für } i = 1, \dots, i_2 - 1 \quad \text{und} \quad \sigma_1(i_2) > i_2$$

induktiv erhält man $k \leq n$ Transpositionen

$$\begin{aligned} & \mathcal{T}_1, \dots, \mathcal{T}_k \text{ mit} \\ & \sigma_k = \mathcal{T}_k \circ \dots \circ \mathcal{T}_1 \circ \sigma = \text{id} \\ \Rightarrow & \sigma = \mathcal{T}_1^{-1} \circ \dots \circ \mathcal{T}_k^{-1} = T_1 \circ \dots \circ T_k \end{aligned}$$

□

Die Transpositionen sind nicht eindeutig bestimmt! Für die Entscheidung über das Vorzeichen

$$\det(I_\sigma) = \pm 1$$

müssen wir zeigen, dass die Anzahl der Transpositionen immer gerade oder immer ungerade ist.

Definition 7.7: Fehlstand

Seien $n \in \mathbb{N}$, $n \geq 2$ und $\sigma \in S_n$. Ein Paar $(i, j) \in N \times N$ mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$ nennt man **Fehlstand** von σ . Ist k die Anzahl der Fehlstände von σ , so heißt $\text{sgn}(\sigma)$ das **Signum** oder Vorzeichen von σ . Für $n = 1$ setzt man $\text{sgn}([1]) := 1 = (-1)^0$

Beobachtung: Für $\sigma = [3 \ 2 \ 1]$ ($3 = \sigma(1)$, $2 = \sigma(2)$, $1 = \sigma(3)$) gilt

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{2-3}{2-1} \cdot \frac{1-3}{3-1} \cdot \frac{1-2}{3-2} = (-1)^3 = -1 = \text{sgn}(\sigma)$$

Diese Eigenschaft kann man allgemein beweisen:

Lemma 7.8: Seien $n \in \mathbb{N}$ und $\sigma \in S_n$. Dann gilt

$$\text{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

Beweis:

$n = 1$: Wegen der Definition von sgn und der Definition des leeren Produktes folgt $1 = 1$.

$n > 1$: Sei $\text{sgn}(\sigma) = (-1)^k$, d.h. die Anzahl der Fehlstände ist k . Dann gilt

$$\begin{aligned} \prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i)) &= (-1)^k \prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i)) \\ &= (-1)^k \prod_{1 \leq i < j \leq n} (j - i) \end{aligned}$$

Bei diesen letzten Gleichungen verwendet man, dass die beiden Produkte bis auf die Reihenfolge die gleichen Faktoren enthalten.

□

Satz 7.9: Seien $n \in \mathbb{N}$ und $\mathcal{T}, \sigma \in S_n$. Dann gilt

$$\text{sgn}(\mathcal{T} \circ \sigma) = \text{sgn}(\mathcal{T}) \cdot \text{sgn}(\sigma)$$

Beweis: Wegen Lemma 7.8 gilt:

$$\begin{aligned} \operatorname{sgn}(\mathcal{T} \circ \sigma) &= \prod_{1 \leq i < j \leq n} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{j - i} \\ &= \underbrace{\prod_{1 \leq i < j \leq n} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)}}_{= \operatorname{sgn}(\mathcal{T}) ?} \underbrace{\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}}_{= \operatorname{sgn}(\sigma)} \end{aligned}$$

zu zeigen: ? (s.o.)

$$\begin{aligned} \prod_{1 \leq i < j \leq n} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)} &= \prod_{\substack{1 \leq i < j \leq n \\ \sigma(i) > \sigma(j)}} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{1 \leq i < j \leq n \\ \sigma(i) < \sigma(j)}} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &= \prod_{\substack{1 \leq i < j \leq n \\ \sigma(i) < \sigma(j)}} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{1 \leq j < i \leq n \\ \sigma(i) < \sigma(j)}} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)} = \prod_{\substack{1 \leq i < j \leq n \\ \sigma(i) < \sigma(j)}} \frac{\mathcal{T}(\sigma(j)) - \mathcal{T}(\sigma(i))}{\sigma(j) - \sigma(i)} \\ &\stackrel{*}{=} \prod_{1 \leq i < j \leq n} \frac{\mathcal{T}(j) - \mathcal{T}(i)}{j - i} \stackrel{\text{Lemma 7.8}}{=} \operatorname{sgn}(\mathcal{T}) \end{aligned}$$

★: σ ist bijektiv, damit enthält das letzte Produkt bis auf die Reihenfolge die gleichen Faktoren.

Lemma 7.10: Sei $n \in \mathbb{N}$. Dann gilt

1. Ist $\mathcal{T} \in S_n$ eine Transposition, so gilt $\operatorname{sgn}(\mathcal{T}) = -1$
2. Ist $\sigma \in S_n$ mit $\sigma = \mathcal{T}_1 \circ \dots \circ \mathcal{T}_k$ mit Transpositionen $\mathcal{T}_i \in S_n$, so ist $\operatorname{sgn}(\sigma) = (-1)^k$

Beweis:

zu 1): Sei \mathcal{T} die Transposition, die k und l mit $1 \leq k < l \leq n$ vertauscht. Damit ist $l = k + j$ für ein $j \geq 1$ und \mathcal{T} ist gegeben durch

$$\mathcal{T} = \begin{bmatrix} 1 & 2 & k-1 & \overbrace{k+j}^{=l} & k+1 & \dots & \overbrace{k+j-1}^{l-1} & k & \overbrace{k+j+1}^{l+1} & \dots & n \end{bmatrix}$$

Damit hat \mathcal{T} die Fehlstände

$$\underbrace{(k \quad k+1), \dots, (k \quad k+j), j}_{(k \quad k+j), j \text{ Fehlstände}} \underbrace{(k+1 \quad k+j), \dots, (k+j-1 \quad k+j)}_{j-1 \text{ Fehlstände}}$$

Insgesamt folgt: Anzahl der Fehlstände

$$j + j - 1 = 2j - 1$$

$$\implies \operatorname{sgn}(\mathcal{T}) = (-1)^{2j-1} = -1$$

zu 2) Folgt aus Satz 7.9.

□

Wir brauchen noch:

Definition 7.11: Charakteristik

Ist R ein Ring mit $1 \neq 0$, so ist seine Charakteristik definiert durch

$$\text{char}(R) = \begin{cases} 0, & \text{falls } n - 1 \neq 0 \quad \forall n \in \mathbb{N} \\ \min\{n \in \mathbb{N} \mid n - 1 = 0\} & \text{sonst} \end{cases}$$

Beispiel: \mathbb{F}_2 (vgl. Bsp. 2.8) gilt

$$1 + 1 = 1 - 1 + 1 \cdot 1 = \underbrace{(1 + 1)}_{=0} \cdot 1 = 0$$

$$1 = -1 \hat{=} \text{ inverses Element bzgl. Addition}$$

Allgemein: “Bei einem Ring mit Charakteristik 2 darf man Vorzeichen ignorieren.”

Korollar 7.12: Sei $n \in \mathbb{N}$. Für jede Permutation $\sigma \in S_n$ gilt

$$\det(I_\sigma) = \det(e_{\sigma(n)}^\top, \dots, e_{\sigma(1)}^\top) = \begin{cases} \text{sgn}(\sigma) & \text{für } \text{char}(K) \neq 2 \\ 1 & \text{für } \text{char}(K) = 2 \end{cases}$$

Insgesamt folgt: Anzahl der Fehlstände

$$j + j - 1 = 2j - 1$$

$$\Rightarrow \text{sgn}(\mathcal{T}) = (-1)^{2j-1} = -1$$

zu 2) Folgt aus Satz 7.9.

□

Satz 7.13: Leibniz-Formel

Ist K ein Körper und $n \geq 1$, so gibt es genau eine Determinante $\det : K^{n,n} \rightarrow K$. Dies ist für $A \in K^{n,n}$ die Abbildung

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)} \quad (\star)$$

Für $\text{char}(K) = 2$ sind die Faktoren $\text{sgn}(\sigma)$ in den Summanden überflüssig.

Beweis:

Eindeutigkeit:

Dazu: Beweis, dass (\star) aus Axiomen folgt. Dafür: Darstellung jedes Zeilenvektors a_i von A als

$$a_i = a_{i1}e_1^\top + a_{i2}e_2^\top + \dots + a_{in}e_n^\top$$

Aus D1 folgt:

$$\begin{aligned}
 \det(A) &= \det \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{j_1=1}^n a_{1j_1} \det \begin{pmatrix} e_{j_1}^\top \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \\
 &= \sum_{j_1=1}^n a_{1j_1} \cdot \left(\sum_{j_2=1}^n a_{2j_2} \det \begin{pmatrix} e_{j_1}^\top \\ e_{j_2}^\top \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \right) \\
 &= \dots \\
 &= \sum_{j_1=1}^n \sum_{j_2=1}^n \sum_{j_n=1}^n a_{1j_1} a_{2j_2} a_{3j_3} \dots a_{nj_n} \det \underbrace{\begin{pmatrix} e_{j_1}^\top \\ e_{j_2}^\top \\ \vdots \\ e_{j_n}^\top \end{pmatrix}}_{(1)}
 \end{aligned}$$

$$(1): \operatorname{sgn}(\sigma) = \operatorname{sgn}([j_1 \ j_2 \ \dots \ j_n])$$

$$\det \begin{pmatrix} e_{1j_1}^\top \\ \vdots \\ e_{nj_n}^\top \end{pmatrix} = \begin{cases} 0 & \exists k \\ l \in \{1, \dots, n\} \ j_k = j_l & \\ \operatorname{sgn}(\sigma) & \sigma = [j_1 \ \dots \ j_n] \end{cases}$$

D1 + D2 sichern damit die Eindeutigkeit.

Existenz: zu zeigen, dass (\star) die Axiome D1-D3 aus Def 7.1 erfüllt.

D1.1:

$$\begin{aligned}
 \det \begin{pmatrix} \vdots \\ a'_i + a''_i \\ \vdots \end{pmatrix} &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \dots a_{n\sigma(n)} \\
 &= \sum_{\sigma \in S_1} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a'_{i\sigma(i)} \dots a_{n\sigma(n)} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a''_{i\sigma(i)} \dots a_{n\sigma(n)} \\
 &= \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_i \\ \vdots \end{pmatrix}, \text{ D1.2 kann man analog nachrechnen}
 \end{aligned}$$

D2:

Angenommen in $A \in K^{n,n}$ sind Zeilen k und l mit $k < l$ gleich. Ist \mathcal{T} die Transposition, die k und l vertauscht, so gilt $\operatorname{sgn}(\mathcal{T}) = -1$. Betrachte

$$A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$$

Dann gilt für $\sigma \in S_n$ mit $\operatorname{sgn}(\sigma) = -1$ dass $\sigma \in A_n^C := \{\tilde{\sigma} \circ \mathcal{T} \mid \tilde{\sigma} \in S_n\}$, denn

$$\operatorname{sgn}(\sigma \circ \mathcal{T}^{-1}) = 1 \quad \text{und} \quad \sigma = \underbrace{\sigma \circ \mathcal{T}^{-1}}_{\in A_n} \circ \mathcal{T}$$

Dann ist

$$S_n = A_n \cup A_n^C \quad \text{und} \quad A_n \cap A_n^C = \emptyset$$

Daraus folgt:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in A_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} - \sum_{\sigma \in A_n^C} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in A_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} - \sum_{\sigma \in A_n} a_{1\sigma(\mathcal{T}(1))} \cdots a_{n\sigma(\mathcal{T}(n))} \end{aligned}$$

$$a_l = a_k \implies$$

$$\begin{aligned} &\overbrace{a_{1\sigma(1)}}^{\sigma(\mathcal{T}(1))=\sigma(1)} \cdots a_{k\sigma(\mathcal{T}(k))} \cdots a_{l\sigma(\mathcal{T}(l))} \cdots \overbrace{a_{n\sigma(n)}}^{\sigma(n)=\sigma(\mathcal{T}(n))} \\ &= a_{1\sigma(1)} \cdots a_{k\sigma(l)} \cdots a_{l\sigma(k)} \cdots \underbrace{a_{n\sigma(\mathcal{T}(n))}}_{=a_{n\sigma(n)}} \end{aligned}$$

Damit heben sich die Summanden in (\star) gegenseitig auf und es folgt $\det(A) = 0$.

D3: Für $\sigma \in S_n$ gilt

$$\delta_{1\sigma(1)} \cdot \delta_{2\sigma(2)} \cdots \delta_{n\sigma(n)} = \begin{cases} 0 & \text{für } \sigma \neq \text{id} \\ 1 & \text{für } \sigma = \text{id} \end{cases}$$

Damit folgt

$$\det(I_n) = \det(\delta_{ij}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)} = 1^n$$

□

Berechnung von Determinanten?

Man kann sich überlegen, dass die Leibniz-Formel $n!$ Summanden hat \rightarrow Komplexität der Berechnung? $\hat{=}$ Rechenaufwand ?? (?)

Beispiel: lösen von $Ax = b$ mit $A \in \mathbb{R}^{n,n}$, $n = 20$, $b \in \mathbb{R}^n$ pro Operation ($\hat{=}$ Multiplikation, Addition, ...) 10^9 Sekunden

- Cramersche Regel:

$$x_j = \frac{\det(A_j)}{\det(A)}$$

d.h. ersetze in A_j die j -te Spalte von A durch b .

Leibniz-Formel: $\geq n \cdot n!$ Operationen

$n = 20$: $n \cdot n! \approx 5 \cdot 10^{19}$ Operationen

\Rightarrow Rechenzeit von 1500 Jahren

- Laplacescher Entwicklungssatz:

$$\det(A) = \sum (-1)^{i+j} a_{ij} \det(A_{ij})$$

$A_{ij} \hat{=}$ Streichen i -te Zeile, j -te Spalte

Operationen: $n \cdot 2^n$

$n = 20$: $n \cdot 2^n \approx 21000000$

Rechenzeit: 0.021 Sekunden

- Gauß-Verfahren

etwa n^3 Operationen (siehe numerische lineare Algebra)

$n = 20$: 8000 Operationen

Rechenzeit: 0.000008 Sekunden