Lineare Algebra 1

Daniele Agostini

23. Juli 2024

Vorwort

Diese Notizen basieren auf den Notizen des Kurses von Hannah Markwig. Unsere Notizen sind eher grob und ersetzen in keiner Weise ein gutes Buch über lineare Algebra. Falls Sie Fehler finden, teilen Sie mir diese (auch die offensichtlichen) bitte mit!

Inhaltsverzeichnis

Kapitel 1

Grundlagen

1.1 Aussagen

Siehe Analysis I.

1.2 Mengen

Das Konzept der Menge ist eines der grundlegenden Konzepte der Mathematik, und es gibt einen ganzen Zweig der Logik, die Mengenlehre, der sich mit ihren Studien beschäftigt. In diesem Kurs (und im Großteil der Mathematik) werden wir die naive Definition einer Menge verwenden:

Als Menge wird in der Mathematik ein abstraktes Objekt bezeichnet, das aus der Zusammenfassung einer Anzahl einzelner Objekte hervorgeht. Diese werden dann als die Elemente der Menge bezeichnet. (Wikipedia)

Man schreibt

$$\begin{aligned} x \in M &\iff x \text{ ein Element in } M \text{ ist }, \\ x \notin M &\iff x \text{ kein Element in } M \text{ ist }. \end{aligned}$$

Beispiel 1.2.1 (Mengen von Zahlen)**.** Die Menge von *natürliche Zahlen* ist

$$\mathbb{N} = \{0, 1, 2, \dots\}$$ [1]

Die Menge der *ganze Zahlen* ist

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$

Die Menge der *rationelle Zahlen* ist

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

Die Menge der *reelle Zahlen* ist

$$\mathbb{R}.$$

[1] Vorsicht: manchmal ist die Menge von natürliche Zahlen als $\{1, 2, 3 \dots\}$ definiert. In unserem Kurs 0 ist eine natürliche Zahl.

Als Beispiel, sehen wir dass $-4 \in \mathbb{Z}$ aber $-4 \notin \mathbb{N}$. Man kann auch zeigen dass $\sqrt{2} \in \mathbb{R}$ aber $\sqrt{2} \notin \mathbb{Q}$.

Beispiel 1.2.2 (Viele andere Mengen)**.** Viele Beispiele von andere Mengen:

- Die zwei Mengen $\{1, 2, 3, 4\} = \{1, 1, 2, 3, 4\}$ sind gleich, weil sie die gleiche Elementen haben.
- Die menge von gerade natürliche Zahlen: $\{n \in \mathbb{N} \,|\, n \text{ gerade } \} = \{0, 2, 4, 6, \dots\}$.
- Die leere Menge $\emptyset$, die keine Elemente enthält.
- Die Menge M von alle Menschen auf der Erde: M hat circa 8.1 Milliarde Elementen. Die Menge $M_{>200} = \{m \in M \,|\, m \text{ ist mehr als 200 Jahre alt } \} = \emptyset$.

Definition 1.2.3 (Teilmenge)**.** Seien M_1, M_2 Mengen. Wir sagen dass M_1 eine *Teilmenge* von M_2 ist falls alle Elemente von M_1 auch Elemente von M_2 sind. Wir schreiben $M_1 \subseteq M_2$. Mit Symbole:

$$M_1 \subseteq M_2 \iff (x \in M_1 \implies x \in M_2)$$

Wir sagen dass M_1 eine *echte Teilmenge* ist, falls $M_1 \subseteq M_2$ und $M_1 \neq M_2$. Wir schreiben $M_1 \subsetneq M_2$.

Beispiel 1.2.4. Man kann zeigen dass $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$.

Bemerkung 1.2.5. Seien M_1, M_2 zwei Mengen, dann

$$M_1 = M_2 \iff M_1 \subseteq M_2 \text{ und } M_2 \subseteq M_1.$$

Definition 1.2.6 (Schnitt, Vereinigung, Differenz, Komplement)**.** Seien M_1, M_2 Mengen.

- Der *Schnitt* $M_1 \cap M_2$ enthält alle Elemente, die sowohl in M_1 als auch in M_2 sind:

$$M_1 \cap M_2 = \{x \,|\, x \in M_1, x \in M_2\}.$$

- Die *Vereinigung* enthält alle Elemente die in M_1 oder in M_2 sind:

$$M_1 \cup M_2 = \{x \,|\, x \in M_1 \text{ oder } x \in M_2\}$$

- Die *Differenz* $M_1 \setminus M_2$ enthält alle Elemente in M_1 die nicht in M_2 enthält sind:

$$M_1 \setminus M_2 = \{x \in M_1 \,|\, x \notin M_2\}.$$

- Wenn $M_2 \subseteq M_1$, das *Komplement* von M_1 in M_2 ist

$$M_1^c = M_1 \setminus M_2$$

Bemerkung 1.2.7. Die Notation M^c ist immer abhängig vom Kontext: das Komplement von $\mathbb{N}$ in $\mathbb{Z}$ ist $\mathbb{N}^c = \{x \in \mathbb{Z} \,|\, x \notin \mathbb{N}\} = \{\dots, -3, -2, -1\}$, aber das Komplement von $\mathbb{N}$ in $\mathbb{N}$ ist $\mathbb{N}^c = \{x \in \mathbb{N} \,|\, x \notin \mathbb{N}\} = \emptyset$.

Wir können auch der Schnitt und der Vereinigung von beliebige viele Mengen $M_i, i \in I$ definieren:

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

Definition 1.2.8. Zwei Mengen M_1, M_2 heißen *disjunkt*, falls $M_1 \cap M_2 = \emptyset$. Wenn M_1, M_2 disjunkt sind, wir schreiben manchmal die Vereinigung als $M_1 \sqcup M_2$ oder $M_1 \dot{\cup} M_2$.

Lemma 1.2.9 (Eigenschaften des Schnittes und der Vereinigung)**.** *Seien M_1, M_2, M_3 Mengen. Der Schnitt und der Vereinigung sind:*

- *Kommutativ:*

$$M_1 \cap M_2 = M_2 \cap M_1,$$
$$M_1 \cup M_2 = M_2 \cup M_1.$$

- *Associativ:*

$$(M_1 \cap M_2) \cap M_3 = M_1 \cup M_2 \cup M_3 = M_1 \cap (M_2 \cap M_3),$$
$$(M_1 \cup M_2) \cup M_3 = M_1 \cap M_2 \cap M_3 = M_1 \cap (M_2 \cap M_3).$$

- *Distributiv:*

$$(M_1 \cap M_2) \cup M_3 = (M_1 \cup M_3) \cap (M_2 \cup M_3),$$
$$(M_1 \cup M_2) \cap M_3 = (M_1 \cap M_3) \cup (M_2 \cap M_3).$$

Beweis. Als Beispiel, zeigen wir dass $(M_1 \cap M_2) \cup M_3 = (M_1 \cup M_3) \cap (M_2 \cup M_3)$. Wir zeigen die zwei Inklusionen:

- $\subseteq$: wir wollen zeigen, dass $(M_1 \cap M_2) \cup M_3 \subseteq (M_1 \cup M_3) \cap (M_2 \cup M_3)$. Sei $x \in (M_1 \cap M_2) \cup M_3$, dann $x \in M_1 \cap M_2$ oder $x \in M_3$. Wenn $x \in (M_1 \cap M_2)$ dann $x \in M_1 \cup M_3$ und $x \in M_2 \cup M_3$, so dass $x \in (M_1 \cup M_3) \cap (M_2 \cup M_3)$. Wenn $x \in M_3$, dann $x \in M_1 \cup M_3$ und $x \in M_2 \cup M_3$, so dass $x \in (M_1 \cup M_3) \cap (M_2 \cup M_3)$.
- $\supseteq$: wir wollen zeigen, dass $(M_1 \cap M_2) \cup M_3 \supseteq (M_1 \cup M_3) \cap (M_2 \cup M_3)$. Sei $x \in (M_1 \cup M_3) \cap (M_2 \cup M_3)$, dann $x \in M_1 \cup M_3$ und $x \in M_2 \cup M_3$. Wenn $x \in M_3$, dann $x \in (M_1 \cap M_2) \cup M_3$ auch. Wenn $x \notin M_3$, wir wissen dass $x \in M_1 \cup M_3$, so dass $x \in M_1$. Wir wissen auch, dass $x \in M_2 \cup M_3$, so dass $x \in M_2$. Das zeigt dass $x \in M_1 \cap M_2$ und dann $x \in (M_1 \cap M_2) \cup M_3$.

$\square$

Lemma 1.2.10 (Eigenschaften des Komplementes)**.** *Seien M_1, M_2, M_3 zwei Mengen so dass $M_1 \subseteq M_3, M_2 \subseteq M_3$. Die Komplementen in M_3 haben die Eigenschaften:*

- $M^1 \subseteq M_2 \iff M_1^c \supseteq M_2^c$.
- $(M_1 \cup M_2)^c = M_1^c \cap M_2^c$.
- $(M_1 \cap M_2)^c = M_1^c \cup M_2^c$.

- $(M_1^c)^c = M_1$.

Beweis. Im Repetitorium. □

Definition 1.2.11 (Kartesisches Produkt)**.** Seien M_1, M_2 Mengen. Das kartesisches Produkt von M_1 und M_2 ist die Menge definiert als

$$M_1 \times M_2 \colon = \{(x_1, x_2) \,|\, x \in M_1, x \in M_2\}$$

Hier (x_1, x_2) ist ein geordnetes Paar: d.h. $(x_1, y_1) = (x_2, y_2)$ genau dann, wenn $x_1 = y_1, x_2 = y_2$. Falls $M = N$ schreiben wir $M \times M = M^2$.

Beispiel 1.2.12. Seien $M_1 = \{1, 2, 3\}$ und $M_2 = \{4, 5\}$. Dann

$$M_1 \times M_2 = \{(1,4), (1,5), (2,4), (2,5), (3,4), (3,5)\}.$$

Beispiel 1.2.13. Die Menge $\mathbb{Z}^2 = \{(n, m) \,|\, n, m \in \mathbb{Z}\}$ ist eine Teilmenge von $\mathbb{R}^2 = \{(x, y) \,|\, x, y \in \mathbb{R}\}$

1.3 Funktionen

Definition 1.3.1 (Funktion, Abbildung)**.** Eine Funktion oder Abbildung $f \colon X \to Y$ von einer Menge X zu einer Menge Y ist eine Zuordnung, die jedem Element $x \in X$ ein Element $y = f(x) \in Y$ zuordnet. Wir schreiben

$$f \colon X \to Y, \qquad x \mapsto f(x)$$

Die Menge X heißt Definitionsbereich, die Menge Y heißt Zielbereich oder Wertemenge

Bemerkung 1.3.2. das Definitionsbereich und das Zielbereich sind teil von der Definition einer Abbildung. Das bedeutet dass zwei Abbildungen

$$f \colon X \to Y, \qquad g \colon X' \to Y'$$

gleich sind, genau dann, wenn $X = X', Y = Y'$ und $f(a) = g(a)$ für alle $a \in X$. Z.B. die zwei Abbildungen

$$f \colon \mathbb{N} \to \mathbb{N}, x \mapsto x + 1 \qquad g \colon \mathbb{N} \to \mathbb{Z}, x \mapsto x + 1$$

sind verschiedene Abbildungen, trotzdem $f(x) = g(x)$ für alle $x \in \mathbb{N}$.

Beispiel 1.3.3. Die Funktion $f \colon \{1, 2, 3\} \to \{4, 7\}$ so dass $f(1) = 7, f(2) = 4, f(3) = 4$

Beispiel 1.3.4. Die Funktion $f \colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$.

Andere Beispiele von Funktionen:

Definition 1.3.5 (Konstante Funktion)**.** Eine Funktion $f \colon X \to Y$ heißt konstant, wenn $k \in Y$ existiert so dass $f(x) = k$ für alle $x \in X$.

Definition 1.3.6 (Identität)**.** Sei X eine Menge. Die Identität von X ist die Abbildung

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

Definition 1.3.7 (Einschränkung)**.** Seien $f\colon X \to Y$ eine Funktion und $A \subseteq X$ eine Teilmenge. Die Einschränkung von f auf A ist die Funktion

$$f_{|A}\colon A \to Y, \qquad a \mapsto f(a)$$

Definition 1.3.8 (Bild)**.** Sei $f\colon X \to Y$ eine Funktion und $A \subseteq X$ eine Teilmenge. Das Bild von A unter f ist

$$f(A)\colon = \{f(a) \mid a \in A\} = \{y \in Y \mid y = f(a) \text{ für ein } a \in A\}$$

Das Bild von X unter f ist auch Bild von f genannt:

$$\operatorname{Im} f = f(X)$$

Definition 1.3.9 (Urbild)**.** Sei $f\colon X \to Y$ eine Funktion und $B \subseteq Y$ eine Teilmenge. Das Urbild von B unter f ist

$$f^{-1}(B)\colon = \{x \in X \mid f(x) \in B\}$$

Wenn $B = \{b\}$ nur ein Element enthält, schreiben wir auch $f^{-1}(b) = f^{-1}(\{b\})$.

Beispiel 1.3.10. Sei $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$. Das Bild von f ist

$$f(\mathbb{R}) = \{x^2 \mid x \in \mathbb{R}\}$$

Wir sehen dass $f(\mathbb{R}) \subseteq \mathbb{R}_{\geq 0} = \{y \in \mathbb{R} \mid y \geq 0\}$ weil jeder Quadrat in $\mathbb{R}$ nicht negativ ist. Aber wir wissen auch dass jedes $y \in \mathbb{R}_{\geq 0}$ ein Wurzel hat: es gibt $x \in \mathbb{R}$ so dass $y = x^2 = f(x)$. Das zeigt, dass

$$f(\mathbb{R}) = \mathbb{R}_{\geq 0};$$

Wir können auch manche Urbilder explizit bestimmen:

$$\begin{aligned} f^{-1}(0) &= \{x \in \mathbb{R} \mid x^2 \in \{0\}\} = \{x \in \mathbb{R} \mid x^2 = 0\} = \{0\} \\ f^{-1}(4) &= \{x \in \mathbb{R} \mid x^2 \in \{4\}\} = \{x \in \mathbb{R} \mid x^2 = 4\} = \{-2, +2\} \\ f^{-1}([-2,-1]) &= \{x \in \mathbb{R} \mid -2 \leq x \leq -1\} = \emptyset. \end{aligned}$$

Definition 1.3.11 (Injektiv, surjektiv, bijektiv)**.** Eine Funktion $f\colon X \to Y$ heißt

- Injektiv: falls verschiedene Elemente von X verschiedene Bilder haben: d.h. für alle $x_1, x_2 \in X$:
  $$x_1 \neq x_2 \implies f(x_1) \neq f(x_2).$$
  Das ist äquivalent zu
  $$f(x_1) = f(x_2) \implies x_1 = x_2.$$
- Surjektiv: falls alle Elemente von Y, Bilder von Elementen von X sind:
  $$\text{Für alle } y \in Y, \text{ existiert } x \in X \text{ s.d. } y = f(x).$$
  Das ist äquivalent zu $f(X) = Y$.
- Bijektiv: falls f injektiv und surjektiv ist. Das bedeutet dass
  $$\text{für alle } y \in Y \text{ existiert genau ein } x \in X \text{ s.d. } y = f(x).$$

♣♣♣♣♣♣♣♣♣♣♣♣♣ **Ende Vorlesung 19.04.2024** ♣♣♣♣♣♣♣♣♣♣♣♣♣

Beispiel 1.3.12. Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist nicht injektiv, weil $f(1) = f(-1) = 1$ und ist nicht surjektiv, weil $f(x) = x^2 \geq 0$ für alle $x \in \mathbb{R}$ so dass kein $x \in \mathbb{R}$ existiert mit $f(x) = x^2 = -1$.

Die Abbildung $f_{|\mathbb{R}_{\geq 0}}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}, x \mapsto x^2$ ist injektiv aber nicht surjektiv.

Die Abbildung $f_{|\mathbb{R}_{\geq 0}}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, x \mapsto x^2$ ist injektiv und surjektiv, d.h. bijektiv.

Lemma 1.3.13. *Sei $f\colon X \to Y$ eine Funktion. Dann gilt*

(a) f ist injektiv, genau dann, wenn $f^{-1}(y)$ entweder 1 oder 0 Elemente enthält, für alle $y \in Y$.

(b) f ist surjektiv, genau dann, wenn $f^{-1}(y) \neq \emptyset$ für alle $y \in Y$.

(c) f ist bijektiv, genau dann, wenn $f^{-1}(y)$ genau ein Element enthält, für alle $y \in Y$.

Beweis. (a) ($\Longrightarrow$) Sei $y \in Y$. Wir wollen zeigen, dass $f^{-1}(y)$ höchstens ein Element enthält. Seien $x_1, x_2 \in f^{-1}(y)$. Dann $f(x_1) = y = f(x_2)$. Da f injektiv ist, das zeigt dass $x_1 = x_2$.

($\Longleftarrow$) Seien $x_1, x_2 \in X$ so dass $f(x_1) = f(x_2) = y$. Dann $x_1, x_2 \in f^{-1}(y)$. Da $f^{-1}(y)$ höchstens ein Element enthält, es muss sein dass $x_1 = x_2$. Das zeigt dass f injektiv ist.

(b) Die Funktion f ist surjektiv, genau dann, wenn für alle $y \in Y$ ein $x \in X$ existiert, so dass $f(x) = y$. Die Gleichung $f(x) = y$ bedeutet genau dass $x \in f^{-1}(y)$. Das zeigt dass f surjektiv ist, genau dann, wenn $x \in X$ existiert so dass $xinf^{-1}(y)$. Äquivalent, genau dann, wenn $f^{-1}(y) \neq \emptyset$ für alle $y \in Y$.

(c) Die Funktion f ist bijektiv, genau dann, wenn f injektiv und surjektiv ist. Punkt (a) zeigt dass, f injektiv ist, genau dann, wenn $f^{-1}(y)$ höchstens ein Element enthält, für alle $y \in Y$. Punkt (b) zeigt dass f surjektiv ist, genau dann, wenn $f^{-1}(y)$ mindestens ein Element enthält, für alle $y \in Y$. Dann, ist f bijektiv, genau dann, wenn $f^{-1}(y)$ genau ein Element enthält, für alle $y \in Y$.

□

Definition 1.3.14 (Komposition)**.** Seien $f\colon X \to Y$ und $g\colon Y \to Z$ zwei Abbildungen. Ihre Komposition $g \circ f$ ist die Abbildung

$$g \circ f\colon X \to Z, \qquad x \mapsto g(f(x))$$

Beispiel 1.3.15. Seien $f\colon \mathbb{R} \to \mathbb{R}$ und $g\colon \mathbb{R} \to \mathbb{R}$ durch $f(x) = x + 2, g(x) = x^2$. Dann

$$\begin{aligned} f \circ g\colon \mathbb{R} \to \mathbb{R}, &\quad x \mapsto x^2 + 2, \\ g \circ f\colon \mathbb{R} \to \mathbb{R}, &\quad x \mapsto (x+2)^2 \end{aligned}$$

Lemma 1.3.16. *(a) Sei $f\colon X \to Y$ eine Abbildung. Dann*

$$f \circ \mathrm{id}_X = f, \quad \mathrm{id}_Y \circ f = f.$$

(b) (Assoziativität) Seien $f\colon X \to Y, g\colon Y \to W, h\colon W \to Z$ Abbildungen. Dann gilt

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

Beweis. (a) Die zwei Abbildungen $f \circ \mathrm{id}_X$ und f haben Definitionsbereich X und Zielbereich X. Wir müssen zeigen dass $(f \circ \mathrm{id}_X)(x) = f(x)$ für alle $x \in X$. Aber $(f \circ \mathrm{id}_X)(x) = f(\mathrm{id}_X(x)) = f(x)$. Eine änliche Begrundung zeigt dass $\mathrm{id}_Y \circ f = f$.

(b) Die Zwei Abbildungen $h \circ (f \circ g)$ und $(h \circ f) \circ g$ haben Definitionsbereich X und Zielbereich Z. Sei $x \in X$: dann

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

□

Definition 1.3.17 (Invertierbare Funktion)**.** Eine Funktion $f\colon X \to Y$ heißt invertierbar oder umkehrbar, falls eine Funktion $g\colon Y \to X$ existiert so dass

$$(g \circ f) = \mathrm{id}_X, \qquad (f \circ g) = \mathrm{id}_Y$$

Die Funktion g heißt dann, Umkehrfunktion oder inverse Funktion von f. Wir schreiben auch f^{-1}.

Lemma 1.3.18. *Sei $f\colon X \to Y$ eine invertierbare Funktion. Dann ist die inverse Funktion eindeutig.*

Beweis. Seien $g_1\colon Y \to X$ und $g_2\colon Y \to X$ zwei inverse Funktionen von f. Dann gilt

$$g_1 = g_1 \circ \mathrm{id_Y} = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = \mathrm{id_X} \circ g_2 = g_2.$$

□

Bemerkung 1.3.19. Die Gleichung $f^{-1} \circ f = \mathrm{id}_X$ zeigt dass $f^{-1}(f(x)) = \mathrm{id}_X(x) = x$ für alle $x \in X$. Die Gleichung $f \circ f^{-1} = \mathrm{id}_Y$ zeigt dass $f(f^{-1}(x)) = \mathrm{id}_Y(y) = y$ für alle $y \in Y$.

Definition 1.3.20 (Inverse Funktion)**.** Wenn $f\colon X \to Y$ eine invertierbare Funktion ist, schreiben wir die eindeutige inverse Funktion als $f^{-1}\colon Y \to X$.

Bemerkung 1.3.21 (Vorsicht!). Sei $f\colon X \to Y$ eine Funktion. Das Urbild $f^{-1}(B)$ von einer Teilmenge $B \subseteq Y$ ist *immer* definiert, für *jede* Funktion. Die inverse Funktion $f^{-1}\colon Y \to X$ ist definiert *nur* wenn f invertierbar ist.

Wenn die Funktion invertierbar ist, stimmen diese beiden Begriffe wie folgt überein

Lemma 1.3.22. *Sei $f\colon X \to Y$ eine invertierbare Funktion mit inverse Funktion $f^{-1}\colon Y \to X$. Sei auch $B \subseteq Y$ eine Teilmenge. Das Urbild von B unter f, und das Bild von B unter f^{-1} sind gleich.*

Beweis. Wir schreiben die inverse Abbildung als $g\colon Y \to X$, um Verwirrung zwischen das Urbild und die inverse Funktion zu vermeiden. Wir wollen zeigen dass, das Bild $g(B)$ und das Urbild $f^{-1}(B)$ gleich sind:

$$g(B) = f^{-1}(B)$$

Wir zeigen zuerst dass $g(B) \subseteq f^{-1}(B)$: sei $y \in B$, dann $f(g(y)) = (f \circ g)(y) = \mathrm{id}_Y(y) = y$. Das bedeutet dass $g(y)$ im Urbild $f^{-1}(B)$ ist.

Wir zeigen jetzt dass $g(B) \supseteq f^{-1}(B)$: sei $x \in f^{-1}(B)$, sodass $y = f(x) \in B$. Dann $x = \mathrm{id}_X(x) = (g \circ f)(x) = g(f(x)) = g(y)$. Das zeigt dass $x \in g(B)$. □

Satz 1.3.23. *Eine Funktion $f\colon X \to Y$ ist invertierbar, genau dann, wenn f bijektiv ist.*

Beweis. Wir zeigen die zwei Implikationen:

($\Longrightarrow$) Wir zeigen dass, wenn f invertierbar ist, dann f auch bijektiv ist. Sei $f^{-1}\colon Y \to X$ die inverse Funktion von f.Wir zeigen zuerst dass f injektiv ist: seien $x_1, x_2 \in X$ so dass $f(x_1) = f(x_2)$. Dann $f^{-1}(f(x_1)) = f^{-1}(f(x_2)$. Da $f^{-1}(f(x)) = x$ für alle $x \in X$, sehen wir dass $x_1 = x_2$. Wir zeigen dass f surjektiv ist: sei $y \in Y$. Dann $f^{-1}(y) \in X$ und $f(f^{-1}(y)) = y$.

($\Longleftarrow$) Sei f bijektiv, und sei $y \in Y$. Lemma 1.3.13 shows that das Urbild $f^{-1}(y)$ genau ein Element enthält, wir nennen es $g(y)$. Das definiert eine Funktion $g\colon Y \to X$. Wir wollen zeigen, dass $f \circ g = \mathrm{id}_Y$ und $g \circ f = \mathrm{id}_X$. Sei $y \in Y$, dann $g(y) \in f^{-1}(y)$ so dass $f(g(y)) = y$. Das zeigt dass $f \circ g = \mathrm{id}_Y$. Sei $x \in X$, dann $x \in f^{-1}(f(x))$ so dass $x = g(f(x))$. Das zeigt, dass $g \circ f = \mathrm{id}_X$. □

1.4 Relationen

Siehe Analysis I.

1.5 Vollständige Induktion

Das Prinzip der Vollständige Induktion ist einfach, aber sehr wirkungsvoll

Prinzip der Vollständige Induktion. *Für jedes $n \in \mathbb{N}$ sei $\mathcal{A}(n)$ eine Aussage. Angenommen dass:*

- ***Induktionsanfang:*** *$\mathcal{A}(0)$ wahr ist,*
- ***Induktionsschritt:*** *für jedes $n \in \mathbb{N}$, wenn $\mathcal{A}(n)$ wahr ist, dann ist $\mathcal{A}(n+1)$ auch wahr:*

$$\mathcal{A}(n) \implies \mathcal{A}(n+1) \qquad \text{für alle } n \in \mathbb{N},$$

dann ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$. In dem Induktionsschritt die Annahme dass $\mathcal{A}(n)$ wahr ist heißt ***Induktionsvoraussetzung****.*

Wir zeigen jetzt mehrere Anwendungen und Beispielen:

Beispiel 1.5.1 (Summe der erste n Zahlen)**.** Wir zeigen dass

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}, \qquad \text{für alle } n \in \mathbb{N},$$

durch Induktion.

- **Induktionsanfang:** Für $n = 0$ die linke Seite ist 0, weil eine Summe von null Elemente, Null ist. Die rechte Seite ist auch 0, weil $\frac{0 \cdot 1}{2} = 0$. Die Aussage gilt für $n = 0$.

- **Induktionsschritt:** Nehmen wir an, dass die Aussage für n gilt. Wir wollen zeigen dass die Aussage für $n+1$ auch gilt:

$$\begin{aligned}1+2+\cdots+n+(n+1) &= (1+2+\cdots+n)+(n+1)\\ &= \frac{n(n+1)}{2}+(n+1) \qquad (\textbf{Induktionsvoraussetzung})\\ &= \frac{n(n+1)+2(n+1)}{2} = \frac{(n+2)(n+1)}{2}\end{aligned}$$

  und das ist genau die Aussage für $n+1$.

Bemerkung 1.5.2. Es gibt Varianten des Prinzips der Vollständige Induktion. Z.B. für jedes $n \in \mathbb{N}$ sei $\mathcal{A}(n)$ eine Aussage. Angenommen dass

- $\mathcal{A}(n_0)$ gilt für ein $n_0 \in \mathbb{N}$,
- $\mathcal{A}(n) \implies \mathcal{A}(n+1)$ für alle $n \geq n_0$,

dann gilt $\mathcal{A}(n)$ für alle $n \geq n_0$.

Satz 1.5.3 (Bernoullische Ungleichung)**.** *Seien $x \in \mathbb{R}, x \geq 1$ und $n \in \mathbb{N}$. Dann*

$$(1+x)^n \geq 1+nx.$$

Ist zusätzlich $n \geq 2$ und $x \neq 0$, so gilt sogar

$$(1+x)^n > 1+nx.$$

Beweis. Wenn $n=0$ oder $n=1$, es ist klar dass $(1+x)^n = 1+nx$. Das ist auch klar, wenn $x=0$. Wir müssen dann zeigen dass

$$(1+x)^n \geq 1+nx \qquad \text{für alle } x \geq -1, x \neq 0, \text{ und } n=2$$

und wir zeigen das durch Induktion:

- **Induktionsanfang:** Für $n=2$ gilt

$$(1+x)^2 = 1+2x+x^2 > 1+2x$$

  weil $x^2 > 0$ (hier ist wichtig dass $x \neq 0$).

- **Induktionsschritt:** Angenommen dass die Ungleichung für ein $n \geq 2$ gilt, wollen wir zeigen dass die für $n+1$ auch gilt:

$$(1+x)^{n+1} = (1+nx)^n(1+x) \geq (1+nx)(1+x) = 1+x+nx+nx^2 > 1+(n+1)x.$$

  Hier haben wir die Induktionsvoraussetzung $(1+nx)^n \geq 1+nx$ benutzt.

$\square$

1.6 Mächtigkeit

Beispiel 1.6.1. Wir betrachten die zwei Mengen $S = \{$ Studierende im Horsaal $\}$ und $P = \{$ Plätze im Horsaal $\}$. Wann haben diese zwei Mengen die gleiche Anzahl von Elemente? Wir können entweder die Studierende und die Plätze Anzahlen. Oder wir können auch sagen dass jede Studierende ein Platz hat, zwei Studierende sitzen nicht am gleichen Platz, und dass alle Platze voll sind. Mathematisch gesagt, es gibt eine bijektive Funktion

$$p\colon S \to P, \qquad s \mapsto p(s) = \text{ platz von Studierende } s$$

Dieses kleines Beispiel fuhrt zu einen sehr wichtiges Konzept in der Mathematik:

Definition 1.6.2 (Gleichmächtige Menge)**.** Zwei Mengen X, Y sind gleichmächtig wenn es eine bijektive Funktion $f\colon X \to Y$ existiert.

Lemma 1.6.3. *Zwei Mengen X, Y sind gleichmächtig, genau dann, wenn sie gleich viele Elemente besitzen.*

Beweis. Wir zeigen die zwei Implikationen:

($\Longrightarrow$) Sei $f\colon X \to Y$ eine bijektive Abbildung. Wir schreiben $X = x_1, \dots, x_n$, mit paarweise verschiedene x_i, so dass X genau n Elemente hat. Da f surjektiv ist, $Y = f(X) = \{f(x_1), \dots, f(x_n)\}$, und da f injektiv ist, die $f(x_i)$ sind auch paarweise verschiedene. Das bedeutet dass Y genau n Elemente hat.

($\Longleftarrow$) Wenn X und Y beide n Elemente haben, dann $X = \{x_1, \dots, x_n\}$ mit paarweise verschiedene x_i und $Y = \{y_1, \dots, y_n\}$ mit paarweise verschiedene y_i. Wir definieren zwei Funktionen

$$f\colon X \to Y, \quad x_i \mapsto y_i, \quad g\colon y_i \to x_i$$

und wir sehen dass $f \circ g = \mathrm{id}_Y$ und $g \circ f = \mathrm{id}_X$. Das bedeutet dass f invertierbar ist, und deswegen bijektiv. $\square$

Definition 1.6.4 (Mächtigkeit)**.** Sei M eine endliche Menge. Der Anzahl von Elemente in M ist auch die **Machtigkeit** oder Kardinalität von M genannt. Man schreibt $|M|$ für die Mächtigkeit.

Eine Menge M heißt **abzählbar unendlich**, wenn M gleichmächtig zu $\mathbb{N}$ ist.

Eine Menge M heißt **überabzählbar** wenn sie weder endlich oder abzählbar unendlich ist.

Beispiel 1.6.5. Die Menge $A = \{1, \sqrt{2}, \pi\}$ hat 3 Elemente, so dass $|A| = 3$.

Beispiel 1.6.6. In der Mengenlehre zeigt man dass $\mathbb{Z}$ und $\mathbb{Q}$ unendlich abzählbar sind, und dass $\mathbb{R}$ unendlich überabzählbar ist.

1.6.1 Fakultät und Binomialkoeffizienten

Definition 1.6.7 (Fakultät)**.** Sei $n \in \mathbb{N}$. Das Fakultät $n!$ (sprich "n Fakultät") ist definiert als

$$n! := \begin{cases} 1, & \text{falls } n = 0, \\ 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n, & \text{falls } n > 0 \end{cases}$$

Bemerkung 1.6.8. Eine nützliche Konvention in der Mathematik ist dass die leere Summe 0 ist und dass das leeres Produkt 1 ist. Dann können wir auch schreiben

$$n! = n \cdot (n-1) \cdot \dots \cdot 1 \qquad \text{für alle } n \in \mathbb{N}$$

weil wenn $n = 0$, das Produkt an der rechte Seite ein leeres Produkt ist, und ein leeres Produkt ist 1.

Definition 1.6.9 (Binomialkoeffizient)**.** Seien $n, k \in \mathbb{N}$. Das Binomialkoeffizient $\binom{n}{k}$ (sprich "n über k" oder "k aus n") ist definiert als

$$\binom{n}{k} := \begin{cases} 0, & \text{falls } k > n. \\ \dfrac{n!}{k!(n-k)!}, & \text{falls } k \leq n. \end{cases}$$

Lemma 1.6.10 (Eigenschaften des Binomialkoeffizientes)**.** *Seien $n, k \in \mathbb{N}$ mit $k \leq n$. Dann gilt:*

(a) $\binom{n}{0} = 1$ *und* $\binom{n}{1} = n$ *für alle* n.

(b) $\binom{n}{k} = \dfrac{n(n-1) \cdot \ldots \cdot (n-k+1)}{k!}$

(c) $\binom{n}{k} = \binom{n}{n-k}$.

(d) $\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}$.

Beweis. (a) Wir berechnen

$$\binom{n}{0} = \frac{n!}{0!n!} = \frac{n!}{1 \cdot n!} = 1, \qquad \binom{n}{1} = \frac{n!}{1!(n-1)!} = \frac{n}{1} = n.$$

(b) Wir berechnen

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n-k)! \cdot (n-k+1) \cdot \ldots \cdot (n-1)n}{k!(n-k)!} = \frac{(n-k+1) \cdot \ldots \cdot (n-1) \cdot n}{k!}.$$

(c) Wir berechnen

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-(n-k))!(n-k)!} = \binom{n}{n-k}.$$

(d) Wir berechnen

$$\begin{aligned}
\binom{n}{k+1} + \binom{n}{k} &= \frac{n!}{(k+1)!(n-k-1)!} + \frac{n!}{k!(n-k)!} \\
&= \frac{n!}{(k+1)k!(n-k-1)!} + \frac{n!}{k!(n-k-1)!(n-k)} = \\
&= \frac{n!}{k!(n-k-1)!} \cdot \left(\frac{1}{k+1} + \frac{1}{n-k}\right) \\
&= \frac{n!}{k!(n-k-1)!} \cdot \frac{n-k+k+1}{(k+1)(n-k)} = \frac{n!}{k!(n-k-1)!} \cdot \frac{n+1}{(k+1)(n-k)} \\
&= \frac{(n+1)!}{(k+1)!(n-k)!} = \binom{n+1}{k+1}.
\end{aligned}$$

□

Bemerkung 1.6.11. Die letze Gleichung im vorherigen Lemma lässt sich gut anhand des Pascalsches Dreieck veranschaulichen. Siehe Tafel oder Wikipedia: `https://de.wikipedia.org/wiki/Pascalsches_Dreieck`.

Binomialkoeffizienten und Faktorzahlen sind in der Kombinatorik allgegenwärtig.Ganz informell ist dies der Teil der Mathematik, der sich mit dem Abzählen von Objekten beschäftigt. Wir werden hier nicht viele Objekte zählen, aber wir können ein sehr einfaches, aber sehr wichtiges Prinzip aufstellen

Lemma 1.6.12. *Sei $f\colon X \to Y$ eine Funktion, dann*

$$X = \bigsqcup_{y \in Y} f^{-1}(y).$$

Das heißt: $X = \cup_{y\in Y} f^{-1}(y)$ und die $f^{-1}(y)$ sind paarweise disjunkt.

Beweis. Wir zeigen zuerst dass $X = \bigcup_{y\in Y} f^{-1}(y)$: seien $x \in X$ und $y = f(x)$. Dann $x \in f^{-1}(y)$. Wir zeigen jetzt, dass die $f^{-1}(y)$ paarweise disjunkt sind: sei $x \in f^{-1}(y) \cap f^{-1}(y')$, dann $y = f(x) = y'$. □

Korollar 1.6.13. *Sei $f\colon X \to Y$ eine Funktion zwischen endliche Menge. Dann*

$$|X| = \sum_{y\in Y} |f^{-1}(y)|.$$

Beweis. Lemma 1.6.12 zeigt dass $X = \sqcup_{y\in Y} f^{-1}(y)$, sodass $|X| = \sum_{y\in Y} |f^{-1}(y)|$. □

Korollar 1.6.14. *Seien X, Y endliche Menge. Dann $|X \times Y| = |X| \times |Y|$.*

Beweis. Wir betrachten die Abbildung $\mathrm{p}_X\colon X \times Y \to X$. Für jedes $x \in X$, gilt $|f^{-1}(x)| = |\{(x,y) \mid y \in Y\}| = |Y|$.

$$|X \times Y| = \sum_{x\in X} |p_X^{-1}(x)| = \sum_{x\in X} |Y| = |X| \cdot |Y|$$

□

Proposition 1.6.15. *Sei A eine endliche Menge mit $|A| = n$ und sei $k \in \mathbb{N}, k \geq 1$. Die Anzahl von k-Tupeln,*

$$(a_1, a_2, \dots, a_k), \qquad a_i \in A, \quad a_i \textit{ paarweise verschiedene,}$$

ist $n \cdot (n-1) \cdot \dots \cdot (n-k+1)$.

Beweis. Sei $T_k(A) = \{(a_1, \dots, a_k) \,|\, a_i \in A, a_i \text{ paarweise verschiedene } \}$. Wir wollen zeigen, dass

$$|T_k(A)| = n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$$

Wenn $k > n$, dann $|T_k(A)| = 0$, weil $T_k(A) = \emptyset$ und $n \cdot (n-1) \dots \cdot (n-k+1) = 0$, weil eine der Faktoren gleich 0 ist. Wir müssen dann die Aussage beweisen für alle $1 \leq k \leq n$. Wir zeigen sie durch Induktion auf k:

- **Induktionsanfang:** Wenn $k = 1$, dann hat $T_1(A) = \{(a_1) \,|\, a_1 \in A\}$ genau n Elemente.
- **Induktionsschritt:** Wir nehmen an, dass die Aussage fur $1 \leq k \leq n-1$ gilt und wir wollen zeigen dass es für $k+1$ gilt. Wir betrachten die Funktion

  $$f \colon T_{k+1}(A) \to T_k(A), \quad (a_1, \dots, a_{k+1}) \mapsto (a_1, \dots, a_k)$$

  (Warum ist diese Funktion wohldefiniert?). Korollar 1.6.13 zeigt dass

  $$|T_{k+1}(A)| = \sum_{t \in T_k(A)} |f^{-1}(t)|$$

  Sei $t = (a_1, \dots, a_k) \in T_k(A)$. Dann $f^{-1}(t) = \{(a_1, \dots, a_k, a_{k+1}) \,|\, a_{k+1} \notin A \setminus \{a_1, \dots, a_k\}\}$, so dass $|f^{-1}(t)| = |A \setminus \{a_1, \dots, a_k\}| = (n-k)$. Dann haben wir

  $$|T_{k+1}(A)| = \sum_{t \in T_k(A)} (n-k) = |T_k(A)| \cdot (n-k)$$

  und die **Induktionsvoraussetzung** zeigt dass $|T_k(A)| = n(n-1) \cdot (n-k+1)$. Am Ende haben wir

  $$|T_{k+1}(A)| = n(n-1) \cdot \dots \cdot (n-k)$$

  und das ist genau was wir zeigen wollten.

$\square$

Proposition 1.6.16. *Sei A eine endliche Menge mit n Elemente. Dann A hat genau $\binom{n}{k}$ Teilmenge mit k Elemente.*

Beweis. Für eine endliche Menge A schreibt man

$$\binom{A}{k} = \{B \subseteq A \,|\, B \text{ Teilmenge mit } k \text{ Elemente } \} = \{B \subseteq A \,|\, |B| = k\}\}$$

Wir wollen zeigen dass

$$\left|\binom{A}{k}\right| = \binom{n}{k}.$$

Wenn $k > n$, dann sind beide gleich 0. Wenn $k \leq n$, betrachten wir die Funktion

$$f\colon T_k(A) \to \binom{A}{k}, \quad (a_1, \ldots, a_k) \mapsto \{a_1, \ldots, a_k\}$$

(Warum ist die Funktion wohldefiniert?). Wir wollen Korollar 1.6.13 benutzen: sei $B \subseteq A$ eine Teilmenge mit $|B| = k$ Elemente. Dann $f^{-1}(B) = \{(b_1, \ldots, b_k) \mid b_i \in B \text{ paarweise verschiedene }\} = T_k(B)$, und Proposition 1.6.15 zeigt dass $|T_k(B)| = k!$. Dann Korollar 1.6.13 zeigt dass

$$|T_k(A)| = \sum_{B \in \binom{A}{k}} |f^{-1}(B)| = \sum_{B \in \binom{A}{k}} k! = \left|\binom{A}{k}\right| \cdot k!,$$

und dann

$$\begin{aligned}
\left|\binom{A}{k}\right| &= \frac{|T_k(A)|}{k!} \\
&= \frac{n(n-1)\ldots(n-k+1)}{k!} && \text{(Prop. 1.6.15)} \\
&= \binom{n}{k} && \text{(Lemma 1.6.10.)}
\end{aligned}$$

□

Bemerkung 1.6.17. Dieser Satz kann suggestiv wie folgt formuliert werden: sei A eine endliche Menge. Dann

$$\left|\binom{A}{k}\right| = \binom{|A|}{k}.$$

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 23.04.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Kapitel 2

Matrizen und das Gaußsches Verfahren

Wir haben zwei Beispiele von lineare Gleichungssysteme schon in der erste Vorlesung gesehen:

$$\begin{cases} h+k & = 40, \\ 2h+4k & = 120. \end{cases}, \qquad \begin{cases} 2x+3y-z+4w-5v=10, \\ x-2y+3z-w+2v=5, \\ 3x+2y-z+w-4v=-3, \\ 4x-y+2z+3w+v=8 \end{cases}$$

Man kann die einzige Lösung des erstes Systems relativ einfach berechnen: $h = 30, k = 10$. Was ist das bestes Weg, um das zweites System zu lösen? Gibt's uberhaupt Lösungen? In diesem Teil der Vorlesung, werden wir lineare Gleichungssysteme systematisch betrachten.

In diesem Kapitel werden wir das Symbol $\mathbb{K}$ für $\mathbb{R}$ verwenden, was bedeutet, dass $\mathbb{K} = \mathbb{R}$. Nachdem wir definiert haben, was ein Körper ist, werden wir sehen, dass die Ergebnisse des Kapitels für ein beliebiges Körper $\mathbb{K}$ gelten.

2.1 Gleichungssysteme, Vektoren, Matrizen

Die Gleichungssysteme oben sind bestimmt von den Koeffizienten von der Unbekannte und von den Werten das die Gleichungen. Wir organisieren diese Daten in Vektoren und Matrizen:

Definition 2.1.1 (Vektor)**.** Ein **Spaltenvektor** oder einfach **Vektor** in $\mathbb{K}^{m\times 1}$ ist ein Schema

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

mit **Koeffizienten** $b_i \in \mathbb{K}$. Die Koeffizienten b_i heißen auch **Koordinaten** von b.

Ein **Zeilenvektor** oder manchmal **Kovektor** in $\mathbb{R}^{1\times n}$ ist ein Schema

$$a = \begin{pmatrix} a_1 & a_2 & \dots & a_m \end{pmatrix}$$

mit Koeffizienten $a_j \in \mathbb{K}$. mit

Bemerkung 2.1.2. Beide Vektoren und Kovektoren sind Tupeln von Elementen in $\mathbb{K}$. Wir schreiben die Vektoren vertikal und die Kovektoren horizontal. In der Mathematik ist es üblich, die Elemente in $\mathbb{K}^m$ als Spaltenvektoren zu betrachten. **Ab jetzt, alle Elementen in $\mathbb{K}^m$ sind für uns Spaltenvektoren**:

$$b \in \mathbb{K}^m \iff b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \text{ mit } b_i \in \mathbb{K}.$$

Anders gesagt, $\mathbb{K}^m = \mathbb{K}^{m \times 1}$. Die Menge von Reihevektoren schreiben wir weiter als $\mathbb{K}^{1 \times n}$.

Die **Summe** von zwei Vektoren $b, c \in \mathbb{K}^m$ ist definiert als

$$b + c = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} b_1 + c_1 \\ \vdots \\ b_m + c_m \end{pmatrix}$$

Fur jedes $\lambda \in \mathbb{K}$ und jeden Vektor $b \in \mathbb{K}^n$, definieren wir das Produkt $\lambda \cdot b$ als

$$\lambda \cdot b = \lambda \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} \lambda \cdot b_1 \\ \vdots \\ \lambda \cdot b_m \end{pmatrix}$$

Wir können auch die analogen Operationen für Zeilenvektoren definieren.

Definition 2.1.3 (Matrix)**.** Eine $m \times n$ Matrix mit Koeffizienten in $\mathbb{K}$ ist eine Schema $A \in \mathbb{K}^{m \times n} = \mathrm{Mat}(m \times n, \mathbb{K})$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \qquad a_{ij} \in \mathbb{K}$$

mit m Zeilen und n Spalten.

Bemerkung 2.1.4. Vektoren sind Matrizen mit nur eine Spalte und Zeilenvektoren sind Matrizen mit nur eine Zeile.

Definition 2.1.5 (Matrix-Vektor-Produkt)**.** Seien $A \in \mathbb{K}^{m \times n}$ eine Matrix und $c \in \mathbb{K}^n$ ein Vektor. Wir definieren das Produkt $A \cdot c$ als

$$A \cdot c = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a_{11} \cdot c_1 + a_{12} \cdot c_2 + \dots + a_{1n} \cdot c_n \\ a_{21} \cdot c_1 + a_{22} \cdot c_2 + \dots + a_{2n} \cdot c_n \\ \vdots \\ a_{m1} \cdot c_1 + a_{m2} \cdot c_2 + \dots + a_{mn} \cdot c_n \end{pmatrix}$$

Bemerkung 2.1.6. Um das Produkt $A \cdot c$ zu definiere, die Matrix A muss n Spalten haben und der Vektor c muss n Reihen haben: die Anzahl von Spalten von A und die Anzahl von Zeilen von c sind gleich. Z.B., das Produkt von einer 2×4 Matrix mit einem 3×1-Vektor ist nicht definiert.

Definition 2.1.7 (Lineares Gleichungssystem)**.** Seien $A = (a_{ij}) \in \mathbb{K}^{m\times n}$ eine Matrix und $b = (b_i) \in \mathbb{K}^m$ ein Vektor. Das lineares Gleichungssystem (LGS) von A und b ist besteht auf m lineare Gleichungen mit n Unbekannte oder Variablen:

$$LGS(A,b)\colon \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

mit $a_{ij}, b_i \in \mathbb{K}$. Wir köennen dieses System auch als

$$A \cdot x = b, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

schreiben. Die Matrix A heißt **Koeffizientenmatrix** des Systems und die Matrix $(A|b) \in \mathbb{K}^{m\times(n+1)}$ heißt **erweiterte Koeffizientenmatrix**. Das LGS heißt **homogen** falls $b = 0$ und **inhomogen** falls $b \neq 0$.

Die **Menge von Lösungen** des Systems ist die Menge von alle Vektoren $x \in \mathbb{K}^n$ die alle Gleichungen erfüllen:

$$\mathrm{Los}(A,b) = \{x \in \mathbb{K}^n \mid Ax = b\}.$$

Bemerkung 2.1.8. Ein homogenes system hat immer die Lösung $x = 0$: $A \cdot 0 = 0$ für alle $A \in \mathbb{K}^{m\times n}$.

2.2 Die Zeilenstufenform

Beispiel 2.2.1. Wir wollen alle Lösungen des folgendes LGS bestimmen:

$$\begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_3 + 5x_5 & = -10 \\ x_4 - x_5 & = 2 \end{cases}. \tag{2.2.1}$$

Das ist sehr einfach: dieses System ist äquivalent zu

$$\begin{cases} x_1 & = -x_2 - 3x_5 - 6 \\ x_3 & = -5x_5 - 10 \\ x_4 & = x_5 + 2 \end{cases}$$

so dass die Menge von Lösungen des LGSs ist:

$$\mathrm{Los} = \left\{ \begin{pmatrix} -x_2 - 3x_5 - 6 \\ x_2 \\ -5x_5 - 10 \\ x_5 + 2 \\ x_5 \end{pmatrix} \mid x_2, x_5 \in \mathbb{K} \right\}$$

. Das nennt man eine **Parametrisierung** der Lösungsmenge: die Lösungen sind parametrisiert von den “freie” Variabeln x_2, x_5.

Beispiel 2.2.2. Wir betrachten jetzt das LGS

$$\begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_1 + x_2 + x_3 + 8x_5 & = -16 \\ x_1 + x_2 + x_3 + 2x_4 + 6x_5 & = -12 \end{cases}$$

Wir können dieses System nicht sofort lösen. Wenn wir jedoch die erste Gleichung von der zweiten und dritten Gleichung subtrahieren, erhalten wir

$$\begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_1 + x_2 + x_3 + 8x_5 & = -16 \\ x_1 + x_2 + x_3 + 2x_4 + 6x_5 & = -12 \end{cases} \iff \begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_3 + 5x_5 & = -10 \\ x_3 + 2x_4 + 3x_5 & = -6 \end{cases}$$

Dies ist eine Äquivalenz, denn um vom zweiten System zum ersten überzugehen, genügt es, die erste Gleichung mit der zweiten und der dritten zu addieren. Nun können wir Nun können wir die zweite Gleichung von der dritten subtrahieren (was ebenfalls eine umkehrbare Operation ist) und erhalten

$$\begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_3 + 5x_5 & = -10 \\ x_3 + 2x_4 + 3x_5 & = -6 \end{cases} \iff \begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_3 + 5x_5 & = -10 \\ 2x_4 - 2x_5 & = 4 \end{cases}$$

Nun können wir die dritte Gleichung mit $\frac{1}{2}$ multiplizieren (ebenfalls eine umkehrbare Operation) und erhalten

$$\begin{cases} x_1 + x_2 + 3x_5 & = -6 \\ x_3 + 5x_5 & = -10 \\ 2x_4 - 2x_5 & = 4 \end{cases} \iff \begin{cases} x_1 + x_2 - x_5 & = -4 \\ x_3 + 5x_5 & = -12 \\ x_4 - x_5 & = 2 \end{cases}$$

Wir haben das System (2.2.1) wieder gefunden. Wir hatten also ein erstes System, das sehr einfach zu lösen war, und ein weiteres System, das wir als äquivalent zum ersten System zeigen konnten. In diesem Kapitel werden wir sehen, dass jedes lineare System einem äquivalenten System entspricht, das sehr einfach zu lösen ist.

Die Tatsache, dass das erste lineare System sehr einfach zu lösen ist, ist auf seine Form zurückzuführen: die erweiterte Koeffizientenmatrix ist in reduzierter Zeilenstufenform:

$$(A \mid b) = \begin{pmatrix} 1 & 1 & 0 & 0 & 3 & -6 \\ 0 & 0 & 1 & 0 & 5 & -10 \\ 0 & 0 & 0 & 1 & -1 & 2 \end{pmatrix},$$

Definition 2.2.3 (Zeilenstufenform)**.** Eine Matrix $A = (a_{ij}) \in \mathbb{K}^{m \times n}$ hat **Zeilenstufenform**, falls es Zahlen r mit $0 \le r \le n$ und $j_1, j_2, \ldots, j_r$ mit $1 \le j_1 < j_2 < \cdots < j_r \le n$ gibt, so daß:

(a) $a_{ij} = 0$ für alle $i \le r$ und $j < j_i$.

(b) $a_{ij} = 0$ für alle $i > r$ und alle j.

(c) $a_{ij_i} \neq 0$ für alle $i = 1, \ldots, r$.

Die Zahl r heißt der **Rang** der Matrix in Zeilenstufenform: wir schreiben $r = \mathrm{rk}(A)$. die Zahlen $a_{ij_i} \neq 0$ heißen **Pivots**.

$$A = \begin{pmatrix}
0 & \dots & 0 & a_{1j_1} & * & * & * & * & \dots & * & * & * & \dots & * & * & \dots & * \\
0 & \dots & 0 & 0 & \dots & 0 & a_{2j_2} & * & \dots & * & * & * & \dots & * & * & \dots & * \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \dots & 0 & a_{3j_3} & * & \dots & * & \dots & * \\
\vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & a_{rj_r} & \dots & * \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0
\end{pmatrix}$$

Die Matrix A hat **reduzierte Zeilenstufenform** wenn sie Zeilenstufenform hat, und außerdem

(d) alle Pivots sind 1: $a_{ij_i} = 1$ für alle $i = 1, \dots, r$.

(e) alle Koeffizienten oben die Pivots sind 0: $a_{hj_i} = 0$ für alle $i = 1, \dots, r$ und $h < i$.

$$A = \begin{pmatrix}
0 & \dots & 0 & 1 & * & * & 0 & * & \dots & * & * & 0 & \dots & * & 0 & \dots & * \\
0 & \dots & 0 & 0 & \dots & 0 & 1 & * & \dots & * & * & 0 & \dots & * & 0 & \dots & * \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \dots & 0 & 1 & * & \dots & 0 & \dots & * \\
\vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \dots & * \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0
\end{pmatrix}$$

Bemerkung 2.2.4. Jetz haben wir der Rang definiert nur für eine Matrix in Zeilenstufenform. Wir werden später der Rang für alle Matrizen definieren.

Proposition 2.2.5 (Lösungen von LGS in reduzierte Zeilenstufenform)**.** *Wir betrachten ein LGS* $\{Ax = b\}$ *mit erweiterte Koeffizientenmatrix* $(A|b)$ *in reduzierte Zeilenstufenform.*

(a) Die Koeffizientenmatrix A hat reduzierte Zeilenstufenform, und die Pivots von A sind auch Pivots von $(A|b)$. Insbesondere $\mathrm{rk}(A) \leq \mathrm{rk}(A|b) \leq \mathrm{rk}(A) + 1$.

(b) Wenn $\mathrm{rk}(A) \neq \mathrm{rk}(A|b)$, *dann* $\mathrm{Los}(A, b) = \emptyset$.

(c) Wenn $\mathrm{rk}(A) = \mathrm{rk}(A|b)$, *dann hat* $\mathrm{Los}(A, b)$ *eine Parametrisierung, bei der die Variablen, die den Spalten der Pivots entsprechen, durch die Variablen der anderen Spalten ausgedrückt werden können.*

Beweis. (a) Hausaufgabe.

(b) Wenn $\mathrm{rk}(A) \neq \mathrm{rk}(A|b)$, dann $\mathrm{rk}(A|b) = \mathrm{rk}(A) + 1$. Sei $r = \mathrm{rk}(A)$. Dann ist die $(r+1)$-te Zeile von A Null aber die $(r+1)$-te Zeile von $(A|b)$ ist nicht Null. Dann ist die $(r+1)$-te Gleichung

$$0 \cdot x_1 + \cdots + 0 \cdot x_n = b_{r+1}$$

mit $b_{r+1} \neq 0$. Diese Gleichung hat keine Lösungen.

(c) Wenn $\mathrm{rk}(A) = \mathrm{rk}(A|b)$, dann sind alle Pivots von $(A|b)$ auch Pivots von A. Sei $A \in \mathbb{K}^{m\times n}$ und seien $a_{1j_1}, \ldots, a_{rj_r}$ die Pivots von A, mit $1 \le j_1 < \cdots < j_r \le n$. Das LGS von $(A|b)$ ist:

$$\begin{cases} a_{1j_1}x_{j_1} + \sum_{k=j_1+1}^{n} a_{1k}x_k & = b_1 \\ a_{2j_2}x_{j_2} + \sum_{k=j_2+1}^{n} a_{2k}x_k & = b_2 \\ \vdots \\ a_{rj_r}x_{j_r} + \sum_{k=j_r+1}^{n} a_{rk}x_k & = b_r \end{cases}$$

Da $(A|b)$ reduzierte Zeilenstufenform hat, $a_{ij_i} = 1$ für alle $i = 1, \ldots, r$, und $a_{hj_i} = 0$ für alle $i = 1, \ldots, r$ und $h < i$. Sei $F = \{\{1, \ldots, n\} \setminus \{j_1, j_2, \ldots, j_r\}$. Wir können das LGS wie folgt schreiben:

$$\begin{cases} x_{j_1} & = -\sum_{k\in F}^{n} a_{1k}x_k + b_1 \\ x_{j_2} & = -\sum_{k\in F}^{n} a_{2k}x_k + b_2 \\ \vdots \\ x_{j_r} & = -\sum_{k\in F}^{n} a_{rk}x_k + b_r \end{cases}$$

und das ist eine Parametrisierung der Menge der Lösungen $\mathrm{Los}(A, b)$.

□

Bemerkung 2.2.6. Diese Proposition zeigt, dass es sehr einfach ist, die Lösung eines Systems in reduzierter Zeilenstufenform zu parametrisieren. Wenn wir also ein lineares System in reduzierte Zeilenstufenform bringen können, können wir es auch lösen. Dies ist immer möglich,und wir sehen das in der nächsten Sektion.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 26.04.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

2.3 Elementare Zeilenunformungen und das Gauß-Verfahren

Definition 2.3.1 (Elementare Zeilenumformungen)**.** Die elementare Zeilenumformungen in einer Matrix sind:

(GZ_1) Vertauschen zweier Zeilen.

(GZ_2) Addition des λ-fachen einer Zeile zu einer andere, verschiedene Zeile, $\lambda \in \mathbb{K}$.

(GZ_3) Multiplikation einer Zeile mit $\lambda \in \mathbb{K}, \lambda \neq 0$.

Proposition 2.3.2. *Entsteht* $(A'|b')$ *aus* $(A|b)$ *durch endlich viele elementare Zeilenunformungen, so ist*

$$\mathrm{Los}(A, b) = \mathrm{Los}(A', b')$$

Beweis. Dank der Induktion, es genugt, die Aussage für eine einzelne elementare Zeitumformung zu Zeigen. Wir nehmen also an, $(A'|b')$ entsteht aus $(A|b)$ durch eine elementare Zeilenumformung vom Typ:

(GZ_1) Die zwei Systeme $\{Ax = b\}$ und $\{A'x = b'\}$ haben die selbe Gleichungen, nur in einer anderen Reihenfolge.

(GZ_2) Nehmen wir an, dass die h-te Zeile von $(A'|b')$ die Summe des h-te Zeile von $(A|b)$ und des λ-fachen der k-te Zeile von $(A|b)$ ist. Alle Zeilen außer h und k bleiben unverändert. Es genugt, daher, zu zeigen, daß die Lösungsmengen von

$$\begin{cases} a_{h1}x_1 + \cdots + a_{hn}x_n & = b_h \\ a_{k1}x_1 + \cdots + a_{kn}x_n & = b_k \end{cases}, \quad \text{und} \quad \begin{cases} (a_{h1} + \lambda a_{k1})x_1 + \cdots + (a_{hn} + \lambda a_{kn})x_n & = b_h + \lambda b_k \\ a_{k1}x_1 + \cdots + a_{kn}x_n & = b_k \end{cases},$$

gleich sind. Addiert man ein λ-Vielfaches der zweiten Gleichung links zur ersten Gleichung links, so erhält man das System rechts. Somit ist jede Lösung des linken Systems auch eine Lösung des rechten Systems. Umgekehrt erhält man das linke System, wenn man ein λ-Vielfaches der zweiten Gleichung auf der rechten Seite von der ersten Gleichung auf der rechten Seite subtrahiert. Somit ist jede Lösung des rechten Systems auch eine Lösung des linken Systems.

(GZ_3) Nehmen wir an, dass die h-te Zeile von $(A'|b')$ das λ-Vielfaches der h-the Zeile von $(A|b)$ ist, mit $\lambda \in \mathbb{K}, \lambda \neq 0$. Dann, analog zu ($GZ_2$), müssen wir zeigen dass die zwei Gleichungen

$$a_{h1}x_1 + \cdots + a_{hn}x_n = b_h, \quad \text{und} \quad \lambda a_{h1}x_1 + \cdots + \lambda a_{hn}x_n = \lambda b_h$$

die gleiche Lösungen haben. Multipliziert man jedoch die erste Gleichung mit λ, erhält man die erste Gleichung, und multipliziert man die zweite Gleichung mit λ^{-1} (existiert weil $\lambda \neq 0$), erhält man die erste Gleichung. Dann haben die beiden Gleichungen die gleichen Lösungen.

□

Das folgende Satz ist grundlegend für die lineare Algebra

Satz 2.3.3 (Gaußches Eliminationsverfahren)**.** *Jede Matrix A läßt sich durch endliche viele elementare Zeilenumformungen vom Typ (GZ_1) und (GZ_2) auf Zeilenstufenform bringen, mit (GZ_1),(GZ_2) und (GZ_3) sogar auf reduzierte Zeilenstufenform.*

Beweis. Der Beweis ist konstruktiv, durch das Gaußsches Eliminationsverfahren. Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix. Ist $A = 0$, so ist A schon auf reduzierte Zeilenstufenform. Für den allgemeinen Fall beweisen wir zunächst, dass die Matrix in die Zeilenstufenform gebracht werden kann. Wenn $m = 1$ so das A nur eine Zeile hat, so ist A schon in Zeilenstufenform. Wenn $m > 1$, führen wir die Schritte des **Gaußsches Eliminationsverfahren** durch:

- **Schritt 1.**: Die Spalten von links nach rechts durchgehen, bis die erste Spalte A^{j_1} , die nicht Null ist, gefunden wird.
- **Schritt 2.**: Zwei Zeilen vertauschen (Zeilenumformung (GZ_1)), so dass das Element am Anfang der Spalte A^{j_1} ungleich Null ist. Sei A' die neue Matrix. Dann $a'_{1j_1} \neq 0$.

$$A \longrightarrow A' = \begin{pmatrix} 0 & 0 & \dots & 0 & a'_{1j_1} & * & * & \dots & * \\ 0 & 0 & \dots & 0 & a'_{2j_1} & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a'_{mj_1} & * & * & \dots & * \end{pmatrix}$$

- **Schritt 3.**: Für alle $i > 1$, subtrahiere von der i-ten Zeile von A' das $\frac{a_{ij_1}}{a_{1j_1}}$-Vielfache der ersten Zeile (GZ_2), so dass alle Einträge unter a_{1,j_1} zu Null werden. Sei A'' die neue Matrix:

$$A' \longrightarrow A'' = \begin{pmatrix} 0 & 0 & \dots & 0 & a'_{1j_1} & * & * & \dots & * \\ 0 & 0 & \dots & 0 & 0 & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & * & * & \dots & * \end{pmatrix}$$

Sei A_1 die $(m-1) \times (n - j_1 + 1)$ Matrix unter rechts von a_{1j_1} (siehe Tafel). Wenn A_1 in Zeilenstufenform ist, dann ist die ganze Matrix A'' in Zeilenstufenform auch. Wenn A_1 nicht in Zeilenstufenform ist, das Gauß-verfahren wiederholen. Da die erste Matrix m Zeilen hat, müssen wir den Algorithmus höchstens $m-1$ mal wiederholen, bis wir fertig sind (eine Matrix mit nur einer Zeile ist bereits in Zeilenstufenform).

Sobald die Matrix in Zeilstufenform ist, können wir die Zeile des Pivots a_{ij_i} mit $a_{ij_1}^{-1}$ multiplizieren (dies ist möglich, weil $a_{ij_i} \neq 0$), so dass die Pivots eins werden. Danach können wir Schritt 3 wiederholen, aber diesmal mit den Zeilen oberhalb der Pivots. Dadurch wird die Matrix von einer Zeilenstufenform zu einer reduzierten Zeilenstufenform □

Beispiel 2.3.4. Wir wenden den Algorithmus auf die folgende Matrix an:

$$\begin{pmatrix} 0 & 0 & 2 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 2 & 6 & 7 \end{pmatrix} \xrightarrow{Z1 \leftrightarrow Z2} \begin{pmatrix} 0 & 1 & 4 & 5 \\ 0 & 0 & 2 & 3 \\ 0 & 2 & 6 & 7 \end{pmatrix} \xrightarrow{Z3 \to Z3 - 2 \cdot Z1} \begin{pmatrix} 0 & 1 & 4 & 5 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & -2 & -3 \end{pmatrix} \xrightarrow{Z3 \to Z3 + Z2} \begin{pmatrix} 0 & 1 & 4 & 5 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Dies bringt die Matrix in Zeilenstufenform. Um es in reduzierte Zeilstufenform zu bringen gehen wir weiter:

$$\begin{pmatrix} 0 & 1 & 4 & 5 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{Z2 \to \frac{1}{2} \cdot Z2} \begin{pmatrix} 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{Z1 \to Z1 - 4 \cdot Z2} \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dank dieses Algorithmus haben wir eine allgemeine Methode zur Lösung linearer Systeme: Wir bringen die erweiterte Koeffizientenmatrix in eine reduzierte Zeilenstufenform und verwenden dann die Proposition 2.2.5. Wenn wir nur wissen wollen, ob ein System Lösungen hat oder nicht, reicht es eigentlich aus, eine Zeilenstufenform zu berechnen:

Lemma 2.3.5. *Ein LGS mit erweiterte Koeffizientenmatrix* $(A|b)$ *in Zeilenstufenform hat eine Losung genau dann, wenn* $\mathrm{rk}(A) = \mathrm{rk}(A|b)$.

Beweis. Der gleiche Beweis (Aufgabe) wie in Satz 2.2.5 zeigt, dass wenn $(A|b)$ Zeilenstufenform hat, dann hat auch A Zeilenstufenform, so dass es sinnvoll ist, von ihren Rängen zu sprechen[1]. Wir können den Algorithmus von Gauß verwenden, um die Matrix $(A|b)$ in eine andere Matrix (A'—b') in reduzierter Zeilenstufenform zu bringen. Betrachtet man den Algorithmus, so sieht man, dass die Pivots von $(A|b)$ die gleichen sind wie die Pivots von $(A'|b')$, also $\mathrm{rk}(A) = \mathrm{rk}(A')$ und $\mathrm{rk}(A|b) = \mathrm{rk}(A'|b')$. Proposition 2.3.2 zeigt, dass $\mathrm{Los}(A, b) \neq \emptyset$ genau dann, wenn $\mathrm{Los}(A', b') \neq \emptyset$, und Satz 2.2.5 zeigt, dass $\mathrm{Los}(A', b') \neq \emptyset$ genau dann, wenn $\mathrm{rk}(A') = \mathrm{rk}(A'|b')$. Da $\mathrm{rk}(A') = \mathrm{rk}(A)$ und $\mathrm{rk}(A|b) = \mathrm{rk}(A'|b')$, ist dies äquivalent zu $\mathrm{rk}(A|b) = \mathrm{rk}(A'|b')$. □

[1]In der Zukunft werden wir sehen, wie man den Rang einer beliebigen Matrix definiert, aber im Moment haben wir ihn nur für die Matrizen in Zeilenstufenform definiert

Bemerkung 2.3.6 (Lösung eines LGS)**.** Wir fassen zusammen, wie man ein lineares System mit dem Gauß-Algorithmus löst:

- Schritt 1.: Die erweiterte Koeffizientenmatrix $(A|b)$ in Zeilenstufenform bringen.
- Schritt 2.: Wenn $\mathrm{rk}(A) \neq \mathrm{rk}(A|b)$, dann $\mathrm{Los}(A|b) = \emptyset$.
- Schritt 3.: Wenn $\mathrm{rk}(A) = \mathrm{rk}(A|b)$, dann die erweiterte Koeffizientenmatrix $(A|b)$ in reduzierte Zeilenstufenform bringen, und die Lösungen durch Proposition 2.2.5 bestimmen.

Beispiel 2.3.7. Wir anwenden das Algorithmus fürs LGS

$$\begin{cases} x + 2y - 2z &= -15 \\ 2x + y - 5z &= -21 \\ x - 4y + z &= 18 \end{cases}$$

Wir bringen die erweiterte Koeffizientenmatrix in Zeilenstufenform:

$$(A|b) = \begin{pmatrix} 1 & 2 & -2 & | & -15 \\ 2 & 1 & -5 & | & -21 \\ 1 & -4 & 1 & | & 18 \end{pmatrix} \xrightarrow[Z3 \to Z3 - Z1]{Z2 \to Z2 - 2\cdot Z1} \begin{pmatrix} 1 & 2 & -2 & | & -15 \\ 0 & -3 & -1 & | & 9 \\ 0 & -6 & 3 & | & 33 \end{pmatrix}$$
$$\xrightarrow{Z3 \to Z3 - 2\cdot Z2} \begin{pmatrix} 1 & 2 & -2 & | & -15 \\ 0 & -3 & -1 & | & 9 \\ 0 & 0 & 5 & | & 15 \end{pmatrix}$$

Die erweiterte Koeffizientenmatrix ist jetzt in Zeilenstufenform und hat Rang 3. Die Koeffizientenmatrix hat auch Rang 3, sodass Lösungen existieren. Um alle Lösungen zu bestimmen, bringen wir die erweiterte Koeffizientenmatrix in reduzierte Zeilenstufenform:

$$\begin{pmatrix} 1 & 2 & -2 & | & -15 \\ 0 & -3 & -1 & | & 9 \\ 0 & 0 & 5 & | & 15 \end{pmatrix} \xrightarrow{Z3 \to \frac{1}{5} Z3} \begin{pmatrix} 1 & 2 & -2 & | & -15 \\ 0 & -3 & -1 & | & 9 \\ 0 & 0 & 1 & | & 3 \end{pmatrix} \xrightarrow[Z1 \to Z1 + 2\cdot Z3]{Z2 \to Z2 + Z3} \begin{pmatrix} 1 & 2 & 0 & | & -9 \\ 0 & -3 & 0 & | & 12 \\ 0 & 0 & 1 & | & 3 \end{pmatrix}$$
$$\xrightarrow{Z2 \to -\frac{1}{3} Z2} \begin{pmatrix} 1 & 2 & 0 & | & -9 \\ 0 & 1 & 0 & | & -4 \\ 0 & 0 & 1 & | & 3 \end{pmatrix} \xrightarrow{Z1 \to Z1 - 2\cdot Z2} \begin{pmatrix} 1 & 0 & 0 & | & -1 \\ 0 & 1 & 0 & | & -4 \\ 0 & 0 & 1 & | & 3 \end{pmatrix}$$

Das LGS hat jetzt reduzierte Zeilenstufenform

$$\begin{cases} x_1 &= -1 \\ x_2 &= -4 \\ x_3 &= 3. \end{cases}$$

sodass das System eine einzige Lösung hat:

$$\mathrm{Los} = \left\{ \begin{pmatrix} -1 \\ -4 \\ 3 \end{pmatrix} \right\}$$

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 30.04.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Beispiel 2.3.8. Für welche $b \in \mathbb{R}^3$ hat das LGS

$$\begin{cases} x + 2y + 3z &= b_1 \\ 4x + 5y + 6z &= b_2 \\ -2x - y &= b_3 \end{cases}$$

eine Lösung? Wir können diese Frage mit Hilfe des Gauß-Algorithmus beantworten: Wir bringen zuerst die erweiterte Koeffizientenmatrix in Zeilenstufenform

$$\begin{pmatrix} 1 & 2 & 3 & | & b_1 \\ 4 & 5 & 6 & | & b_2 \\ -2 & -1 & 0 & | & b_3 \end{pmatrix} \xrightarrow[Z3\mapsto Z3+2\cdot Z1]{Z2\mapsto Z2-4\cdot Z1} \begin{pmatrix} 1 & 2 & 3 & | & b_1 \\ 0 & -3 & -6 & | & b_2 - 4b_1 \\ 0 & 3 & 6 & | & b_3 + 2b_1 \end{pmatrix}$$
$$\xrightarrow{Z3\to Z3+Z2} \begin{pmatrix} 1 & 2 & 3 & | & b_1 \\ 0 & -3 & -6 & | & b_2 - 4b_1 \\ 0 & 0 & 0 & | & b_3 + b_2 - 2b_1 \end{pmatrix}$$

Diese Matrix hat Zeilenstufenform, und wir sehen dass die Koeffizientenmatrix hat Rang 2. Die erweiterte Koeffizientenmatrix hat Rang 2 genau dann, wenn $b_3 + b_2 - 2b_1 = 0$. Hence, das LGS hat eine Lösung genau dann, wenn $b_3 + b_2 - 2b_1 = 0$.

In diesem Fall können wir auch alle Lösungen beschreiben: wir bringen die erweiterte Koeffizientenmatrix in reduzierte Zeilenstufenform (die dritte Zeile ist Null und wir koennen sie ignorieren)

$$\begin{pmatrix} 1 & 2 & 3 & | & b_1 \\ 0 & -3 & -6 & | & -4b_1 + b_2 \end{pmatrix} \xrightarrow{Z2\mapsto -\frac{1}{3}Z2} \begin{pmatrix} 1 & 2 & 3 & | & b_1 \\ 0 & 1 & 2 & | & \frac{4}{3}b_1 - \frac{1}{3}b_2 \end{pmatrix}$$
$$\xrightarrow{Z1\mapsto Z1-2\cdot Z2} \begin{pmatrix} 1 & 0 & -1 & | & -\frac{5}{3}b_1 + \frac{2}{3}b_2 \\ 0 & 1 & 2 & | & \frac{4}{3}b_1 - \frac{1}{3}b_2 \end{pmatrix}$$

Das System hat jetzt die Form

$$\begin{cases} x &= z - \frac{5}{3}b_1 + \frac{2}{3}b_2 \\ y &= -2z + \frac{4}{3}b_1 - \frac{1}{3}b_2 \end{cases}$$

und die Lösungsmenge ist

$$\text{Los} = \left\{ \begin{pmatrix} z - \frac{5}{3}b_1 + \frac{2}{3}b_2 \\ -2z + \frac{4}{3}b_1 - \frac{1}{3}b_2 \\ z \end{pmatrix} \mid z \in \mathbb{R} \right\}$$

Bemerkung 2.3.9. Wir haben gesehen, dass wir jedes lineare Gleichungssystem über $\mathbb{R}$ mit dem Gauß-Algorithmus lösen können, insbesondere mithilfe der drei Zeilenoperationen. Gibt es etwas Besonderes an den reellen Zahlen? Wenn wir uns die Zeilenoperationen ansehen, stellen wir fest, dass was wir brauchen ist das wir zwei reelle Zahlen a, b addieren und multiplizieren können, um andere reelle Zahlen $a + b$ und $a \cdot b$ zu erhalten. Außerdem, brauchen wir dass jede von null verschiedene reelle Zahl $a \neq 0$ eine multiplikative Inverse $\frac{1}{a}$ besitzt, die wiederum eine reelle Zahl ist. Diese Eigenschaften gelten auch für die rationalen Zahlen $\mathbb{Q}$, aber zum Beispiel gilt die letzte nicht für die ganzen Zahlen $\mathbb{Z}$. Die allgemeine Struktur, die diese Eigenschaften ermöglicht, ist die eines Körpers, den wir im nächsten Kapitel kennenlernen werden.

Kapitel 3

Vektorräume

3.1 Abelsche Gruppe, Ringe, Körpern

Definition 3.1.1 (Abelsche Gruppe)**.** Eine abelsche Gruppe ist eine Menge A zusammen mit einer Verknüpfung

$$+\colon A \times A \to A, \qquad (a, b) \mapsto a + b$$

mit den folgenden Eigenschaften:

- **Assoziativität:** $a + (b + c) = (a + b) + c$ für alle $a, b, c \in A$.
- **Neutrales Element:** Es gibt ein Element $0 \in A$ so, dass $a + 0 = 0 + a = a$ für alle $a \in A$ gilt.
- **Inverses Element:** Für jedes $a \in A$ gibt es ein Element $a' \in A$ so, dass $a + a' = a' + a = 0$ gilt. Man schreibt $a' = -a$.
- **Kommutativität:** $a + b = b + a$ für alle $a, b \in A$.

Wir bezeichnen eine solche Gruppe mit $(A, +)$ oder manchmal nur mit A, wenn die Verknüpfung $+$ klar ist.

Bemerkung 3.1.2. Eine Menge A mit einer Verknüpfung, die die ersten drei Eigenschaften erfüllt, nennt man eine Gruppe. Eine abelsche Gruppe ist eine Gruppe, deren Verknüpfung kommutativ ist.

Bemerkung 3.1.3. Wenn $(A, +)$ eine abelsche Gruppe ist, sagt uns die Assoziativität, dass $(a + b) + c = a + (b + c)$ für alle $a, b, c \in A$. Insbesondere können wir dieses Element einfach als $a + b + c$ bezeichnen. Auf die gleiche Weise können wir $a_1 + a_2 + \cdots + a_n$ für jede endliche Sammlung von Elementen $a_i \in A, i = 1 \ldots, n$ definieren. Als weitere Notation schreiben wir $a - b$ für $a + (-b)$.

In der Definition der Gruppe fragen wir nach der Existenz eines neutralen Elements und eines inversen Elements. Tatsächlich sind diese Elemente, falls sie existieren, eindeutig bestimmt:

Lemma 3.1.4. *Sei $(A, +)$ eine abelsche Gruppe.*

1. *Das neutrale Element ist eindeutig bestimmt: Wenn $0, 0' \in A$ zwei neutrale Elemente sind, dann gilt $0 = 0'$.*

2. *Das inverse Element ist eindeutig bestimmt: Wenn a', a'' zwei inverse Elemente von $a \in A$ sind, dann gilt $a' = a''$.*

Beweis. 1. Nach der Definition eines neutralen Elements: $0 = 0 + 0' = 0'$.

2. Wir nutzen die Eigenschaften einer Gruppe:

$$\begin{aligned} a' &= a' + 0 && \text{(Neutrales Element)} \\ &= a' + (a + a'') && \text{(Inverses Element)} \\ &= (a' + a) + a'' && \text{(Assoziativität)} \\ &= 0 + a'' && \text{(Inverses Element)} \\ &= 0 + a'' = a''. && \text{(Neutrales Element)} \end{aligned}$$

Hier haben wir explizit erwähnt, wo wir die Eigenschaften der Gruppenoperation verwendet haben. In Zukunft werden wir nicht so explizit sein und die Eigenschaften nicht mehr explizit erwähnen.

□

Beispiel 3.1.5. Die Mengen der ganzen Zahlen $\mathbb{Z}$, rationalen Zahlen $\mathbb{Q}$ und reellen Zahlen $\mathbb{R}$ mit der üblichen Addition bilden abelsche Gruppen $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +)$. Die Menge der natürlichen Zahlen $\mathbb{N}$ mit der üblichen Addition bildet keine Gruppe, da es kein additives Inverses gibt: z.B. $-1 \notin \mathbb{N}$.

Beispiel 3.1.6. Die Mengen der nicht nullen reellen Zahlen $\mathbb{R}^\times = \mathbb{R} \setminus 0$ mit der Multiplikation $(\mathbb{R}^\times, \cdot)$ und die Menge der nicht nullen rationalen Zahlen $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$ mit der Multiplikation $(\mathbb{Q}^\times, \cdot)$ bilden ebenfalls abelsche Gruppen. Die Menge der nicht nullen ganzen Zahlen mit der Multiplikation ist jedoch keine Gruppe: Es fehlt das inverse Element, z.B. $\frac{1}{2} \notin \mathbb{Z}$.

Bemerkung 3.1.7. Wie das vorherige Beispiel zeigt, ist das Symbol $+$ für die Operation in einer abelschen Gruppe nur eine Konvention. Manchmal kann die Operation durch andere Symbole angegeben werden, wie z.B. in $(\mathbb{Q}^\times, \cdot)$. Wenn wir von einer allgemeinen abelschen Gruppe sprechen, werden wir weiterhin die Notation $+$ verwenden, aber wenn wir mit expliziten Beispielen umgehen, ist es wichtig zu beachten, was die Operation ist.

Wir geben einige grundlegende Eigenschaften abelscher Gruppen an, die zeigen, dass sie sich in vielen Aspekten wie die Gruppe der ganzen Zahlen $\mathbb{Z}$ verhalten.

Lemma 3.1.8. *Sei $(A, +)$ eine abelsche Gruppe.*

1. *Für alle $a, b \in A$ gilt $-(a + b) = -a - b$ und $-(-a) = a$.*

2. *Für alle $a, b, c \in A$ gilt $a + b = c$ genau dann, wenn $a = c - b$. Insbesondere gilt $a + c = b + c$ genau dann, wenn $a = b$.*

3. *Für jedes $n \in \mathbb{N}$ und $a \in A$ definieren wir*

$$n \cdot a = a + a + \cdots + a \quad (n \text{ mal }), \quad (-n) \cdot a = (-a) + (-a) + \cdots + (-a) \quad (n \text{ mal })$$

Dann gilt

$$(-1) \cdot a = (-a), \quad (n + m) \cdot a = n \cdot a + m \cdot a, \quad n \cdot (m \cdot a) = (n \cdot m) \cdot a$$

Beweis. 1. Um zu zeigen, dass $-(a+b) = -a-b$ gilt, müssen wir beweisen, dass $(a+b)+(-a-b) = 0$, aber $a+b-a-b = a-a+b-b = 0+0 = 0$. Um zu zeigen, dass $-(-a) = a$ gilt, müssen wir beweisen, dass $-a+a = 0$, was klar ist.

2. Wenn $a+b = c$, dann $a+b-b = c-b$, sodass $a = a+0 = c-b$. Umgekehrt, wenn $a = c-b$, dann $a+b = c-b+b = c$.

3. Übung.

□

Definition 3.1.9 (Ring)**.** Ein kommutativer Ring mit Eins ist eine Menge A zusammen mit zwei Operationen, Addition und Multiplikation

$$+\colon A \times A \to A, \qquad \cdot\colon A \times A \to A$$

so dass

- $(A, +)$ eine abelsche Gruppe ist.
- **Assoziativität der Multiplikation**: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in A$.
- **Multiplikative Eins**: Es existiert ein Element $1 \in A$ so dass $a \cdot 1 = 1 \cdot a = a$ für alle $a \in A$.
- **Kommutativität der Multiplikation**: $a \cdot b = b \cdot a$ für alle $a, b \in A$.
- **Distributivität**: $a \cdot (b+c) = a \cdot b + a \cdot c$ und $(a+b) \cdot c = a \cdot c + b \cdot c$ für alle $a, b, c \in A$.

Definition 3.1.10. Wenn die Multiplikation nur assoziativ und distributiv bezüglich der Addition ist, erhalten wir einen Ring. Ein Ring, in dem die Multiplikation kommutativ ist, wird als kommutativer Ring bezeichnet, ein Ring mit einer multiplikativen Eins wird als Ring mit Eins bezeichnet.

Definition 3.1.11 (Körper)**.** Ein Körper ist ein kommutativer Ring mit Eins $(\mathbb{K}, +, \cdot)$, in dem $1 \neq 0$ ist und jedes nicht null Element ein multiplikatives Inverses hat:

- **Multiplikatives Inverses**: Für jedes $a \in \mathbb{K}, a \neq 0$ gibt es ein Element $a^{-1} \in \mathbb{K}$ so dass $aa^{-1} = 1$.

Wir sagen auch, dass jedes nicht null Element bezüglich der Multiplikation invertierbar ist. Wir schreiben auch $\frac{1}{a} = a^{-1}$ und $\frac{b}{a} = ba^{-1}$.

Beispiel 3.1.12. Die Menge $\mathbb{Z}$ mit der üblichen Addition und Multiplikation ist ein kommutativer Ring mit Eins, aber kein Körper, da wir z.B. kein multiplikatives Inverses für 2 haben. Die Menge der rationalen Zahlen $\mathbb{Q}$ und die Menge der reellen Zahlen $\mathbb{R}$ sind Körper.

Lemma 3.1.13. *Sei $(A, +, \cdot)$ ein Ring.*

1. *Wenn die multiplikative Eins existiert, ist sie eindeutig.*
2. *Wenn das multiplikative Inverse eines Elements $a \in A$ existiert, ist es eindeutig.*

Außerdem gilt für alle $a, b, c, d, e \in A$:

3. *$a \cdot 0 = 0 \cdot a = 0$.*

4. *$(-a) \cdot b = a \cdot (-b) = -(ab)$.*

5. *$(-a) \cdot (-b) = ab$.*

6. *$a \cdot (b - c) = ab - ac$ und $(a - b)c = ac - bc$.*

7. *$a \cdot (n \cdot b) = (n \cdot a) \cdot b = n \cdot (ab)$ für alle $n \in \mathbb{Z}$.*

Außerdem, wenn A ein Körper ist, gilt:

8. *$(a^{-1})^{-1} = a$ für jedes $a \neq 0$.*

9. *$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$, für jedes $b, d \neq 0$.*

10. *$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$, für jedes $b, d \neq 0$.*

11. *$ab = 0$ genau dann, wenn $a = 0$ oder $b = 0$.*

12. *$ab = ac$ genau dann, wenn $a = 0$ oder $b = c$.*

Beweis.

1. Seien $1, 1'$ zwei multiplikative Einheiten. Dann $1 = 1 \cdot 1' = 1'$.

2. Seien a', a'' zwei multiplikative Inversen von a. Dann $a' = a' \cdot 1 = a' \cdot (a \cdot a'') = (a' \cdot a) \cdot a'' = 1 \cdot a'' = a''$.

3. Wir sehen, dass $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, also $a \cdot 0 = 0$. Etwas Ähnliches gilt für $0 \cdot a$.

4. Um zu zeigen, dass $(-a) \cdot b = -(ab)$ gilt, müssen wir zeigen, dass $(-a) \cdot b + ab = 0$, und wir können dies wie folgt tun: $(-a) \cdot b + ab = (-a + a) \cdot b = 0 \cdot a = 0$. Eine ähnliche Argumentation beweist die andere Gleichheit.

5. Aus dem zuvor Bewiesenen folgt: $(-a) \cdot (-b) = -(a \cdot (-b)) = -(-a \cdot b) = a \cdot b$.

6. Aus dem zuvor Bewiesenen folgt: $a \cdot (b - c) = a \cdot b + a \cdot (-c) = a \cdot b - a \cdot c$.

7. Übung.

8. Da $aa^{-1} = 1$, muss $a = (a^{-1})^{-1}$ sein.

9. Wir sehen, dass $\frac{a}{b} \cdot \frac{c}{d} = ab^{-1}cd^{-1} = acb^{-1}d^{-1}$, also reicht es zu beweisen, dass $b^{-1}c^{-1} = (bc)^{-1}$. Dies ist einfach, weil $b^{-1}c^{-1}bc = b^{-1}bc^{-1}c = 1 \cdot 1 = 1$.

10. Wir sehen, dass $\frac{ad+bc}{bd} = (ad + bc) \cdot (bd)^{-1} = (ad) \cdot (bd)^{-1} + (bc) \cdot (bd)^{-1} = adb^{-1}d^{-1} + bcb^{-1}d^{-1} = ab^{-1} + cd^{-1} = \frac{a}{b} + \frac{c}{d}$.

11. Wenn $ab = 0$ und $b \neq 0$, dann $0 = 0 \cdot b^{-1} = ab \cdot b^{-1} = a$.

12. Wir sehen, dass $ab = ac$ genau dann, wenn $a(b - c) = 0$. Dann folgt die Schlussfolgerung aus dem vorherigen Punkt.

$\square$

Bemerkung 3.1.14. Die vorherigen Eigenschaften zeigen, dass wenn $\mathbb{K}$ ein Körper ist, dann die Menge der nicht null Elemente $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$ mit der Multiplikation

$$\mathbb{K}^\times \times \mathbb{K}^\times \to \mathbb{K}^\times, \qquad (a, b) \mapsto ab$$

eine abelsche Gruppe ist. Sie können diese Aussage als Übung beweisen..

3.1.1 Vektoren, Matrizen und lineare Gleichungssysteme über einem Körper

Die zuvor gegebenen Definitionen von Vektoren und Matrizen können über einem beliebigen Körper $\mathbb{K}$ anstelle von nur über $\mathbb{R}$ gegeben werden. Daher können wir von Vektoren $b \in \mathbb{K}^m$ und von Matrizen $A \in \mathbb{K}^{m\times n}$ sprechen. Wir können auch von dem entsprechenden linearen Gleichungssystem und Lösungsraum sprechen

$$A \cdot x = b, \qquad \mathrm{Los}(A, b) = \{x \in \mathbb{K}^n \mid A \cdot x = b\}$$

Die Algorithmen, die wir zuvor über $\mathbb{R}$ vorgestellt haben, übertragen sich ohne Änderungen auf jeden Körper $\mathbb{K}$: Tatsächlich basierten diese Algorithmen auf den drei elementaren Zeilenumformungen auf einer Matrix, die über einem beliebigen Körper $\mathbb{K}$ durchgeführt werden können. Insbesondere kann der Gauß-Eliminationsverfahren über einem beliebigen Körper verwendet werden, um eine Matrix in (reduzierte) Zeilenstufenform zu bringen, und dies kann verwendet werden, um die Menge der Lösungen $\mathrm{Los}(A, b)$ über einem beliebigen Körper $\mathbb{K}$ zu bestimmen.

3.1.2 Endliche Körper

Die bisher gesehenen Beispiele von Körpern sind unendliche Körper: $\mathbb{Q}, \mathbb{R}$. Nicht alle Körper sind unendlich:

Beispiel 3.1.15. Betrachten Sie eine Menge $\mathbb{F}_2 = \{[0], [1]\}$ mit zwei verschiedenen Elementen $[0], [1]$. Wir definieren eine Summe und eine Multiplikation auf $\mathbb{F}_2$ wie folgt

$$\begin{aligned} [0] + [0] &= [0], & \qquad [0] \cdot [0] &= [0], \\ [0] + [1] &= [1], & [0] \cdot [1] &= [0], \\ [1] + [0] &= [1], & [1] \cdot [0] &= [0], \\ [1] + [1] &= [0], & [1] \cdot [1] &= [1]. \end{aligned}$$

Dann ist $\mathbb{F}_2$ mit diesen Operationen ein Körper. Um ein Gefühl für diesen Körper zu bekommen, ersetzen Sie [0] durch "gerade" und [1] durch "ungerade".

Allgemein kann bewiesen werden, dass wenn $\mathbb{K}$ ein endlicher Körper ist, dann muss er p^n Elemente haben, wobei p eine Primzahl ist und $n \geq 1$. Umgekehrt gibt es für jede Primzahl p und jedes $n \geq 1$ einen Körper $\mathbb{F}_{p^n}$ mit p^n Elementen.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 3.05.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

3.1.3 Die komplexen Zahlen

Ein wichtiges Beispiel für einen Körper sind die komplexen Zahlen:

Definition 3.1.16 (Komplexe Zahlen)**.** Eine komplexe Zahl ist ein formaler Ausdruck der Form

$$z = a + ib, \qquad a, b \in \mathbb{R}$$

Die reelle Zahl a wird als der Realteil von z bezeichnet und die reelle Zahl b wird als der Imaginärteil von z bezeichnet:

$$a = \Re(z), \qquad b = \Im(z)$$

Zwei komplexe Zahlen sind gleich, genau dann, wenn sie den gleichen Real- und Imaginärteil haben. Die komplexe Zahl $i := 0 + i \cdot 1$ wird als die imaginäre Einheit bezeichnet. Die Menge aller komplexen Zahlen wird durch $\mathbb{C}$ bezeichnet, und $\mathbb{R}$ kann als eine Teilmenge von $\mathbb{C}$ betrachtet werden durch die injektive Abbildung

$$\mathbb{R} \hookrightarrow \mathbb{C}, \qquad a \mapsto a + i \cdot 0$$

so dass $\mathbb{R} = \{z \in \mathbb{C} \mid \Im(z) = 0\}$.

Bemerkung 3.1.17. Der Real- und Imaginärteil bilden Bijektionen

$$\mathbb{R}^2 \to \mathbb{C}, \quad (a, b) \mapsto a + ib, \qquad \mathbb{C} \to \mathbb{R}^2, \quad z \mapsto (\Re(z), \Im(z))$$

so dass $\mathbb{C}$ auch als $\mathbb{R}^2$ betrachtet werden kann. In diesem Sinne nennt man $\mathbb{C}$ die komplexe Ebene. Die Teilmenge $\mathbb{R} = \{z \in \mathbb{C} \mid \Im(z) = 0\}$ wird als die reale Achse bezeichnet und $\{z \in \mathbb{C} \mid \Re(z) = 0\}$ wird als die imaginäre Achse bezeichnet (siehe Tafel).

Definition 3.1.18 (Addition und Multiplikation von komplexen Zahlen)**.** Die Addition und Multiplikation auf $\mathbb{C}$ sind definiert durch

$$\begin{aligned} +\colon \mathbb{C} \times \mathbb{C} \to \mathbb{C} \qquad & (a + ib, c + id) \mapsto (a + c) + i(b + d) \\ \cdot\colon \mathbb{C} \times \mathbb{C} \to \mathbb{C} \qquad & (a + ib, c + id) \mapsto (ac - bd) + i(ad + bc) \end{aligned}$$

Bemerkung 3.1.19. Die Art und Weise, wie man diese Operationen betrachtet, ist, dass man die Ausdrücke $a + ib$ wie erwartet multipliziert, jedoch mit der zusätzlichen Bedingung, dass $i^2 = -1$. Tatsächlich überprüfen wir zunächst, dass $i^2 = -1$ ist:

$$i \cdot i = (0 + i \cdot 1)(0 + i \cdot 1) = -1 + i \cdot 0 = -1$$

Jetzt, wenn wir $(a + ib)$ und $(c + id)$ wie üblich addieren und multiplizieren, erhalten wir

$$\begin{aligned} (a + ib) + (c + id) &= a + ib + c + id = a + c + ib + id = (a + c) + i(b + d). \\ (a + ib) \cdot (c + id) &= ac + iad + ibc + i^2bd = ac + iad + ibc - bd = (ac - bd) + i(ad + bc). \end{aligned}$$

Proposition 3.1.20. *Die Menge der komplexen Zahlen mit diesen beiden Operationen ist ein Körper.*

Beweis. Die Überprüfung der Eigenschaften eines Körpers bleibt als Übung. Wir zeigen nur, wie man das Inverse eines von Null verschiedenen Elements $z = a + ib$ berechnet. Wir beobachten, dass

$$(a + ib) \cdot (a - ib) = a^2 - (ib)^2 = a^2 - (i)^2b^2 = a^2 - (-1)b^2 = a^2 + b^2$$

Insbesondere, wenn $z \neq 0$, dann $a \neq 0$ oder $b \neq 0$, so dass $a^2 + b^2$ reell und positiv ist und wir sein Inverses $(a^2 + b^2)^{-1}$ nehmen können. Dies zeigt, dass

$$(a + ib) \cdot \frac{a - ib}{a^2 + b^2} = \frac{a^2 + b^2}{a^2 + b^2} = 1.$$

□

Definition 3.1.21 (Komplexe Konjugation und Betrag)**.** Die komplexe Konjugierte einer komplexen Zahl $z = a + ib$ ist

$$\bar{z} = a - ib.$$

Der Betrag einer komplexen Zahl $z = a + ib$ ist

$$|z| = \sqrt{a^2 + b^2}$$

Lemma 3.1.22 (Eigenschaften der komplexen Konjugation und des Betrags)**.** *Für alle $z, w \in \mathbb{C}$ gilt:*

1. $\overline{z+w} = \overline{z} + \overline{w}$ *und* $\overline{zw} = \overline{z} \cdot \overline{w}$.
2. $\overline{\overline{z}} = z$.
3. $z + \overline{z} = 2\Re(z)$ *und* $z - \overline{z} = 2\Im(z)$.
4. $z = \bar{z}$ *genau dann, wenn* $z \in \mathbb{R}$.
5. $z \cdot \overline{z} = |z|^2$.
6. $|z| \geq 0$ *und* $|z| = 0$ *genau dann, wenn* $z = 0$
7. $z^{-1} = \frac{\overline{z}}{|z|^2}$ *falls* $z \neq 0$.
8. $|\overline{z}| = |z|$.
9. $|zw| = |z| \cdot |w|$.
10. $\Re(z) \leq |z|, \Im(z) \leq |z|$
11. $|z + w| \leq |z| + |w|$ *(Dreiecksungleichung).*

Beweis. Wir beweisen nur die Dreiecksungleichung und lassen den Rest als Übung. Wir sehen, dass

$$\begin{aligned}|z+w|^2 &= (z+w)\overline{(z+w)} = (z+w)(\overline{z}+\overline{w}) = z \cdot \overline{z} + z \cdot \overline{w} + \overline{z} \cdot w + w \cdot \overline{w} \\ &= |z|^2 + |w|^2 + z \cdot \overline{w} + \overline{z} \cdot w = |z|^2 + |w|^2 + 2\Re(z \cdot \overline{w}) \leq |z|^2 + |w|^2 + |z \cdot \overline{w}| \\ &= |z|^2 + |w|^2 + 2|z| \cdot |w| = (|z| + |w|)^2.\end{aligned}$$

Dies zeigt, dass $|z + w|^2 \leq (|z| + |w|)^2$ ist, und durch Ziehen von Wurzeln sehen wir, dass $|z + w| \leq |z| + |w|$. □

3.1.4 Polynome

Ein wichtiges Beispiel eines Rings ist der Ring der Polynome mit Koeffizienten in einem Körper.

Definition 3.1.23 (Polynom)**.** Sei $\mathbb{K}$ ein Körper. Ein Polynom mit Koeffizienten in $\mathbb{K}$ in der Variablen x ist eine formale Summe

$$f = a_0 + a_1 \cdot x + \cdots + a_n \cdot x^n = \sum_{i=0}^{n} a_i x^n, \qquad a_i \in \mathbb{K}$$

Die a_i werden als Koeffizienten des Polynoms bezeichnet. Der Grad eines nicht nullen Polynoms f ist

$$\deg(f) = \max\{i \mid a_i \neq 0\}$$

Wir setzen auch den Grad des Nullpolynoms auf $\deg(0) = -\infty$. Wenn $\deg(f) = n$, dann wird a_n als der Leitkoeffizient bezeichnet. Ein Polynom mit Leitkoeffizienten 1 wird monisch genannt. Der Koeffizient a_0 wird als der konstante Koeffizient bezeichnet. Ein Polynom heißt konstant, falls $f = a_0$. Die Menge aller Polynome mit Koeffizienten in $\mathbb{K}$ und in der Variablen x wird mit $\mathbb{K}[x]$ bezeichnet.

Bemerkung 3.1.24. Manchmal werden wir ein Polynom wie oben als $f(x)$ bezeichnen, um zu betonen, dass es von der Variablen x abhängt. Es sollte jedoch nicht als Funktion betrachtet werden, es ist nur ein formaler Ausdruck. Wir werden jedoch später sehen, dass es manchmal als Funktion betrachtet werden kann.

Definition 3.1.25 (Summe und Produkt von Polynomen)**.** Für zwei Polynome $f, g \in \mathbb{K}[x]$

$$f = \sum_{i=0}^{n} a_i x^i, \qquad g = \sum_{i=0}^{m} a_i x^i$$

definieren wir ihre Summe und ihr Produkt als

$$f + g \coloneqq \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) x^i, \qquad f \cdot g = \sum_{i=0}^{n+m} \left(\sum_{j=0}^{i} a_j b_{i-j} \right) x^i$$

Hier verwenden wir die Konvention, dass $a_k = 0$ für alle $k > \deg(f)$ und $b_h = 0$ für alle $h > \deg(g)$.

Bemerkung 3.1.26. Die Definition des Produkts mag kompliziert aussehen, aber es ist genau das, was wir erhalten, wenn wir das Produkt

$$(a_0 + a_1 x + \cdots + a_n x^n) \cdot (b_0 + b_1 x + \cdots + b_m x^m)$$

auf die erwartete Weise ausmultiplizieren. Zum Beispiel

$$\begin{aligned}(a_0 + a_1 x) \cdot (b_0 + b_1 x + b_2 x^2) &= a_0 b_0 + a_0 b_1 x + a_0 b_2 x^2 + a_1 b_0 x + a_1 b_1 x^2 + a_1 b_2 x^3 \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) x + (a_0 b_2 + a_1 b_1) x^2 + (a_1 b_2) x^3.\end{aligned}$$

Insbesondere können wir berechnen, dass $x^n \cdot x^m = x^{n+m}$.

Proposition 3.1.27. *Die Menge $\mathbb{K}[x]$ zusammen mit der obigen Addition und Multiplikation ist ein kommutativer Ring mit Eins. Weiterhin gilt für $f, g \in \mathbb{K}[x]$, dass $\deg(f \cdot g) = \deg(f) + \deg(g)$.*

Beweis. Die Tatsache, dass $\mathbb{K}[x]$ ein Ring ist, wird als Übung überlassen. Wir beweisen die Aussage über die Grade. Wenn $f = 0$, dann $0 \cdot g = 0$, so dass $\deg(0 \cdot g) = \deg(0) = -\infty = \deg(0) + \deg(g)$. Das Gleiche gilt, wenn $g = 0$. Wenn $f, g \neq 0$, dann $\deg(f) = n, \deg(g) = m$ und

$$f = \sum_{i=0}^{n} a_i x^i, \qquad g = \sum_{i=0}^{m} b_i x^i$$

mit $a_n, b_m \neq 0$. Dann ist $a_n b_m \neq 0$ (da $\mathbb{K}$ ein Körper ist) und die Formel für das Produkt zeigt, dass $\deg(f \cdot g) = n + m$. $\square$

Sei $f \in \mathbb{K}[x]$ ein Polynom, $f = \sum_{i=0}^{n} a_i x^i$. Wir können eine entsprechende Polynomfunktion definieren:

$$f\colon \mathbb{K} \to \mathbb{K}, \qquad x_0 \mapsto f(x_0) = \sum_{i=0}^{n} a_i x_0^i$$

die wir immer noch mit demselben Zeichen bezeichnen, auch wenn es mehrdeutig ist.

Definition 3.1.28 (Nullstelle eines Polynoms)**.** Sei $f \in \mathbb{K}[x]$ ein Polynom. Ein Element $x_0 \in \mathbb{K}$ ist eine Nullstelle des Polynoms, wenn die entsprechende Polynomfunktion an dem Element Null ist: $f(x_0) = 0$.

Beispiel 3.1.29. Nehmen Sie $f(x) = x^2 - 3x + 2 \in \mathbb{Q}[x]$. Dann ist $f(2) = 2^2 - 3 \cdot 2 + 2 = 4 - 6 + 2 = 0$, so dass 2 eine Nullstelle von f ist.

Proposition 3.1.30 (Abspalten von Nullstellen)**.** *Sei $f(x) \in \mathbb{K}[x]$. Dann gilt $f(x_0) = 0$ genau dann, wenn wir schreiben können*

$$f(x) = (x - x_0) \cdot g(x)$$

für ein Polynom $g(x) \in \mathbb{K}[x]$.

Bevor wir dies beweisen, benötigen wir ein Lemma:

Lemma 3.1.31. *Sei A ein kommutativer Ring mit Eins. Für jedes $a, b \in A$ und $n \in \mathbb{N}, n \geq 2$ gilt*

$$a^n - b^n = (a - b) \cdot (a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$$

Beweis. Wir können explizit berechnen

$$\begin{aligned}(a - b) \cdot (a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}) &= a\left(\sum_{i=0}^{n-1} a^i b^{n-1-i}\right) + b\left(\sum_{i=0}^{n-1} a^i b^{n-1-i}\right)\\ &= \sum_{i=0}^{n-1} a^{i+1} b^{n-1-i} - \sum_{i=0}^{n-1} a^i b^{n-i}\\ &= a^n + \sum_{i=1}^{n-1} a^i b^{n-i} - \sum_{i=1}^{n-1} a_i b^{n-i} - b^n = a^n - b^n.\end{aligned}$$

□

Nun können wir die Proposition beweisen:

Beweis der Proposition 3.1.30. Wenn $f(x) = (x - x_0) \cdot g(x)$, dann $f(x_0) = 0 \cdot g(x_0) = 0$. Umgekehrt, nehmen Sie an, $f(x_0) = 0$. Wenn wir $f(x) = \sum_{i=0}^{n} a_i x^i$ schreiben, erhalten wir $f(x_0) = \sum_{i=0} a_i x_0^i = 0$. Dann

$$f(x) = f(x) - 0 = \sum_{i=0}^{n} a_i x^i - \sum_{i=0}^{n} a_i x_0^i = \sum_{i=0}^{n} a_i (x^i - x_0^i) = \sum_{i=0}^{n} a_i (x - x_0) \cdot g_i(x)$$

wobei $g_i(x) = \sum_{h=0}^{i-1} x_0^h x^{i-1-h}$ aus dem vorherigen Lemma stammt. Dann können wir schreiben

$$f(x) = \sum_{i=0}^{n} a_i (x - x_0) \cdot g_i = (x - x_0) \cdot (\sum_{i=0}^{n} a_i g_i) = (x - x_0) \cdot g(x)$$

□

Korollar 3.1.32. *Sei $f \in \mathbb{K}[x]$ ein Polynom.*

1. *Für $x_1, \ldots, x_m \in \mathbb{K}$ paarweise verschieden gilt $f(x_1) = \cdots = f(x_n) = 0$ genau dann, wenn*
$$f(x) = (x - x_1) \cdot \cdots \cdot (x - x_m) \cdot g(x)$$
für ein bestimmtes $g(x) \in \mathbb{K}[x]$.

2. *Wenn f nicht null und vom Grad n ist, kann es höchstens n Nullstellen haben.*

Beweis. 1. Induktion über m. Wenn $m = 1$, haben wir das bereits zuvor bewiesen. Wenn $m > 1$, dann ist $f(x_m) = 0$, so dass
$$f(x) = (x - x_m) \cdot h(x)$$
Für $i = 1, \ldots, m-1$ muss $0 = f(x_i) = (x_i - x_m) \cdot h(x_i)$ sein. Da $x_i \neq x_m$ sein muss, muss $h(x_i) = 0$ sein. Nach Induktionshypothese $h(x) = (x - x_1) \cdot (x - x_{m-1}) \cdot g(x)$ und $f(x) = (x - x_1) \cdot (x - x_m) \cdot g(x)$.

2. Seien $x_1, \ldots, x_m$ die verschiedenen Nullstellen von f. Dann haben wir bewiesen, dass $f(x) = (x - x_1) \ldots (x - x_m) g(x)$, so dass $n = \deg(f) = \sum_{i=1}^{m} \deg(x - x_i) + \deg(g) = m + \deg(g)$. Dies zeigt $m \leq n$.

□

Korollar 3.1.33. [1] *Sei $\mathbb{K}$ ein unendlicher Körper und seien $f, g \in \mathbb{K}[x]$. Dann gilt $f = g$ in $\mathbb{K}[x]$ genau dann, wenn $f(x_0) = g(x_0)$ für alle $x_0 \in \mathbb{K}$ gilt. Anders ausgedrückt, sind zwei Polynome gleich, wenn ihre Polynomfunktionen gleich sind.*

Beweis. Betrachten Sie das Polynom $h(x) = f(x) - g(x)$. Wir wissen, dass $h(x_0) = 0$ für alle $x_0 \in \mathbb{K}$. Das bedeutet, dass h unendlich viele Nullstellen hat, und das vorherige Lemma zeigt, dass h das Nullpolynom sein muss. □

3.1.5 Algebraisch abgeschlossene Körper

Definition 3.1.34 (Algebraisch abgeschlossener Körper)**.** Ein Körper $\mathbb{K}$ heißt algebraisch abgeschlossen, wenn jedes nichtkonstante Polynom $f \in \mathbb{K}[x]$ eine Nullstelle in $\mathbb{K}$ hat.

Proposition 3.1.35. [2] *Wenn $\mathbb{K}$ algebraisch abgeschlossen ist, dann zerlegt sich jedes nichtkonstante Polynom als Produkt von linearen Faktoren: D.h. für jedes nichtkonstante $f(x) \in \mathbb{K}[x]$ gibt es $x_1, \ldots, x_m \in \mathbb{K}$ zusammen mit Exponenten $e_1, \ldots, e_m \in \mathbb{N}$ und $a \in \mathbb{K}^\times$, sodass*
$$f(x) = a \cdot (x - x_1)^{e_1} \cdot \cdots \cdot (x - x_m)^{e_m}$$

Beweis. Durch Induktion über den Grad von f. Wenn $\deg(f) = 1$, dann ist $f(x) = a_1 x + a_0$ mit $a_1 \neq 0$ und $f(x) = a_1 \cdot (x - (-\frac{a_0}{a_1}))$. Wenn $\deg(f) > 1$ dann hat f eine Nullstelle, da $\mathbb{K}$ algebraisch abgeschlossen ist. Dann können wir $f(x) = (x - x_1) \cdot g(x)$ für ein Polynom $g(x)$ schreiben. Da $\deg(g) = \deg(f) - 1$ zeigt die Induktionshypothese, dass g ein Produkt von linearen Faktoren ist. Dann ist f auch ein Produkt von linearen Faktoren. □

[1] Dies wurde in den Vorlesungen nicht behandelt, wird aber im Repetitorium besprochen
[2] Dies wurde in den Vorlesungen nicht behandelt, wird aber im Repetitorium besprochen

Satz 3.1.36 (Fundamentalsatz der Algebra)**.** *Der Körper der komplexen Zahlen $\mathbb{C}$ ist algebraisch abgeschlossen.*

♣♣♣♣♣♣♣♣♣♣♣♣ **Ende Vorlesung 7.5.2024** ♣♣♣♣♣♣♣♣♣♣♣♣

3.2 Vektorräume und Lineare Abbildungen

Definition 3.2.1 (Vektorraum)**.** Sei $\mathbb{K}$ ein Körper. Ein $\mathbb{K}$-Vektorraum (Vektorraum über $\mathbb{K}$) besteht aus einer Menge V sowie eine Vektoraddition

$$+\colon V \times V \to V, \qquad (v, w) \mapsto v + w$$

und einer Skalarmultiplikation

$$\cdot\colon \mathbb{K} \times V \to V, \qquad (\lambda, v) \mapsto \lambda \cdot v$$

mit den folgenden Eigenschaften:

1. $(V, +)$ ist eine **abelsche Gruppe**.
2. **Assoziativität der Skalarmultiplikation**: $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$ für alle $\lambda, \mu \in \mathbb{K}, v \in V$.
3. 1 **und Skalarmultiplikation** :$1 \cdot v = v$ für alle $v \in V$.
4. **Distributivität**: $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$ und $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ für alle $\lambda, \mu \in \mathbb{K}, v, w \in V$.

Die Elemente von V nennt man Vektoren, die aus $\mathbb{K}$ Skalare.

Beispiel 3.2.2. Ein Körper $\mathbb{K}$ ist selbst ein $\mathbb{K}$-Vektorraum, mit der Multiplikation in $\mathbb{K}$ als Skalarmultiplikation

Beispiel 3.2.3. Für $n \geq 1$ ist die Menge $\mathbb{K}^n$ von Spaltenvektoren ein $\mathbb{K}$-Vektorraum mit der komponentenweise definierten Vektoraddition und Skalarmultiplikation sind

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}, \qquad \lambda \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot a_1 \\ \lambda \cdot a_2 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}$$

Insbesondere ist $\mathbb{R}^n$ ein $\mathbb{R}$-Vektorraum, $\mathbb{Q}^n$ ein $\mathbb{Q}$-Vektorraum und $\mathbb{C}^n$ ein $\mathbb{C}$-Vektorraum. Die Vektorraumstruktur auf $\mathbb{R}, \mathbb{R}^2$ und $\mathbb{R}^3$ kann grafisch dargestellt werden (siehe Tafel)

Beispiel 3.2.4. Die Menge $\{0\}$ ist ein $\mathbb{K}$-Vektorraum mit der Summe $0 + 0 = 0$ und der Skalarmultiplikation $\lambda \cdot 0 = 0$ für alle $\lambda \in \mathbb{K}$. Wir schreiben auch $\mathbb{K}^0 = \{0\}$.

Beispiel 3.2.5. Die Menge $\mathbb{K}^{m\times n}$ von $m \times n$-Matrizen mit Einträge in $\mathbb{K}$ ist ein $\mathbb{K}$- Vektorraum mit der komponentenweise definierten Addition und Skalarmultiplikation:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

$$\lambda \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda\cdot a_{11} & \lambda\cdot a_{12} & \dots & \lambda\cdot a_{1n} \\ \lambda\cdot a_{21} & \lambda\cdot a_{22} & \dots & \lambda\cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda\cdot a_{m1} & \lambda\cdot a_{m2} & \dots & \lambda\cdot a_{mn} \end{pmatrix}$$

Beispiel 3.2.6. Der Polynomring $\mathbb{K}[x]$ ist ein $\mathbb{K}$-Vektorraum mit der Addition von Polynome und mit der Skalarmultiplikation $\lambda \cdot f$ womit $\lambda \in \mathbb{K}$ als konstante Polynome betrachtet ist:

$$\lambda \cdot (a_0 + a_1x + \dots + a_nx^n) = \lambda a_0 + \lambda a_1 x + \dots + \lambda a_n x^n$$

Beispiel 3.2.7. Sei M eine beliebige Menge. Die Menge von alle Funktionen $\mathrm{Fun}(M, \mathbb{K}) = \{f : M \to \mathbb{K}\}$ ist ein $\mathbb{K}$-Vektorraum mit der punktweise definierten Addition und Skalarmultiplikation: für alle $f, g \in \mathrm{Fun}(M, \mathbb{K})$ und

$$\begin{aligned} f + g : M \to \mathbb{K}(f+g)(x) &= f(x) + g(x) \\ \lambda \cdot f : M \to \mathbb{K}(\lambda \cdot f)(x) = \lambda \cdot f(x) \end{aligned}$$

Dasselbe gilt für die Menge $\mathrm{Fun}(M, V)$, womit V ein $\mathbb{K}$-Vektorraum ist.

Lemma 3.2.8. *Sei V ein $\mathbb{K}$-Vektorraum, und seien 0_V das neutrales Element für die Addition in V und $0_{\mathbb{K}}$ das neutrales Element für die Addition in $\mathbb{K}$. Seien $\lambda \in \mathbb{K}$ und $v \in V$, dann*

1. $0_K \cdot v = 0_V$ *und* $\lambda \cdot 0_V = 0_V$.
2. $\lambda \cdot v = 0_V$ *genau dann, wenn* $\lambda = 0_K$ *oder* $v = 0_V$.
3. $(-1) \cdot v = -v$.

Beweis.
1. $0_K \cdot v = (0_K + 0_K) \cdot v = 0_K \cdot v + 0_K \cdot v$ so dass $0_K \cdot v = 0_V$. Außerdem, $\lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V$ so dass $\lambda \cdot 0_V = 0_V$.
2. Wenn $\lambda \neq 0$ dann ist $\lambda \in \mathbb{K}$ invertierbar, da $\mathbb{K}$ ein Körper ist. Dann $v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1}(\lambda \cdot v) = \lambda^{-1} \cdot 0_V = 0_V$.
3. Wir sehen dass $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1-1) \cdot v = 0_K \cdot v = 0_V$, so dass $(-1) \cdot v = -v$.

$\square$

Definition 3.2.9 (Lineare Abbildung)**.** Seien V, W zwei $\mathbb{K}$-Vektorräume. Eine Abbildung $f : V \to W$ heißt ($\mathbb{K}-$)lineare Abbildung falls

$$f(v + v') = f(v) + f(v'), \qquad f(\lambda \cdot v) = \lambda \cdot f(v) \qquad \text{für alle } v, v' \in V, \lambda \in \mathbb{K}$$

Eine lineare Abbildung heißt auch ein Homomorphismus von $\mathbb{K}$-Vektorräume. Die Menge von alle lineare Abbildungen von V nach W nennen wir $\mathrm{Hom}_K(V, W)$.

Beispiel 3.2.10. Sei V ein $\mathbb{K}$-Vektorraum. Die Identität $\mathrm{id}_V : V \to V$ ist eine lineare Abbildung. Die Null-Abbildung

$$0 : V \to V, \qquad v \mapsto 0$$

ist eine lineare Abbildung. Im Allgemeinen, für jede $\lambda \in \mathbb{K}$ ist die Abbildung

$$m_\lambda \cdot V \to V, \qquad v \mapsto \lambda \cdot v$$

eine lineare Abbildung. Wir sehen dass $m_1 = \mathrm{id}_V$ und $m_0 = 0$.

Beispiel 3.2.11. Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix mit Einträge in A. Wir definieren eine Abbildung

$$L_A : \mathbb{K}^n \to \mathbb{K}^m, \qquad x \mapsto A \cdot x$$

Genauer gesagt, wenn $A = (a_{ij})$, dann

$$L_A : \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

Das ist eine lineare Abbildung: für alle $x, y \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$ gilt dass

$$\begin{aligned}
L_A(x+y) = A \cdot (x+y) &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \left(\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \right) \\
&= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \dots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} a_{11}(x_1+y_1) + a_{12}(x_2+y_2) + \dots + a_{1n}(x_n+y_n) \\ a_{21}(x_1+y_1) + a_{22}(x_2+y_2) + \dots + a_{2n}(x_n+y_n) \\ \dots \\ a_{m1}(x_1+y_1) + a_{m2}(x_2+y_2) + \dots + a_{mn}(x_n+y_n) \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \\ \dots \\ a_{m1}y_1 + a_{m2}y_2 + \dots + a_{mn}y_n \end{pmatrix} \\
&= A \cdot x + A \cdot y = L_A(x) + L_A(y).
\end{aligned}$$

$$\begin{aligned}
L_A(\lambda \cdot x) = A \cdot (\lambda \cdot x) &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \left(\lambda \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \right) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \dots \\ \lambda x_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11}(\lambda x_1) + a_{12}(\lambda x_2) + \dots + a_{1n}(\lambda x_n) \\ a_{21}(\lambda x_1) + a_{22}(\lambda x_2) + \dots + a_{2n}(\lambda x_n) \\ \dots \\ a_{m1}(\lambda x_1) + a_{m2}(\lambda x_2) + \dots + a_{mn}(\lambda x_n) \end{pmatrix} = \lambda \cdot \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} \\
&= \lambda \cdot (Ax) = \lambda \cdot L_A(x)
\end{aligned}$$

Wir schreiben dies noch einmal, weil es wichtig ist: Wir haben bewiesen, dass die Abbildung L_A linear ist. Das bedeutet dass

$$A \cdot (x + y) = A \cdot x + A \cdot y, \qquad A \cdot (\lambda \cdot x) = \lambda \cdot (Ax)$$

für alle $x, y \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$.

Beispiel 3.2.12. Die Identitätsmatrix $I_n \in \mathbb{K}^{n \times n}$ ist die Matrix

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Man sieht dass die lineare Abbildung $L_{I_n} : \mathbb{K}^n \to \mathbb{K}^n$ die Identität $\mathrm{id}_{\mathbb{K}^n} : \mathbb{K}^n \to \mathbb{K}^n$ ist: $L_{I_n} = \mathrm{id}_{\mathbb{K}^n}$.

Beispiel 3.2.13. Für alle $m \leq n$ sind die Inklusion $\iota : \mathbb{K}^m \to \mathbb{K}^n$ und die Projektion $\pi : \mathbb{K}^n \to \mathbb{K}^m$

$$\iota : \mathbb{K}^m \to \mathbb{K}^n, \qquad \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \pi : \mathbb{K}^n \to \mathbb{K}^m, \qquad \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix}$$

lineare Abbildungen. Wir können dies durch eine direkte Berechnung beweisen, aber wir können auch sehen, dass dies die linearen Abbildungen sind, die den Matrizen $\left(\begin{smallmatrix} I_m \\ 0_n \end{smallmatrix}\right)$ und $(I_m | 0_n)$ entsprechen, womit I_m die $m \times m$ Identitätsmatrix ist und 0_n die $n \times n$ Nullmatrix ist.

Beispiel 3.2.14. Seien $\theta \in [0, 2\pi)$ und $R_\theta \in \mathbb{R}^{2 \times 2}$ die Matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Wir betrachten die Abbildung $L_\theta : \mathbb{R}^2 \to \mathbb{R}^2$. Wir schreiben ein Vektor $v \in \mathbb{R}^2$ in Polarkoordinaten (siehe Tafel) als

$$v = \begin{pmatrix} r \cdot \cos\alpha \\ r \cdot \sin\alpha \end{pmatrix},$$

so dass

$$R_\theta \cdot v = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} r \cdot \cos\alpha \\ r \cdot \sin\alpha \end{pmatrix} = r \cdot \begin{pmatrix} \cos\theta \cdot \cos\alpha - \sin\theta \sin\alpha \\ \sin\theta \cdot \cos\alpha + \cos\theta \cdot \sin\alpha \end{pmatrix} = r \cdot \begin{pmatrix} \cos(\theta + \alpha) \\ \sin(\theta + \alpha) \end{pmatrix}$$

Das zeigt, dass L_{R_θ} eine Drehung von Winkel θ ist.

Lemma 3.2.15. *Seien U, V zwei $\mathbb{K}$-Vektorräume und $f : U \to V$ eine lineare Abbildung.*

1. $f(0) = 0$.

2. $f(\lambda_1 u_1 + \dots + \lambda_n u_n) = \lambda_1 \cdot f(u_1) + \dots + \lambda_n \cdot f(u_n)$ *für alle* $\lambda_i \in \mathbb{K}, v_i \in U$

3. *Ist* f *bijektiv, so ist* $f^{-1}\colon V \to U$ *auch eine* $\mathbb{K}$*-lineare Abbildung.*

Beweis. 1. $f(0) = f(0+0) = f(0) + f(0)$ so dass $f(0) = 0$.

2. $f(\lambda_1 u_1 + \dots + \lambda_n u_n) = f((\lambda_1 u_1 + \dots + \lambda_{n-1} u_{n-1}) + \lambda_n u_n) = f(\lambda_1 u_1 + \dots + \lambda_{n-1} u_{n-1}) + f(\lambda_n u_n) = f(\lambda_1 u_1 + \dots + \lambda_{n-1} u_{n-1}) + \lambda_n f(u_n) = \dots = \lambda_1 f(u_1) + \dots + \lambda_n f(u_n)$.

3. Seien $v, v' \in V$ und $\lambda \in \mathbb{K}$. Wir wollen zeigen dass $f^{-1}(v + v') = f^{-1}(v) + f^{-1}(v')$ und $f^{-1}(\lambda \cdot v) = \lambda \cdot f^{-1}(v)$. Da f injektiv ist, es reicht zu zeigen dass $f(f^{-1}(v + v')) = f(f^{-1}(v) + f^{-1}(v'))$ und $f(f^{-1}(\lambda \cdot v)) = f(\lambda \cdot f^{-1}(v))$. Wir rechnen dass

$$f(f^{-1}(v + v')) = v + v' = f(f^{-1}(v)) + f(f^{-1}(v')) = f(f^{-1}(v) + f^{-1}(v'))$$
$$f(f^{-1}(\lambda \cdot v)) = \lambda \cdot v = \lambda \cdot f(f^{-1}(v)) = f(\lambda \cdot f^{-1}(v))$$

□

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 10.5.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Lemma 3.2.16. *Seien* U, V *zwei* $\mathbb{K}$*-Vektorraume und* $f\colon U \to V$ *und* $f'\colon U \to V$ *zwei* $\mathbb{K}$*-lineare Abbildungen.*

1. $f + f'$ *ist eine* $\mathbb{K}$*-lineare Abbildung.*

2. $\lambda \cdot f$ *ist eine* $\mathbb{K}$*-lineare Abbildung für alle* $\lambda \in \mathbb{K}$.

Seien W *ein* $\mathbb{K}$*-Vektorraum und* $g\colon V \to W, g'\colon V \to W$ *zwei* $\mathbb{K}$*-lineare Abbildungen*

1. $g \circ f$ *ist eine* $\mathbb{K}$*-lineare Abbildung*

2. $g \circ (\lambda f + \lambda' f') = \lambda(g \circ f) + \lambda'(g \circ f')$ *für alle* $\lambda, \lambda' \in \mathbb{K}$.

3. $(\lambda g + \lambda' g') \circ f = \lambda(g \circ f) + \lambda'(g' \circ f)$ *für alle* $\lambda, \lambda' \in \mathbb{K}$.

4. $\lambda(g \circ f) = (\lambda g) \circ f = g \circ (\lambda f)$.

Beweis. 1. Seien $u, u' \in U$ und $\lambda \in \mathbb{K}$. Wir rechnen dass $(f + f')(u + u') = f(u + u') + f'(u + u') = f(u) + f(u') + f'(u) + f'(u') = (f + f')(u) + (f + f')(u')$ und $(f + f')(\lambda u) = f(\lambda u) + f'(\lambda u) = \lambda f(u) + \lambda f'(u) = \lambda \cdot (f + f')(u)$.

2. Seien $u, u' \in U$ und $\mu \in \mathbb{K}$. Wir rechnen dass $(\lambda \cdot f)(u + u') = \lambda \cdot f(u + u') = \lambda \cdot (f(u) + f(u')) = \lambda f(u) + \lambda f(u') = (\lambda \cdot f)(u) + (\lambda \cdot f)(u')$ und $(\lambda \cdot f)(\mu \cdot u) = \lambda \cdot f(\mu \cdot u) = \lambda \cdot \mu f(u) = \mu \cdot \lambda f(u) = \mu \cdot (\lambda f)(u)$.

3. Seien $u, u' \in U$ und $\lambda \in \mathbb{K}$. Dann $(g \circ f)(u + u') = g(f(u + u')) = g(f(u) + f(u')) = g(f(u)) + g(f(u')) = (g \circ f)(u) + (g \circ f)(u')$, und $(g \circ f)(\lambda u) = g(f(\lambda u)) = g(\lambda f(u)) = \lambda g(f(u)) = \lambda \cdot (g \circ f)(u)$.

4. folgt durch Nachrechnen.

5. folgt durch Nachrechnen.

6. folgt durch Nachrechnen.

□

Bemerkung 3.2.17. Wir haben insbesondere gezeigt dass die Summe von zwei lineare Abbildungen und die Multiplikation einer linearen Abbildung mit einem Skalar, noch lineare Abbildungen ist. Insbesondere, seien V, W zwei $\mathbb{K}$-Vektorräume. Dann sind die Abbildungen

$$\begin{aligned} \mathrm{Hom}_{\mathbb{K}}(V,W) \times \mathrm{Hom}_{\mathbb{K}}(V,W) &\to \mathrm{Hom}_{\mathbb{K}}(V,W) & \qquad (f,g) &\mapsto f+g \\ \mathbb{K} \times \mathrm{Hom}_{\mathbb{K}}(V,W) &\to \mathrm{Hom}_{\mathbb{K}}(V,W) & \qquad (\lambda, f) &\mapsto \lambda \cdot f \end{aligned}$$

wohldefiniert. Die Menge $\mathrm{Hom}_{\mathbb{K}}(V,W)$ mit diesen zwei Abbildungen ist ein $\mathbb{K}\mathbb{K}$-Vektorraum.

Bemerkung 3.2.18. Wir haben gezeigt dass die Komposition von lineare Abbildungen eine lineare Abbildung ist und dass die inverse Abbildung einer bijektiven linearen Abbildung noch eine lineare Abbildung ist.

Definition 3.2.19 (Isomorphismus, Endomorphismus, Automorphismus)**.** Seien V, W $\mathbb{K}$-Vektorräume. Eine bijektive lineare Abbildung $f\colon V \to W$ heißt Isomorphismus. Eine lineare Abbildung $f\colon V \to V$ heißt Endomorphismus von V, und eine bijektive lineare Abbildung $f\colon V \to V$ heißt Automorphismus von V. Die Menge von Endomorphismen ist $\mathrm{End}_K(V)$ und die Menge von Automorphismen ist $GL(V)$.

Beispiel 3.2.20. Für jede $\lambda \in \mathbb{K}$ ist die Abbildung $m_\lambda\colon V \to V, v \mapsto \lambda \cdot v$ ein Endomorphismus. Wenn $\lambda \neq 0$, dann ist m_λ ein Automorphismus, mit inverse Abbildung $m_{\lambda^{-1}}$.

3.3 Untervektorräume

Definition 3.3.1 (Untervektorraum)**.** Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $U \subseteq V$ heißt Untervektorraum (Unterraum), wenn U mit der Einschränkung der Addition und Skalarmultiplikation selbst ein Vektorraum ist.

Proposition 3.3.2. *Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $U \subseteq V$ ist ein Untervektorraum genau dann, wenn*

1. $0 \in U$.

2. $v + w \in U$ *für alle* $v, w \in U$.

3. $\lambda \cdot v \in U$ *für alle* $\lambda \in \mathbb{K}, v \in U$

Beweis. ($\Longrightarrow$) Die Einschränkungen der Vektoraddition und der Skalarmultiplikation sind

$$+\colon U \times U \to V, \qquad \cdot \mathbb{K} \times U \to V$$

Da U mit diesen ein Vektorraum ist, muss es sein, dass diese Funktionen Bilder in U haben, was bedeutet, dass Eigenschaften (2) und (3) gelten. Außerdem existiert $e \in U$ so dass $e + u = u$ für alle $u \in U$. Insbesondere $e + e = e$ und wenn wir diese Gleichung in der Vektorraum V betrachten, sehen wir dass $e = 0$.

($\Longleftarrow$) Eigenschaften (2) und (3) zeigen dass die Einschränkungen von der Addition und der Skalarmultiplikation zwei Abbildungen

$$+: U \times U \to U, \qquad \cdot \mathbb{K} \times U \to U$$

definieren. Wir zeigen dass U mit diesen ein $\mathbb{K}$-Vektorraum ist.

- $(U, +)$ ist eine abelsche Gruppe. Die Addition ist auf U assoziativ und kommutativ, da sie auf dem gesamten V assoziativ und kommutativ ist. Das neutrales Element ist $0 \in U$ und wenn $u \in U$, dann $-u = (-1) \cdot u \in U$.
- Die Eigenschaften der skalaren Multiplikation gelten für U, da sie für das gesamte V gelten.

$\square$

Beispiel 3.3.3. Jeder Vektorraum V hat die Untervektorräume V und $\{0\}$.

Beispiel 3.3.4. Sei V ein $\mathbb{K}$-Vektorraum und sei $v \in V$ ein Vektor. Die Menge

$$\langle v \rangle = \{\lambda \cdot v \,|\, \lambda \in \mathbb{K}\}$$

ist ein Untervektorraum von V. Um das zu zeigen, sehen wir dass $0 = 0 \cdot v \in \langle v \rangle$. Außerdem, wenn $v_1, v_2 \in \langle v \rangle$, dann existieren $\lambda_1, \lambda_2 \in \mathbb{K}$ so dass $v_i = \lambda_i \cdot v$, für $i = 1, 2$. Wir sehen dass $v_1 + v_2 = \lambda_1 \cdot v + \lambda_2 \cdot v = (\lambda_1 + \lambda_2) \cdot v \in \langle v \rangle$. Wenn $\mu \in \mathbb{K}$, dann $\mu \cdot v_1 = \mu \cdot (\lambda \cdot v) = (\mu\lambda_1) \cdot v \in \langle v \rangle$.

Beispiel 3.3.5. Für jede $n \in \mathbb{N}$ ist die Menge $\mathbb{K}[x]_{\leq n} = \{f \in \mathbb{K}[x] \,|\, \deg(f) \leq n\}$ ein Untervektorraum von $\mathbb{K}[x]$: wir sehen zuerst dass 0 Grad $-\infty$ ist, so dass 0 Grad kleiner als n ist. Seien auch $f = \sum_{i=0}^{n} a_i x^i$ und $g = \sum_{i=0}^{n} b_i x^i$ zwei Polynome von Grad kleiner oder gleich als n und sei $\lambda \in \mathbb{K}$. Dann haben $f + g = \sum_{i=0}^{n} (a_i + b_i) x^i$ und $\lambda \cdot f = \sum_{i=0}^{n} \lambda \cdot a_i x^i$ Grad kleiner oder gleich als n.

Definition 3.3.6 (Kern)**.** Sei $f: V \to W$ eine lineare Abbildung von $\mathbb{K}$-Vektorräume. Der Kern von f ist

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

Proposition 3.3.7. *Sei $f: V \to W$ eine lineare Abbildung von $\mathbb{K}$-Vektorräume. Der Kern* $\operatorname{Ker}(f)$ *ist ein Untervektorraum von V und das Bild* $\operatorname{Im}(f)$ *ist ein Untervektorraum von W.*

Beweis. Wir zeigen dass der Kern ein Untervektorraum von V ist. Wir sehen dass $0 \in \operatorname{Ker}(f)$, weil $f(0) = 0$. Seien $v, v' \in \operatorname{Ker}(f)$ und $\lambda \in \mathbb{K}$. Dann $f(v + v') = f(v) + f(v') = 0 + 0 = 0$ und $f(\lambda v) = \lambda f(v) = \lambda \cdot 0 = 0$, so dass $v + v' \in \operatorname{Ker}(f)$ und $\lambda \cdot v \in \operatorname{Ker}(f)$.

Wir zeigen dass das Bild ein Untervektorraum von W ist. Wir sehen dass $0 \in \operatorname{Im} f$, weil $f(0) = 0$. Seien $w, w' \in \operatorname{Im}(f)$ und $\lambda \in \mathbb{K}$. Dann existieren $v, v' \in V$ so dass $f(v) = w, f(v') = w'$ und wir rechnen dass $f(v + v') = f(v) + f(v') = w + w'$ und $f(\lambda \cdot v) = \lambda \cdot f(v) = \lambda \cdot w$, so dass $w + w' \in \operatorname{Im}(f)$ und $\lambda \cdot w \in \operatorname{Im} f$. $\square$

Bemerkung 3.3.8. Sei $f: V \to W$ eine lineare Abbildung von $\mathbb{K}$-Vektorräume. Das Urbild $f^{-1}(w)$ ist ein Untervektorraum von V genau dann, wenn $w = 0$. Wenn $w = 0$ haben wir gezeigt dass $f^{-1}(0) = \operatorname{Ker} f$ ein Untervektorraum ist. Wenn $f^{-1}(w)$ ein Untervektorraum ist, dann $0 \in f^{-1}(w)$, so dass $w = f(0) = 0$.

Beispiel 3.3.9. Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix und $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$ die entsprechende lineare Abbildung. Dann

$$\mathrm{Los}(A,b) = \{x \in \mathbb{K}^n \mid A\cdot x = b\} = \{x \in \mathbb{K}^n \mid L_A(x) = b\} = L_A^{-1}(b)$$

Insbesondere, ist $\mathrm{Los}(A,0) = L_A^{-1}(0) = \operatorname{Ker} L_A$ ein Untervektorraum.

Beispiel 3.3.10. Die Teilmenge $\{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 \mid x - y = 0\}$ ist ein Unterrvektorräum von $\mathbb{R}^2$. Die Teilmenge $\{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 \mid x - y = 1\}$ ist kein Unterrvektorräum.

Der Kernel charakterisiert, ob eine lineare Abbildung injektiv ist oder nicht:

Proposition 3.3.11. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-Vektorräume. Die lineare Abbildung f ist injektiv, genau dann, wenn* $\mathrm{Ker}(f) = \{0\}$.

Beweis. ($\Longrightarrow$) Nehmen wir an, dass f injektiv ist. Sei $v \in \mathrm{Ker}(f)$, so dass $f(v) = 0$. Da f eine lineare Abbildung ist, wissen wir dass $f(0) = 0$, so dass $f(v) = f(0)$. Da f injektiv ist, folgt dass $v = 0$.

($\Longleftarrow$) Nehmen wir an dass $\mathrm{Ker}(f) = \{0\}$. Seien $v, v' \in V$ so dass $f(v) = f(v')$. Dann $f(v - v') = f(v) - f(v') = 0$, so dass $v - v' \in \operatorname{Ker} f$. Das bedeutet $v - v' = 0$, so dass $v = v'$. ☐

Allgemeiner ausgedrückt, können wir zeigen, dass die Vorbilder einzelner Elemente durch eine lineare Anwendung Translationen des Kerns sind.

Proposition 3.3.12. *Sei $f\colon V \to W$ eine lineare Abbildung von $\mathbb{K}$-Vektorräume und sei $w \in W$. Wenn $f^{-1}(w) \neq \emptyset$, sei $v \in f^{-1}(w)$. Dann*

$$f^{-1}(w) = v + \mathrm{Ker}(f) = \{v + u \mid u \in \mathrm{Ker}(f)\}$$

Beweis. Wir zeigen die zwei Inklusionen: ($\subseteq$) sei $v' \in f^{-1}(w)$, so dass $f(v') = w = f(v)$. Dann $f(v' - v) = f(v') - f(v) = 0$, und das bedeutet dass $v' - v = u \in \operatorname{Ker} f$. Dann $v' = v + u \in v + \mathrm{Ker}(f)$.

($\supseteq$) Sei $u \in \mathrm{Ker}(f)$. Dann $f(v + u) = f(v) + f(u) = w + 0 = w$, so dass $v + u \in f^{-1}(w)$. ☐

Bemerkung 3.3.13. Sei $\mathbb{K}$ ein Körper, und sei $A \in \mathbb{K}^{m\times n}$ eine Matrix. Wir betrachten die lineare Abbildung $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$. Wir wissen dass $\operatorname{Ker} L_A = \mathrm{Los}(A,0)$ die Losungsmenge des homogenen Systems $Ax = 0$ ist. Sei $b \in \mathbb{K}^m$, und sei $x_0 \in \mathrm{Los}(A,b)$ eine Losüng des Systems $Ax = b$. Wir wissen da $\mathrm{Los}(A,b) = L_A^{-1}(b)$ so dass die vorherige Proposition zeigt dass

$$\mathrm{Los}(A,b) = x_0 + \mathrm{Los}(A,0)$$

Das bedeutet, dass man alle Lösungen des Systems $Ax = b$ erhält, indem man zu einer einzigen Lösung x_0 alle Lösungen des homogenen linearen Systems $Ax = 0$ addiert.

Beispiel 3.3.14. Die Teilmengen $\{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 \mid x - y = b\}$ mit $b \in \mathbb{R}^2$, sind Translationen der Menge $\{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 \mid x - y = 0\}$ (siehe Tafel).

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 14.05.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Wir zeigen nun, dass einige natürliche Konstruktionen mit Untervektorräumen andere Untervektorräume ergeben

Proposition 3.3.15. *Seien V ein $\mathbb{K}$-Vektorraum und seien $U_1, U_2 \subseteq V$ Untervektorräume. Dann ist $U_1 \cap U_2$ ein Untervektorraum. Im allgemein, sei $(U_i)_{i\in I}$ eine Familie von Untervektorräume $U_i \subseteq V$, dann ist $\cap_{i\in I} U_i$ ein Untervektorraum.*

Beweis. Wir beweisen direkt den Fall einer beliebige Familie. Da U_i ein Untervektorraum ist, $0 \in U_i$ für alle $i \in I$, so dass $0 \in \cap_{i\in I} U_i$ auch. Seien $v, v' \in \cap_{i\in I} U_i$ und $\lambda \in \mathbb{K}$. Dann $u + u' \in U_i, \lambda \cdot u \in U_i$, da U_i ein Untervektorraum ist. Das bedeutet dass $u + u' \in \cap_{i\in I} U_i$ und $\lambda \cdot u \in \cap_{i\in I} U_i$. Das zeigt dass $\cap_{i\in I} U_i$ ein Untervektorraum ist. □

Bemerkung 3.3.16. Die Vereinigung von Unterektorräume ist nicht immer ein Untervektorräum. Z.B. in $\mathbb{R}^2$ die Vereinigung $\langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle \cup \langle \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \rangle$ ist nicht ein Vektorraum: die zwei Vektoren $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ sind in der Vereinigung, die Summe $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ aber nicht.

Definition 3.3.17 (Summe von Untervektorräum)**.** Sei V ein $\mathbb{K}$-Vektorraum, und seien $U_1, U_2 \subseteq V$ zwei Untervektorräume. Die Summe von der Unterkvektorräume ist

$$U_1 + U_2 = \{v_1 + v_2 \mid v_1 \in U_1, v_2 \in U_2\}.$$

Im allgemein, wenn $U_1, \ldots, U_n \subseteq V$ endliche viele Teilmenge von V sind, definieren wir

$$U_1 + U_2 + \cdots + U_n = \{v_1 + \cdots + v_n \mid v_i \in U_i\}$$

Proposition 3.3.18. *Sei V ein $\mathbb{K}$-Vektorraum, und seien $U_1, \ldots, U_n \subseteq V$ Untervektorräume. Die Summe $U_1 + \cdots + U_n$ ist ein Untervektorraum von V. Außerdem, gilt dass $U_1 \subseteq U_1 + U_2 \subseteq \cdots \subseteq U_1 + \cdots + U_n$.*

Beweis. Durch Induktion auf n. Für $n = 1$ ist das klar, für $n = 2$ sehen wir dass $0 \in U_1, 0 \in U_2$, so dass $0 = 0 + 0 \in U_1 + U_2$. Seien $v, v' \in U_1 + U_2, \lambda \in \mathbb{K}$. Dann $v = u_1 + u_2$ und $v' = u_1' + u_2'$ für $u_j, u_j' \in U_j$ für $j = 1, 2$. Dann $v + v' = u_1 + u_2 + u_1' + u_2' = (u_1 + u_1') + (u_2 + u_2') \in U_1 + U_2$ und $\lambda \cdot v = \lambda \cdot (u_1 + u_2) = \lambda \cdot v_1 + \lambda \cdot v_2 \in U_1 + U_2$. Das zeigt dass $U_1 + U_2$ ein Untervektorräum ist. Wenn $n > 2$, können wir $U_1 + \cdots + U_n = (U_1 + \cdots + U_{n-1}) + U_n$ schreiben. Die Induktionsannahme zeigt dass $U_1 + \cdots + U_{n-1}$ ein Untervektorraum ist, und dann was wir für $n = 2$ gezeigt haben, zeigt dass $(U_1 + \cdots + U_{n-1}) + U_n$ ein Untervektorraum ist.

Wir müssen noch die Kette von Inklusionen beweisen: wir zeigen $U_1 \subseteq U_1 + U_2$: da $0 \in U_2$, sehen wir dass $v_1 = v_1 + 0 \in U_1 + U_2$ für alle $v_1 \in U_1$, so dass $U_1 \subseteq U_1 + U_2$. Eine änliche Begrundung zeigt dass $U_1 + \dot{+} U_i \subseteq U_1 + \cdots + U_i + U_{i+1}$ für $i \geq 1$. □

Beispiel 3.3.19. Wir betrachten in $\mathbb{R}^2$ die Untervektorräume: $\langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle = \left\{ \left(\begin{smallmatrix} \lambda_1 \\ 0 \end{smallmatrix}\right) \mid \lambda_1 \in \mathbb{K} \right\}$ und $\langle \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \rangle = \left\{ \left(\begin{smallmatrix} 0 \\ \lambda_2 \end{smallmatrix}\right) \mid \lambda_2 \in \mathbb{K} \right\}$. Die Summe ist

$$\langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle + \langle \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \rangle = \left\{ \left(\begin{smallmatrix} \lambda_1 \\ \lambda_2 \end{smallmatrix}\right) \mid \lambda_i \in \mathbb{K} \right\} = \mathbb{R}^2$$

Wir können die Summe von beliebige viele Untervektorräume auch definiieren:

Definition 3.3.20 (Summe von beliebige viele Untervektorräume)**.** Seien V ein $\mathbb{K}$-Vektorraum und $(U_i)_{i\in I}$ eine Familie von Untervektorräume $U_i \in I$. Die Summe ist die Menge von alle mögliche endliche Summen von Vektoren in den U_i:

$$\sum_{i\in I} U_i = \{v_{i_1} + \cdots + v_{i_n} \mid v_{i_j} \in U_{i_j}, i_j \in I, n \geq 0\} = \bigcup_{\substack{J \subseteq I \\ J \text{ endlich}}} \sum_{j\in J} U_j$$

Lemma 3.3.21. *Seien V ein $\mathbb{K}$-Vektorraum und $(U_i)_{i\in I}$ eine Familie von Untervektorräume $U_i \subseteq V$. Die Summe $\sum_{i\in I} U_i$ ist ein Untervektorraum. Außerdem, wenn die Familie endliche ist, stimmnt diese Definition der Summe mit der vorhergehenden überein.*

Beweis. Für jede $i \in I$, gilt dass $0 \in U_i$, und $U_i \subseteq \sum_{i\in I} U_i$, so dass $0 \in \sum_{i\in I} U_i$ auch. Seien $v, v' \in \sum_{i\in I} U_i$, dann sind $v = u_{i_1} + \cdots + u_{i_n}$ für $u_{i_j} \in U_{i_j}$ und $v' = u_{i'_1} + \cdots + u_{i'_{n'}}$ für $u_{i'_j} \in U_{i'_j}$ endliche Summen von Vektoren in den U_i. Dann ist

$$v + v' = u_{i_1} + \cdots + u_{i_n} + u_{i'_1} + \cdots + u_{i'_{n'}}$$

auch eine endliche Summe von Vektoren in den U_i, so dass $v + v' \in \sum_{i\in I} U_i$. Wenn $\lambda \in \mathbb{K}$, dann ist auch

$$\lambda \cdot v = \lambda \cdot u_{i_1} + \cdots + \lambda \cdot u_{i_n}$$

und $\lambda \cdot u_{i_j} \in U_{i_j}$ weil U_{i_j} ein Untervektorraum ist. Das zeigt dass $\lambda \cdot v$ auch eine endliche Summe von Vektoren in den U_i ist, so dass $\lambda \cdot v \in \sum_{i\in I} U_i$. Das zeigt dass $\sum_{i\in I} U_i$ ein Untervektorraum ist.

Falls $I = \{i_1, \ldots, i_n\}$ endlich ist, wir müssen noch zeigen dass $\bigcup_{J\subseteq I J \text{ endlich}} \sum_{j\in J} U_j$ und die Summe $\sum_{i\in I} U_i$ von Proposition 3.3.18 gleich sind. Für jede endliche Teilmenge $J \subseteq I$, Proposition 3.3.18 zeigt dass $\sum_{j\in J} U_j \subseteq \sum_{i\in I} U_i$, so dass $\bigcup_{\substack{J\subseteq I \\ J \text{ endlich}}} \left(\sum_{j\in J} U_j\right) \subseteq \sum_{i\in I} U_i$. Andererseits, die Menge I ist selbst endlich, so dass $\sum_{i\in I} U_i \subseteq \bigcup_{\substack{J\subseteq I \\ J \text{ endlich}}} \left(\sum_{j\in J} U_j\right)$. □

Definition 3.3.22 (Direkte Summe)**.** Sei V ein $\mathbb{K}$-Vektorraum und seien $U, W \subseteq V$ Untervektorräume. Der Vektorraum V ist die direkte Summe von U, W wenn

1. $V = U + W$.
2. $U \cap W = \{0\}$, für alle i.

Man schreibt $V = U \oplus W$.

Lemma 3.3.23. *Sei V ein $\mathbb{K}$-Vektorraum und seien U, W Untervektorräume. Der Vektorraum V ist die direkte Summe von U, W genau dann, wenn für alle $v \in V$ eindeutige $u \in U, w \in W$ existieren so dass $v = u + w$.*

Beweis. ($\Longrightarrow$) Nehmen wir an dass $V = U \oplus W$ und sei $v \in V$. Da $V = U + W$, existieren $u \in U, w \in W$ so dass $v = u + w$. Seien $u' \in U, w' \in W$ so dass $u + w = u' + w'$. Dann $u - u' = w' - w$ und, da U, W Untervektorräume sind, sehen wir dass $u - u' \in U, w' - w \in W$. Dann $u - u' = w' - w \in U \cap W = \{0\}$ und $u = u', w = w'$.

($\Longleftarrow$) Da jeder Vektor in V eine Summe eines Vektors in U und eines Vektors in W ist, sehen wir dass $V = U + W$. Wir müssen zeigen dass $U \cap W = \{0\}$. Sei $z \in U \cap W$. Wir schreiben $z = z + 0 = 0 + z$ als Summe eines Vektors in U und eines Vektors in W. Da die Summanden eindeutig sind, sehen wir dass $z = 0$. □

Bemerkung 3.3.24. Sei V ein $\mathbb{K}$-Vektorraum und seien $U, W \subseteq V$ Untervektorräume. Wenn $U \cap W = \{0\}$, dann gilt $U + W = U \oplus W$. Wir betrachten dann $U \oplus W$ als Untervektorraum von V.

Auf diese Weise können wir definieren, was es bedeutet, dass V die direkte Summe einer beliebigen endlichen Anzahl von Vektorräumen $U_1, \dots, U_n$ ist: Es bedeutet, dass V die direkte Summe der beiden Vektorräume $(U_1 \oplus \cdots \oplus U_{n-1})$ und U_n ist. Wir schreiben

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

Es kann bewiesen werden, dass dies der Tatsache entspricht, dass jeder Vektor $v \in V$ eindeutig als Summe von Vektoren $u_i \in U_i$ geschrieben werden kann: $v = u_1 + \cdots = u_n$.

3.4 Erzeuger, Lineare Abhängigkeit, Basen

Definition 3.4.1 (Linearkombinationen)**.** Sei V ein $\mathbb{K}$-Vektorraum. Sei $(v_1, \dots, v_n)$ eine Tupel von Vektoren $v_i \in V$. Eine Linearkombination der v_i ist eine Summe

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n, \qquad \lambda_i \in \mathbb{K}.$$

Im Allgemein, sei $(v_i)_{i \in I}$ eine, möglicherweise unendlich Familie von Vektoren $v_i \in V$. Eine Linearkombination der v_i ist eine Linearkombination von endliche viele Vektoren in der Familie. Anders gesagt, das ist eine Summe

$$\sum_{i \in I} \lambda_i v_i$$

mit $\lambda_i \neq 0$ für endliche viele $i \in I$.

Bemerkung 3.4.2. Die Tatsache, dass wir eine Linearkombination eines Tupels und nicht einer Menge definieren, erlaubt uns zum Beispiel zu sagen, dass $2 \cdot v + 3 \cdot v$ eine Linearkombination von (v, v) ist.

Beispiel 3.4.3. Seien $v_1, \dots, v_n \in \mathbb{K}^m$ Vektoren. Wir schreiben

$$v_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad v_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, v_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

so dass eine Linearkombination $\lambda_1 v_1 + \cdots + \lambda_n v_n$ die Form

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = \begin{pmatrix} \lambda_1 a_{11} + \lambda_2 a_{21} + \cdots + \lambda_n a_{1n} \\ \lambda_1 a_{21} + \lambda_2 a_{22} + \cdots + \lambda_n a_{2n} \\ \vdots \\ \lambda_1 a_{m1} + \lambda_2 a_{m2} + \cdots + \lambda_n a_{mn} \end{pmatrix}$$

hat. Sei jetzt $A = (a_{ij}) = (v_1 | v_2 | \dots | v_n)$ die Matrix die $v_1, \dots, v_n$ als Spaltenvektoren hat. Die vorherige Berechnung zeigt dass

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = A \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = (v_1 | v_2 | \dots | v_n) \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Definition 3.4.4 (Erzeugnis, Span)**.** Sei V ein $\mathbb{K}$-Vektorraum und sei $M \subseteq V$ eine Teilmenge. Das Span von M, oder Erzeugnis von M, oder der von M erzeugter Vektorraum ist

$$\langle M \rangle = \bigcap_{\substack{U \subseteq V, \text{Unterraum} \\ M \subseteq U}} U$$

die Schnittmenge von alle Untervektorräume die M enthalten.

Wenn $v_1, \dots, v_n \in I$ sind, schreiben wir $\langle v_1, \dots, v_n \rangle = \langle \{v_1, \dots, v_n\} \rangle$. Im allgemein, wenn $(v_i)_{i \in I}$ eine Familie von Vektoren ist, dann schreiben wir $\langle v_i \,|\, i \in I \rangle = \langle \{v_i \,|\, i \in I\} \rangle$.

Wenn $U \subseteq V$ ein Untervektorraum ist, und wenn $(u_i)_{\in I}$ eine Familie von Vektoren so dass $U = \langle u_i \,|\, i \in I \rangle$ ist, nennen wir die u_i Erzeuger von U und die Familie $(u_i)_{i \in I}$ ist ein Erzeugendensystem von U.

Beispiel 3.4.5. Sei V ein $\mathbb{K}$-Vektorraum und sei $v \in V$ ein Vektor. Wir haben früher der Untervektorraum $\langle v \rangle = \{\lambda \cdot v \,|\, \lambda \in \mathbb{K}\}$ definiert. Wir zeigen dass $\{\lambda \cdot v \,|\, \lambda \in \mathbb{K}\}$ tatsächlich der Untervektorraum erzeugt von v ist. Sei denn $\langle v \rangle$ der Untervektorraum erzeugt von v. Wir wissen schon dass $\{\lambda \cdot v \,|\, \lambda \in \mathbb{K}\}$ ein Untervektorraum der v enthält ist, so dass $\{\lambda \cdot v \,|\, \lambda \in \mathbb{K}\} \supseteq \langle v \rangle$. Andererseits, $\langle v \rangle$ ist ein Untervektorraum, weil er der Durschnitt von Untervektorräume ist. Da $v \in \langle v \rangle$, folgt es dass $\lambda \cdot v \in \langle v \rangle$ für alle $\lambda \in \mathbb{K}$, so dass $\{\lambda \cdot v \,|\, \lambda \in \mathbb{K}\} \subseteq \langle v \rangle$.

Proposition 3.4.6. *Seien V ein $\mathbb{K}$-Untervektorraum.*

1. *Sei $M \subseteq V$ eine Teilmenge. Dann ist $\langle M \rangle$ ein Untervektorraum von V und ist der kleinste Untervektorraum der M enthalt: wenn $U \subseteq V$ ein Untervektorraum ist s.d. $M \subseteq U$ gilt, dann $\langle M \rangle \subseteq U$.*

2. *Seien $v_1, \dots, v_n \in V$ Vektoren. Dann ist $\langle v_1, \dots, v_n \rangle$ die Menge von alle Linearkombinationen den v_i:*
$$\langle v_1, \dots, v_n \rangle = \left\{ \sum_{i=1}^{n} \lambda_i v_i \,\middle|\, \lambda_i \in K \right\} = \sum_{i=1}^{n} \langle v_i \rangle$$

3. *Seien $U_1, \dots, U_n \subseteq V$ Untervektorräume. Dann gilt*
$$\langle U_1 \cup U_2 \cup \dots \cup U_n \rangle = \sum_{i=1}^{n} U_i.$$

Beweis.

1. Erstmals sehen wir dass V selbst ein Untervektorraum der M enthält ist, so dass es gibt mindestens ein Untervektorraum von V der M enthält. Dann sehen wir dass, da $\langle M \rangle$ als Schnittmenge von Untervektorräume definiert ist, ist $\langle M \rangle$ ein Untervektorraum. Endlich, sei $U \subseteq V$ ein Untervektorraum so dass $M \subseteq U$, denn gilt $\langle M \rangle \subseteq U$ dank der Definition von $\langle M \rangle$.

2. Die Menge $\{\sum_{i=1}^n \lambda_i v_i \,|\, \lambda_i \in K\} = \sum_{i=1}^n \langle v_i \rangle$ ist ein Untervektorraum, weil er die Summe von Untervektorräume ist. Außerdem $v_i \in \sum_{i=1}^n \langle v_i \rangle$ für alle i, so dass $\langle v_1, \dots, v_n \rangle \subseteq \sum_{i=1}^n \langle v_i \rangle$. Andererseits, $\lambda_i \cdot v_i \in \langle v_1, \dots, v_n \rangle$ für alle $\lambda_i \in \mathbb{K}$, weil $\langle v_1, \dots, v_n \rangle$ ein Untervektorraum ist, die $v_1, \dots, v_n$ enthält. Das zeigt dass $\langle v_1, \dots, v_n \rangle \supseteq \sum_{i=1}^n \langle v_i \rangle$, so dass $\langle v_1, \dots, v_n \rangle \supseteq \sum_{i=1}^n \langle v_i \rangle$.

3. Wir wissen dass $\sum_{i=1}^{n} U_i$ ein Untervektorraum von V ist, und wir wissen auch dass $U_i \subseteq \sum_{i=1}^{n} U_i$ für alle i, so dass $\langle U_1 \cup \ldots U_n \rangle \subseteq \sum_{i\in I} U_i$. Andererseits, $u_i \in \langle U_1 \cup \cdots \cup U_n \rangle$ für alle $u_i \in U_i$, so dass $u_1 + \cdots + u_n \in \langle U_1 \cup \cdots \cup U_n \rangle$ für alle $u_i \in U_i$, weil $\langle U_1 \cup \cdots \cup U_n \rangle$ ein Untervektorraum ist. Das zeigt dass $\sum_{i=1}^{n} U_i \subseteq \langle U_1 \cup \cdots \cup U_n \rangle$.

□

Diese Proposition gilt auch mit unendliche Familien von Vektoren oder Untervektorräume:

Proposition 3.4.7. [2] *Seien V ein $\mathbb{K}$-Untervektorraum.*

1. *Sei $(v_i)_{i\in I}$ eine Familie von Vektoren $v_i \in V$. Dann ist $\langle v_i \,|\, i \in I \rangle$ die Menge von alle Linearkombinationen den v_i:*

$$\langle v_i \,|\, i \in I \rangle = \left\{ \sum_{i\in I} \lambda_i v_i \,\middle|\, \lambda_i \in \mathbb{K}, \lambda_i \neq 0 \textit{ für endliche viele } i \in I \right\} = \sum_{i\in I} \langle v_i \rangle$$

2. *Seien $(U_i)_{i\in I}$ eine Familie von Untervektorräume $U_i \subseteq V$. Dann gilt*

$$\left\langle \bigcup_{i\in I} U_i \right\rangle = \sum_{i\in I} U_i.$$

Beweis.

1. Die Menge $\left\{ \sum_{i\in I} \lambda_i v_i \,|\, \lambda_i \in K \lambda_i \neq 0 \text{ für endliche viele } i \in I \right\} = \sum_{i=1}^{n} \langle v_i \rangle$ ist ein Untervektorraum, weil er die Summe von Untervektorräume ist. Ausserdem $v_i \in \sum_{i\in I} \langle v_i \rangle$ für alle i, so dass $\langle v_i \,|\, i \in I \rangle \subseteq \sum_{i=1}^{n} \langle v_i \rangle$. Andererseits, sei $W \subseteq V$ ein Untervektorraum so dass $v_i \in W$ für alle i. Dann $\lambda_i \cdot v_i \in W$ für alle $\lambda_i \in \mathbb{K}$ und $\sum_{i\in I} \lambda_i v_i \in W$ wenn $\lambda_i \neq 0$ für endliche viele $i \in I$. Das zeigt dass $W \supseteq \sum_{i\in I} \langle v_i \rangle$, so dass $\langle v_i \,|\, i \in I \rangle \supseteq \sum_{i\in I} \langle v_i \rangle$.

2. Wir wissen dass $\sum_{i\in I} U_i$ ein Untervektorraum von V ist, und wir wissen auch dass $U_i \subseteq \sum_{i\in I} U_i$ für alle i, so dass $\cup_{i\in I} U_i \subseteq \sum_{i\in I} U_i$. Sei $W \subseteq V$ ein Untervektorraum so dass $W \supseteq \cup_{i\in I} U_i$. Dann $W \supseteq \cup_{j\in J} U_j$ für alle endliche Teilmenge $J \subseteq I$, so dass $W \supseteq \langle \cup_{j\in J} U_j \rangle$. Da J endlich ist, haben wir schön gezeigt dass $\langle \cup_{j\in J} U_j \rangle = \sum_{j\in J} U_j$, so dass $W \supseteq \sum_{j\in J} U_j$. Das zeigt dass $W \supseteq \bigcup_{\substack{J\subseteq I \\ J \text{ endlich}}} \sum_{j\in J} U_j = \sum_{i\in I} U_i$, so dass $\langle \bigcup_{i\in I} U_i \rangle \supseteq \sum_{i\in I} U_i$.

□

Definition 3.4.8 (Endlich erzeugter Vektorraum)**.** Ein $\mathbb{K}$-Vektorraum V heißt endlich erzeugt wenn endliche viele Vektoren $v_1, \ldots, v_n \in V$ existieren, so dass

$$V = \langle v_1, \ldots, v_n \rangle$$

Beispiel 3.4.9. Sei $\mathbb{K}$ ein Körper. Wir zeigen dass der Vektorraum $\mathbb{K}^n$ endlich erzeugt für alle $n \in \mathbb{N}$ ist. Wenn $n = 0$, $\mathbb{K}^0 = \{0\}$ ist von 0 erzeugt. Wenn $n \geq 1$ und $1 \leq i \leq n$, sei $e_i \in \mathbb{K}$ der Vektor mit Eintrag 1 in Zeile i und Eintrag 0 überall sonst

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \leftarrow \text{ i-te Zeile }, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

[2] Das war nicht in der Vorlesung gezeigt

Dann gilt für alle $x \in \mathbb{K}^n$ dass

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \cdot e_1 + \cdots + x_n \cdot e_n \qquad \text{für } x_i \in \mathbb{K}$$

Das zeigt dass $\mathbb{K}^n = \langle e_1, \ldots, e_n \rangle$.

Beispiel 3.4.10. Sei $\mathbb{K}$ ein Körper. Der Polynomring $\mathbb{K}[x]$ ist nicht endlich erzeugt als $\mathbb{K}$-Vektorraum. Um das zu zeigen, müssen wir beweisen dass für $f_1, \ldots, f_n \in \mathbb{K}[x]$ endliche viele Polynome, der erzeugter Unterraum $\langle f_1, \ldots, f_n \rangle$ eine echte Teilmenge von $\mathbb{K}[x]$ ist: $\langle f_1, \ldots, f_n \rangle \subsetneq \mathbb{K}[x]$. Wir wissen dass

$$\langle f_1, \ldots, f_n \rangle = \left\{ \sum_{i=1}^{n} \lambda_i f_i \,|\, \lambda_i \in \mathbb{K} \right\}$$

Sei $M = \max\{\deg(f_i) \,|\, i = 1, \ldots, n\}$, dann $\deg(\sum_{i=1}^n \lambda_i f_i) \leq M$ für alle $\lambda_i \in \mathbb{K}$. Das bedeutet insbesondere dass $x^{M+1} \notin \langle f_1, \ldots, f_n \rangle$, so dass $\langle f_1, \ldots, f_n \rangle \subsetneq \mathbb{K}[x]$.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 17.05.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Definition 3.4.11 (Lineare Unabhängigkeit)**.** Sei V ein $\mathbb{K}$-Vektorraum und sei $(v_1, \ldots, v_n)$ eine Tupel von Vektoren in V. Die v_i sind linear unabhängig wenn für alle $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gilt

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0 \iff \lambda_i = 0 \qquad \text{für alle } i$$

Dagegen, sind die v_i linear abhängig wenn $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ existieren, nicht alle 0, so dass

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

Im allgemein, ist eine Familie (v_i), moglicherweise unendlich, von Vektoren $v_i \in V$ linear unabhängig, genau dann, wenn

$$\text{eine lineare Kombination } \sum_{i \in I} \lambda_i v_i = 0 \iff \lambda_i = 0 \text{ für alle } i \in I$$

Beispiel 3.4.12. Die Vektoren $e_1, \ldots, e_n \in \mathbb{K}^n$ sind linear unabhängig:

$$\lambda_1 e_1 + \cdots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = 0 \iff \lambda_i = 0 \text{ für alle } i$$

Beispiel 3.4.13. Die Vektoren $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right) \in \mathbb{K}^2$ sind linear abhängig, weil $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, so dass $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = 0$.

Beispiel 3.4.14. Sei V ein $\mathbb{K}$-Vektorraum. Der Vektor 0 ist linear abhängig, da $1 \cdot 0 = 0$. Jede Familie die 0 enthält ist auch linear abhängig. Wenn $v \in V, v \neq 0$,dann ist v linear unabhängig, da $\lambda \cdot v = 0 \implies \lambda = 0$.

Definition 3.4.15 (Basis)**.** Sei V ein $\mathbb{K}$-Vektorraum. Ein Erzeugendensystem $(v_i)_{i\in I}$ von Vektoren die linear unabhängig sind, heißt eine Basis von V.

Beispiel 3.4.16 (Kanonische Basis oder Standardbasis)**.** Die Vektoren $e_1,\dots,e_n \in \mathbb{K}^n$ sind Erzeuger von $\mathbb{K}^n$ und linear unabhängig. Das bedeutet dass $\mathcal{C} = (e_1,\dots,e_n)$ eine Basis von $\mathbb{K}^n$ ist. Wir nennen sie die Kanonische Basis oder Standardbasis von $\mathbb{K}^n$.

Beispiel 3.4.17 (Monomialbasis)**.** Wir betrachten den Ring $\mathbb{K}[x]$ als $\mathbb{K}$-Vektorraum. Die unendliche Familie von Monome $(1, x, x^2, x^3, \dots) = (x^i \,|\, i \geq 0)$ ist eine Basis von $\mathbb{K}[x]$ als $\mathbb{K}$-Vektorraum: Ein Polynom $f \in \mathbb{K}[x]$

$$f = \sum_{i=0}^{n} a_i \cdot x^i$$

ist eine Linearkombination der x^i mit den Koeffizienten a_i, was bedeutet dass $(x^i | i \geq 0)$ ein Erzeugendensystem ist. Außerdem ist ein Polynom $f = \sum_{i=0}^{n} a_i \cdot x^i$ gleich Null genau dann wenn $a_i = 0$ für alle i, was bedeutet dass die x^i linear unabhängig sind.

Beispiel 3.4.18. Wir betrachten der $\mathbb{Q}$-Vektorraum $\mathbb{Q}[x]_{\leq 1} = \{a_0 + a_1 \cdot x \,|\, a_0, a_1 \in \mathbb{Q}\}$ von alle Polynome von Grad kleiner oder gleich als 1. Die zwei Polynome $f_1 = (1+x)$, $f_2 = x$ sind linear unabhängig:

$$\lambda_1 f_1 + \lambda_2 f_2 = \lambda_1(1+x) + \lambda_2 x = \lambda_1 + (\lambda_1+\lambda_2)x = 0 \iff \begin{cases} \lambda_1 & = 0 \\ \lambda_1 + \lambda_2 & = 0 \end{cases} \iff \lambda_1 = \lambda_2 = 0$$

Außerdem, sehen wir dass f_1, f_2 Erzeugen von $\mathbb{Q}[x]_{\leq 1}$ sind: wir können jedes Polynom $a_0 + a_1 x \in \mathbb{Q}[x]_{\leq 1}$ als Linearkombination

$$a_0 + a_1 x = a_0 + (a_1 - a_0 + a_0)x = a_0 \cdot (1+x) + (a_1 - a_0) \cdot x$$

schreiben. Das zeigt dass (f_1, f_2) eine Basis von $\mathbb{Q}[x]_{\leq 1}$ ist.

Lemma 3.4.19. *Sei V ein $\mathbb{K}$-Vektorraum und sei $(v_1,\dots,v_n)$ eine Tupel von Vektoren in V*

1. *Die (v_i) sind ein Erzeugendensystem von V genau dann, wenn jeder Vektor in V eine Linearkombination der v_i ist:*

$$\text{für alle } v \in V \text{ existieren } \lambda_i \in \mathbb{K} \text{ so dass } v = \sum_{i=1}^{n} \lambda_i$$

2. *Die (v_i) sind linear unhabhängig in V genau dann, wenn falls ein Vektor eine Linearkombination der v_i ist, dann ist er eine Linearkombination auf eine eindeutige Weise.*

$$\sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \mu_i v_i \iff \lambda_i = \mu_i \text{ für alle } i$$

3. *Die (v_i) sind eine Basis von V genau dann, wenn jeder Vektor in V eindeutig eine Linearkombination der v_i ist:*

$$\text{für alle } v \in V \text{ existieren eindeutige } \lambda_i \in \mathbb{K} \text{ so dass } v = \sum_{i=1}^{n} \lambda_i v_i$$

Die gleichen Aussagen gelten für eine unendliche Familie von Vektoren

Beweis. Wir geben den Beweis nur für die endliche Familie $(v_1, \dots, v_n)$. Der Fall der unendlichen Familien ist ähnlich.

1. Wir wissen dass $\langle v_1, \dots, v_n\rangle$ die Menge der Linearkombinationen der v_i ist.

2. Wenn die v_i linear unabhängig sind, und wenn $\sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i$, dann $\sum_{i=1}(\lambda_i - \mu_i)v_i = 0$, so dass $\lambda_i = \mu_i$ für alle i. Andererseits, Die Bedingung im Lemma für $\mu_i = 0$ bedeutet genau, dass die v_i linear unabhängig sind.

3. Das folgt aus den vorherigen Punkten.

□

Wir können lineare Unabhängigkeit in Form von Generatoren charakterisieren und umgekehrt. Wir brauchen zunächst ein Lemma

Lemma 3.4.20. *Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, \dots, v_n, w \in V$ Vektoren. Es gibt eine Linearkombination $\lambda_1 v_1 + \dots + \lambda_n v_n + \mu w = 0$ mit $\mu \neq 0$ genau dann, wenn $w \in \langle v_1, \dots, v_n\rangle$.*

Beweis. ($\Longrightarrow$) Wir schreiben $\mu w = -\lambda_1 v_1 - \dots - \lambda_n v_n$, so dass $w = -\mu^{-1}\lambda_1 v_1 - \dots - \mu^{-1}\lambda_n v_n$. Das zeigt dass $w \in \langle v_1, \dots, v_n\rangle$

($\Longleftarrow$) Wenn $w \in \langle v_1, \dots, v_n\rangle$, existieren dann $\lambda_i \in \mathbb{K}$ so dass $w = \lambda_1 v_1 + \dots + \lambda_n v_n$. Das bedeutet, dass $\sum_{i=1}^n \lambda_i v_i + (-1)\cdot w = 0$. □

Lemma 3.4.21. *Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, \dots, v_n \in V$ Vektoren. Die Folgende Aussagen sind Äquivalent*

1. *$\mathcal{B} = (v_1, \dots, v_n)$ ist eine Basis von V.*

2. *$\mathcal{B}$ ist eine maximale Familie von linear unabhängige Vektoren. Das bedeutet dass die $v_1, \dots, v_n$ linear unabhängig sind, und dass $v_1, \dots, v_n, v_{n+1}$ linear abhängig für alle $v_{n+1} \in V$ sind.*

3. *$\mathcal{B}$ ist eine minimales Erzeugendensystem. Das bedeutet dass $V = \langle v_1, \dots, v_n\rangle$ und $V \neq \langle v_1, \dots, \hat{v_i}, \dots, v_n\rangle$ für alle $i \in \{1, \dots, n\}$.*

Änliche Aussagen gelten für eine beliebige Familie $(v_i)_{i\in I}$ von Vektoren.

Beweis. Wir geben den Beweis nur für eine endliche Familie $(v_1, \dots, v_n)$.

(1) $\Longrightarrow$ (2) : Wir wissen dass $\mathcal{B}$ eine linear unabhängige Familie ist. Sei $w \in V$: Da $\mathcal{B}$ ein Erzeugendensystem ist, das vorheriges Lemma zeigt dass eine Linearkombination $\sum_{i=1}^n \lambda_i v_i + \mu w = 0$ existiert, mit $\mu \neq 0$. Das zeigt dass $(v_1, \dots, v_n, w)$ linear abhängig sind.

(2) $\Longrightarrow$ (3) : Wir zeigen zuerst dass $\mathcal{B}$ ein Erzeugendensystem ist. Sei $v_{n+1} \in V$. Die $v_1, \dots, v_{n+1}$ sind linear abhängig, so dass $\sum_{i=1}^{n+1} \lambda_i v_i = 0$ für $\lambda_i \in \mathbb{K}$ nicht alle null. Wenn $\lambda_{n+1} = 0$, dann $\sum_{i=1}^n \lambda_i v_i = 0$ und $\lambda_i = 0$ für alle i, Widerspruch. Dann $\lambda_i \neq 0$, und das vorheriges Lemma zeigt dass $v_{n+1} \in \langle v_1, \dots, v_n\rangle$. Das zeigt dass $V = \langle v_1, \dots, v_n\rangle$. Wir zeigen jetzt dass $\mathcal{B}$ ein minimales Erzeugendensystem ist: es reicht zu Zeigen dass $v_i \notin \langle v_1, \dots, \hat{v_i}, \dots, v_n\rangle$ für alle $i = 1, \dots, n$. Nehmen wir an, dass $v_i \in \langle v_1, \dots, \hat{v_i}, \dots, v_n\rangle$, dann zeigt das vorheriges Lemma dass $\sum_{j=1}^n \lambda_j v_j = 0$ für $\lambda_j \in \mathbb{K}$ mit $\lambda_i \neq 0$. Widerspruch, da die Vektoren in $\mathcal{B}$ linear unabhängig sind.

(3) $\implies$ (1) : Wir müssen zeigen dass die Vektoren in $\mathcal{B}$ linear unabhängig sind. Seien $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ nicht alle null, so dass $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Wenn $\lambda_i \neq 0$, zeigt das vorheriges Lemma dass $v_i \in \langle v_1, \dots, \hat{v_i}, \dots, v_n \rangle$, Widerspruch. □

Satz 3.4.22. *Sei V ein $\mathbb{K}$-Vektorraum. Jedes endliche Erzeugendensystem von V enthält eine Basis. Insbesondere, hat jeder endlich erzeugte Vektorraum $V \neq \{0\}$ eine Basis.*

Beweis. Sei $\mathcal{B} = (v_1, \dots, v_n)$ ein endliches Erzeugendensystem von V. Wenn $\mathcal{B}$ ein minimales Erzeugendensystem ist, dann ist es auch eine Basis. Wenn nicht, existiert v_i, mit $i = 1, \dots, n$ so dass $\mathcal{B}' = (v_1, \dots, \hat{v_i}, \dots, v_n)$ ein Erzeugendensystem ist. Wir können dieses Verfahren wiederholen bis wir ein minimales Erzeugendensystem finden, das eine Basis ist, oder bis wir alle Vektoren herausnehmen, aber in diesem Fall wird der Vektorraum durch eine leere Menge erzeugt, was bedeutet, dass $V = \{0\}$. □

Bemerkung 3.4.23. Dieses Satz gilt für beliebige Erzeugendensysteme von V und insbesondere hat jeder Vektorraum eine Basis, wenn man das Auswahlaxiom akzeptiert. Wir würden das aber nicht beweisen.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 28.5.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, \dots, v_n \in V$. Wir betrachten den Tupel $(v_1, \dots, v_n)$ und die folgende Umformungen:

- $(GS1)$ Vertauschen zweier Vektoren.
- $(GS2)$ Addition des λ-fachen eines Vektors zu einem anderen, verschiedenen Vektor, $\lambda \in \mathbb{K}$.
- $(GS3)$ Multiplikation eines Vektor mit $\lambda \in \mathbb{K}, \lambda \neq 0$.

Proposition 3.4.24. *Sei $(v_1, \dots, v_n)$ eine Tupel von Vektoren und sei $(v'_1, \dots, v'_n)$ ein Tupel, das man durch Anwendung einer Folge von endlich vielen Elementarumformungen wie oben erhält. Dann*

1. $\langle v_1, \dots, v_n \rangle = \langle v'_1, \dots, v'_n \rangle$.
2. *$v_1, \dots, v_n$ sind linear unhabhängig genau dann, wenn $v'_1, \dots, v'_n$ linear unhabhängig sind.*
3. *$(v_1, \dots, v_n)$ ist eine Basis von V genau dann, wenn $(v'_1, \dots, v'_n)$ eine Basis von V ist.*

Beweis. Es genügt, die Aussagen zu beweisen, wenn $(v'_1, \dots, v'_n)$ aus $(v_1, \dots, v_n)$ durch eine einzige elementare Umformung erhalten wird. Wir prüfen die drei verschiedenen Arten von Umformungen.

- $(GS1)$ Angenommen,$(v'_1, \dots, v'_n)$ ergibt sich aus $(v_1, \dots, v_n)$ durch Vertauschen zweier Vektoren. Dann sind die drei Aussagen klar.
- $(GS2)$ Angenommen,$(v'_1, \dots, v'_n) = (v_1, \dots, v_i, \dots, v'_j, \dots, v_n)$ womit $v'_j = v_j + \lambda v_i$. Wir zeigen dass $\langle v'_1, \dots, v'_n \rangle \subseteq \langle v_1, \dots, v_n \rangle$: ein Vektor in $\langle v'_1, \dots, v'_n \rangle$ hat die Form $\sum_{i=1}^n \mu_i v'_i$ mit $\mu_i \in \mathbb{K}$. Wir sehen dass

$$\sum_{i=1}^{n} \mu_i v'_i = \left(\sum_{i=1}^{n} \mu_i \cdot v_i \right) + \mu_j \lambda v_i = \mu_1 v_1 + \dots + (\mu_i + \lambda \mu_j) v_i + \dots + \mu_n v_n$$

so dass dieses auch eine Linearkombination der v_i ist. Das zeigt dass $\langle v_1', \ldots, v_n' \rangle \subseteq \langle v_1, \ldots, v_n \rangle$. Da wir auch $(v_1, \ldots, v_n)$ erhalten, indem wir eine elementare Umformung vom Typ $(GS2)$ auf $(v_1', \ldots, v_n')$ anwenden, sehen wir, dass auch der andere Inklusion gilt. Dies beweist, dass $\langle v_1, \ldots, v_n \rangle = \langle v_1', \ldots, v_n' \rangle$.

Wir zeigen jetzt dass wenn die $v_1, \ldots, v_n$ linear unhabhängig sind, dann sind die $v_1', \ldots, v_n'$ auch linear unhabhängig. Seien $\mu_1, \ldots, \mu_n \in \mathbb{K}$ so dass $\sum_{i=1}^n \mu_i v_i' = 0$. Die vorherige Rechnung zeigt dass

$$\mu_1 v_1 + \cdots + (\mu_i + \lambda \mu_j) v_i + \cdots + \mu_n v_n = 0$$

und da die $v_1, \ldots, v_n$ linear unhabhängig sind, sehen wir dass $\mu_h = 0$ für alle $h \neq i$ und $\mu_i = -\lambda \mu_j = -\lambda \cdot 0$. Das zeigt dass die $v_1', \ldots, v_n'$ linear unhabhängig sind. Da wir auch $(v_1, \ldots, v_n)$ erhalten, indem wir eine elementare Umformung vom Typ $(GS2)$ auf $(v_1', \ldots, v_n')$ anwenden, sehen wir, dass wenn die $v_1', \ldots, v_n'$ lineare unhabhängig sind, dann sind die $v_1, \ldots, v_n$ auch linear unhabhängig.

Wenn wir schließlich zusammenfassen, was wir zuvor bewiesen haben, sehen wir, dass $(v_1, \ldots, v_n)$ eine Basis ist, genau dann, wenn $(v_1', \ldots, v_n')$ eine Basis ist.

- $(GS3)$ Wir nehmen an dass $(v_1', \ldots, v_n') = (v_1, \ldots, v_j, \ldots, v_n)$ womit $v_j' = \lambda v_j$ mit $\lambda \neq 0$. Wir zeigen dass $\langle v_1', \ldots, v_n' \rangle \subseteq \langle v_1, \ldots, v_n \rangle$: ein Vektor in $\langle v_1', \ldots, v_n' \rangle$ hat die Form $\sum_{i=1}^n \mu_i v_i$ mit $\mu_i \in \mathbb{K}$. Wir sehen dass

$$\sum_{i=1}^n \mu_i v_i' = \left(\sum_{i=1}^n \mu_i \cdot v_i \right) + \mu_j \lambda v_i = \mu_1 v_1 + \cdots + \lambda \mu_j v_i + \cdots + \mu_n v_n$$

so dass dieses auch eine Linearkombination der v_i ist. Das zeigt dass $\langle v_1', \ldots, v_n' \rangle \subseteq \langle v_1, \ldots, v_n \rangle$. Da wir auch $(v_1, \ldots, v_n)$ erhalten, indem wir eine elementare Umformung vom Typ $(GS3)$ auf $(v_1', \ldots, v_n')$ anwenden, sehen wir, dass auch der andere Inklusion gilt. Dies beweist, dass $\langle v_1, \ldots, v_n \rangle = \langle v_1', \ldots, v_n' \rangle$.

Wir zeigen jetzt dass wenn die $v_1, \ldots, v_n$ linear unhabhängig sind, dann sind die $v_1', \ldots, v_n'$ auch linear unhabhängig. Seien $\mu_1, \ldots, \mu_n \in \mathbb{K}$ so dass $\sum_{i=1}^n \mu_i v_i' = 0$. Die vorherige Rechnung zeigt dass

$$\mu_1 v_1 + \cdots + \lambda \mu_j v_i + \cdots + \mu_n v_n = 0$$

und da die $v_1, \ldots, v_n$ linear unhabhängig sind, sehen wir dass $\mu_h = 0$ für alle $h \neq j$ und $\lambda \mu_i = 0$. Da $\lambda \neq 0$, folgt dass $\mu_j = 0$. Das zeigt dass die $v_1', \ldots, v_n'$ linear unhabhängig sind. Da wir auch $(v_1, \ldots, v_n)$ erhalten, indem wir eine elementare Umformung vom Typ $(GS3)$ auf $(v_1', \ldots, v_n')$ anwenden, sehen wir, dass wenn die $v_1', \ldots, v_n'$ lineare unhabhängig sind, dann sind die $v_1, \ldots, v_n$ auch linear unhabhängig.

Wenn wir schließlich zusammenfassen, was wir zuvor bewiesen haben, sehen wir, dass $(v_1, \ldots, v_n)$ eine Basis ist, genau dann, wenn $(v_1', \ldots, v_n')$ eine Basis ist.

$\square$

Korollar 3.4.25 (Austauschlemma von Steinitz). *Sei V ein endlich erzeugte $\mathbb{K}$-Vektorraum und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Sei auch $v_j' = \sum_{i=1}^n \lambda_i v_i$ mit $\lambda_j \neq 0$. Dann ist*

$$\mathcal{B}' = (v_1, \ldots v_{j-1}, v_j', v_{j+1} \ldots, v_n)$$

auch eine Basis von V. Außerdem können wir $\mathcal{B}'$ aus $\mathcal{B}$ mit einer Folge von Elementarumforumungen erhalten.

Beweis. Wir können $\mathcal{B}'$ aus $\mathcal{B}$ erhalten, indem wir v_j mit λ_j multiplizieren und dann skalare Vielfache der anderen Vektoren dazu addieren. Dies sind alles elementare Umformungen, und das vorherige Lemma zeigt, dass $\mathcal{B}'$ eine Basis ist. $\square$

Satz 3.4.26 (Austauschsatz von Steinitz). *Sei V ein endlich erzeugte $\mathbb{K}$-Vektorraum, sei $(v_1, \ldots, v_n)$ eine Basis von V und seien $w_1, \ldots, w_r \in V$ linear unhabhängig. Dann*

1. *$r \leq n$.*

2. *Es gibt Vektoren $v_{i_1}, \ldots, v_{i_{n-r}}$ in $\{v_1, \ldots, v_n\}$ so dass $(w_1, \ldots, w_r, v_{i_1}, \ldots, v_{i_{n-r}})$ auch eine Basis von V ist.*

Beweis. Wir beweisen die Aussage durch Induktion auf r. Wenn $r = 0$, dann sind die zwei Aussagen wahr. Wenn $r = 1$, dann ist (w_1) linear unhabhängig, d.h. $w_1 \neq 0$. Insbesondere $V \neq 0$ so dass $n \geq 1$. Da $(v_1, \ldots, v_n)$ eine Basis von V ist, können wir $w_1 = \lambda_1 v_1 + \cdots + \lambda_n v_n$ schreiben, und bis auf eine Umbenennung der v_i können wir annehmen, dass $\lambda_1 \neq 0$. Dann zeigt das Austauschlemma von Steinitz, dass $(w_1, v_2, \ldots, v_n)$ eine Basis von V ist.

Wir nehmen nun an, dass die Aussage für $r \geq 1$ gilt, und wir beweisen, dass sie für $r + 1$ gilt. Die $w_1, \ldots, w_r$ sind auch linear unhabhängig und die Induktionsannahme zeigt dass $r \leq n$ und dass, bis auf eine Umbenennung der v_i, die Tupel $(w_1, w_2, \ldots, w_r, v_{r+1}, \ldots, v_n)$ eine Basis von V ist. Insbesondere

$$w_{r+1} = \sum_{i=1}^{r} \lambda_i w_i + \sum_{i=r+1}^{n} \mu_i v_i$$

Wenn $\mu_i = 0$ für alle $i \in \{n + 1, \ldots, r\}$, dann $w_{r+1} \in \langle w_1, \ldots, w_{r+1} \rangle$, und das ist unmöglich, da die $(w_1, \ldots, w_{r+1})$ linear unhabhängig sind. Das bedeutet dass $\mu_i \neq 0$ für ein $i \in \{r + 1, \ldots, n\}$. Insbesondere $r + 1 \leq n$. Bis auf eine Umbenennung der v_i, können wir annehmen dass $\mu_{r+1} \neq 0$. Das Austauschlemma von Steiniz zeigt dann dass $(w_1, \ldots, w_r, w_{r+1}, v_{r+1}, \ldots, v_n)$ eine Basis von V ist. $\square$

Korollar 3.4.27. *Sei V ein endlich erzeugte $\mathbb{K}$-Vektorraum und seien $w_1, \ldots, w_r \in V$ linear unhabhängig. Dann existieren $w_{r+1}, \ldots, w_n \in V$ so dass $(w_1, \ldots, w_n)$ eine Basis von V ist.*

Beweis. Da V endlich erzeugte ist, hat V eine Basis $(v_1, \ldots, v_n)$. Die Aussage folgt jetzt direkt vom Austauschsatz von Steinitz. $\square$

Satz 3.4.28. *Sei V ein endlich erzeugte $\mathbb{K}$-Vektorraum und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V.*

1. *Seien $w_1, \ldots, w_r \in V$ linear unabhängig. Dann $r \leq n$.*

2. *Seien $w_1, \ldots, w_r \in V$ Erzeugern von V. Dann $r \geq n$.*

3. *Alle Basen von V sind endlich und haben genau n Vektoren.*

Beweis. 1. Das folgt direkt vom Austauschsatz von Steinitz.

2. Bis zur Umbenennung der w_i wissen wir, dass es $r' \leq r$ gibt, so dass $(w_1, \dots, w_{r'})$ eine Basis von V ist. Da die $v_1, \dots, v_n$ linear unhabhängig sind, zeigt der Austauschsatz von Steinitz dass $n \leq r'$. Insbesondere $n \leq r$.

3. Punkt (2) zeigt dass eine Basis mindestens n Vektoren haben muss, und Punkt (1) zeigt dass eine Basis maximal n Vektoren haben kann.

$\square$

Bemerkung 3.4.29. Dieses Satz zeigt, dass zwei beliebige Basen eines endlich erzeugten Vektorraums die gleiche Kardinalität haben. Dies gilt auch für nicht endlich erzeugte Vektorräume, aber wir werden es nicht beweisen.

Definition 3.4.30 (Dimension)**.** Sei V ein $\mathbb{K}$-Vektorraum. Die Mächtigkeit einer Basis von V heißt die Dimension von V. Wir schreiben $\dim_{\mathbb{K}} V$ oder nur $\dim V$. Ein Vektorraum heißt endlich-dimensionale wenn $\dim_{\mathbb{K}} V$ endlich ist.

Beispiel 3.4.31. $\dim \mathbb{K}^n = n$.

Proposition 3.4.32. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, mit $\dim_{\mathbb{K}} V = n$ und seien $v_1, \dots, v_n$ Vektoren. Die folgende Aussage sind äquivalent:*

1. *$(v_1, \dots, v_n)$ ist eine Basis.*

2. *$v_1, \dots, v_n$ sind linear unhabhängig.*

3. *$v_1, \dots, v_n$ sind ein Erzeugendensystem.*

Beweis. (1) $\implies$ (2) Die Vektoren in einer Basis sind immer linear unhabhängig.

(2) $\implies$ (3) Wir wissen, dass wir die v_i immer zu einer Basis vervollständigen können, aber da eine Basis genau n Elemente haben muss, muss sie bereits eine Basis sein. Insbesondere sind sie ein System von Generatoren.

(3) $\implies$ (1) Wir wissen, dass wir immer eine Teilmenge von v_i finden können, die eine Basis ist. Da eine Basis genau n Elemente haben muss, muss sie bereits eine Basis sein. $\square$

3.4.1 Dimension und Unterräume

Unterräume eines Vektorraums haben erwartungsgemäß eine geringere Dimension.

Proposition 3.4.33. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und sei $W \subseteq V$ ein Unterraum. Dann*

1. *W ist auch endlich-dimensionaler und $\dim_{\mathbb{K}} W \leq \dim_{\mathbb{K}} V$.*

2. *$W = V$ genau dann, wenn $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} V$.*

Beweis. Sei $n = \dim_{\mathbb{K}} V$.

1. Wir wissen, dass es nicht mehr als n linear unabhängige Vektoren in V geben kann, so dass dies auch für W gelten muss. Insbesondere kann jede Basis von W höchstens n Elemente haben. Dies beweist, dass W eine endliche Dimension hat und dass $\dim_{\mathbb{K}} W \leq n$.

2. Wenn $\dim_{\mathbb{K}} W = n$, dann existiert eine Basis $(w_1, \ldots, w_n)$ von W mit n Vektoren. Diese sind auch n linear unhabhängige Vektoren in V, so dass $(w_1, \ldots, w_n)$ auch eine Basis von V ist. Das zeigt dass $W = \langle w_1, \ldots, w_n \rangle = V$.

$\square$

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 31.5.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Definition 3.4.34. Sei V ein $\mathbb{K}$-Vektorraum und sei $W \subseteq V$ ein Untervektorraum. Ein Komplement von W in V ist ein Untervektorraum $W' \subseteq V$ so dass $V = W \oplus W'$.

Lemma 3.4.35. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und seien $U, W \subseteq V$ zwei Untervektorräume. Seien auch $\mathcal{B}_U = (u_1, \ldots, u_h)$ und $\mathcal{B}_W = (w_1, \ldots, w_k)$ zwei beliebige Basen von U und W. Dann $V = U \oplus W$ genau dann, wenn $\mathcal{B} = (u_1, \ldots, u_h, w_1, \ldots, w_k)$ eine Basis von V ist.*

Beweis. ($\Longrightarrow$) Wir sehen dass $\mathcal{B}$ ein Erzeugendensystem ist, weil $V = U + W$. Seien $\lambda_1, \ldots, \lambda_h, \mu_1, \ldots, \mu_k \in \mathbb{K}$ so dass $\sum_{i=1}^h \lambda_i v_i + \sum_{j=1}^k \mu_j u_j = 0$. Dann $\sum_{i=1}^h \lambda_i v_i = -\sum_{j=1}^k \mu_j u_j$ ist ein Vektor in $U \cap W = \{0\}$. Das bedeutet dass $\sum_{i=1}^h \lambda_i v_i = \sum_{j=1}^k \mu_j u_j = 0$ und, da die v_i linear unabhängig sind, sehen wir dass $\lambda_i = 0$ für alle i, und da die μ_j linear unabhängig sind, sehen wir dass $\mu_j = 0$ für alle j. Das zeigt dass $\mathcal{B}$ linear unabhängig ist.

($\Longleftarrow$) Da $\mathcal{B}$ ein Erzeugendensystem von V ist, sehen wir dass $V = U + W$. Sei $v \in U \cap W$, dann $v = \sum_{i=1}^h \lambda_i u_i = \sum_{i=1}^k \mu_i w_i$ für manche $\lambda_i, \mu_j \in \mathbb{K}$. Insbesondere gilt dass $\sum_{i=1}^h \lambda_i u_i - \sum_{i=1}^k \mu_i w_i = 0$. Da $\mathcal{B}$ linear unabhängig ist, sehen wir dass $\lambda_i = 0$ und $\mu_j = 0$ für alle i, j, so dass $v = 0$. Das zeigt dass $U \cap W = \{0\}$. $\square$

Korollar 3.4.36. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und sei $U \subseteq V$ ein Untervektorraum.*

1. *U hat ein Komplement in V.*

2. *Sei W ein Komplement von U in V, so dass $V = U \oplus W$. Dann* $\dim U + \dim W = \dim V$.

Beweis.

1. Sei $\mathcal{B}_U = (u_1, \ldots, u_h)$ eine Basis von W. Wir können sie zu einer Basis von V ergänzen: $\mathcal{B}_V = (w_1, \ldots, w_h, w_1, \ldots, w_k)$. Sei $W = \langle w_1, \ldots, w_k \rangle$. Das vorheriges Lemma zeigt dass $V = U \oplus W$.

2. Sei $\mathcal{B}_U = (u_1, \ldots, u_h)$ eine Basis von U und sei $\mathcal{B}_W = (w_1, \ldots, w_h)$ eine Basis von W. Das vorheriges Lemma zeigt dass $\mathcal{B} = (u_1, \ldots, u_h, w_1, \ldots, w_k)$ eine Basis von V ist. Insbesondere $\dim V = h + k = \dim U + \dim W$.

$\square$

Bemerkung 3.4.37. Ein Komplement ist nicht eindeutig. Z.B der Untervektorraum $W = \langle e_1 \rangle$ von $\mathbb{R}^2$ hat ein Komplement $W'_\lambda = \langle e_1 + \lambda e_2 \rangle$ für alle $\lambda \in \mathbb{R}, \lambda \neq 0$.

Als nächstes erklären wir eine wichtige Formel für die Dimension der Summe von zwei Unterräumen

Satz 3.4.38. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und seien $U, W \subseteq V$ zwei Untervektorräume. Dann*

$$\dim_{\mathbb{K}}(U+W) + \dim_{\mathbb{K}} U \cap W = \dim_{\mathbb{K}} U + \dim_{\mathbb{K}} W$$

Beweis. Sei U' ein Komplement von $U \cap W$ in U und sei W' ein Komplement von $U \cap W$ in W, so dass $U = (U \cap W) \oplus U'$ und $W = (U \cap W) \oplus W'$. Wir zeigen nun dass $U + W = (U \cap V) \oplus U' \oplus W'$. Wir sehen zuerst dass

$$U + W = ((U \cap W) + U') + ((U \cap W) + W') = U \cap W + U' + W'.$$

Wir wissen schon dass $(U \cap V) \cap U' = \{0\}$. Wir müssen noch zeigen dass $((U \cap V) \oplus U') \cap W' = U \cap W' = 0$. Wir sehen dass $U \cap W' \subseteq U \cap W$, so dass

$$U \cap W' = U \cap W' \cap U \cap W = W' \cap (U \cap W) = \{0\}.$$

Das zeigt dass $U + W = (U \cap V) \oplus U' \oplus W'$: dann sehen wir dass

$$\begin{aligned}
\dim(U+W) &= \dim(U \cap W) + \dim U' + \dim W' \\
&= \dim(U \cap W) + (\dim U - \dim(U \cap W)) + (\dim W - \dim(U \cap W)) \\
&= \dim U + \dim W - \dim(U \cap W)
\end{aligned}$$

□

3.4.2 Lineare Abbildungen und Basen

Proposition 3.4.39. *Seien V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume, $\mathcal{B}_V = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{T} = (w_1, \ldots, w_n)$ eine Tupel von Vektoren in W.*

1. *Es gibt eine einzige lineare Abbildung $f\colon V \to W$ so dass $f(v_i) = w_i$ für alle i.*
2. $\operatorname{Im} f = \langle w_1, \ldots, w_n \rangle$.
3. *f ist injektiv genau dann, wenn $\mathcal{T}$ linear unabhängig ist.*
4. *f ist ein Isomorphismus genau dann, wenn $\mathcal{T}$ eine Basis von W ist.*

Beweis. 1. Sei $v \in V$. Dann existieren eindeutige $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ so dass $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Wir definieren dann $f(\sum_{i=1}^n \lambda_i v_i) \colon = \sum_{i=1}^n \lambda_i w_i$. Dies definiert eine Abbildung

$$f\colon V \to W$$

und wir sehen dass $f(v_i) = w_i$. Wir müssen zeigen dass f eine lineare Abbildung ist. Seien $v = \sum_{i=1}^n \lambda_i v_i, v' = \sum_{i=1}^n \lambda'_i v_i$, und $\lambda \in \mathbb{K}$. Dann

$$f(v+v') = f\left(\sum_{i=1}^n (\lambda_i + \lambda'_i) v_i\right) = \sum_{i=1}^n (\lambda_i + \lambda'_i) w_i = \sum_{i=1}^n \lambda_i w_i + \sum_{i=1}^n \lambda'_i w_i = f(v) + f(v').$$

$$f(\lambda v) = f\left(\sum_{i=1}^n \lambda \lambda_i v_i\right) = \sum_{i=1}^n (\lambda \lambda_i) w_i = \lambda \cdot \sum_{i=1}^n \lambda_i w_i = \lambda \cdot f(w).$$

Das zeigt dass f linear ist. Sei nun $g\colon V \to W$ eine lineare Abbildung so dass $g(v_i) = w_i$ für alle i. Sei $v = \sum_{i=1}^n \lambda_i v_i$ ein Vektor in V. Dann

$$g(v) = g(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i g(v_i) = \sum_{i=1}^n \lambda_i w_i = f(v)$$

so dass $f = g$.

2. Wir sehen dass

$$\operatorname{Im} f = \{f(v) \mid v \in V\} = \left\{ f(\sum_{i=1}^n \lambda_i v_i) \mid \lambda_i \in \mathbb{K} \right\} = \left\{ \sum_{i=1}^n \lambda_i w_i \mid \lambda_i \in \mathbb{K} \right\} = \langle w_1, \ldots, w_n \rangle$$

3. Die lineare Abbildung f ist injektiv genau dann, wenn $\ker f = \{0\}$. Das bedeutet dass $f(\sum_{i=1}^n \lambda_i v_i) = 0$ genau dann, wenn $\lambda_i = 0$ für alle i, oder, anders gesagt, dass $\sum_{i=1}^n \lambda_i w_i = 0$ genau dann, wenn $\lambda_i = 0$ für alle i. Das bedeutet genau dass $w_1, \ldots, w_n$ linear unabhängig sind.

4. Das folgt aus Punkt (2) und Punkt (3).

□

Bemerkung 3.4.40. Seien V, W zwei $\mathbb{K}$-Vektorräume, $\mathcal{B}_V = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{T} = (w_1, \ldots, w_n)$ eine Tupel von Vektoren in W. Wir schreiben $\Phi_{\mathcal{T}}^{\mathcal{B}}\colon V \to W$ für die einzige lineare Abbildung so dass $\Phi_{\mathcal{T}}^{\mathcal{B}}(v_i) = w_i$ für alle i.

Korollar 3.4.41. *Zwei zwei endlich-dimensionale $\mathbb{K}$-Vektorräume V, W sind isomorph genau dann, wenn $\dim_V = \dim W$.*

Beweis. ($\Longrightarrow$) Sei $f\colon V \to W$ ein Isomorphismus, sei $\mathcal{B}_V = (v_1, \ldots, v_n)$ eine Basis von V und seien $w_i = f(v_i)$ für alle i. Die vorherige Proposition zeigt dass $(w_1, \ldots, w_n)$ eine Basis von W ist, so dass $\dim V = n \dim W$.

($\Longleftarrow$) Sei $n = \dim V = \dim W$ und seien $\mathcal{B}_V = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{B}_W = (w_1, \ldots, w_n)$ eine Basis von W. Sei $\Phi_{\mathcal{B}_W}^{\mathcal{B}_V}\colon V \to W$ die lineare Abbildung so dass $\Phi_{\mathcal{B}_W}^{\mathcal{B}_V}(v_i) = w_i$. Die vorherige Proposition zeigt dass $\Phi_{\mathcal{B}_W}^{\mathcal{B}_V}$ ein Isomorphismus ist. □

Satz 3.4.42. *Seien V, W endlich-dimensionale $\mathbb{K}$-Vektorräume und sei $f\colon V \to W$ eine lineare Abbildung. Dann*

$$\dim V = \dim \operatorname{Ker} f + \dim \operatorname{Im} f.$$

Beweis. Sei $U \subseteq V$ ein Komplement vom $\operatorname{Ker} f$, so dass $V = (\operatorname{Ker} f) \oplus U$. Wir zeigen dass $f_{|U}\colon U \to \operatorname{Im} f$ ein Isomorphismus ist.

Wir zeigen zuerst dass $f_{|U}$ surjektiv ist: sei $w \in \operatorname{Im} f$, dann $w = f(v)$ für ein $v \in V$ und wir können $u \in U, z \in \operatorname{Ker} f$ finden so dass $v = u + z$. Dann $w = f(v) = f(u+z) = f(u) + f(z) = f(u) + 0 = f(u)$. Das zeigt dass $f_{|U}$ surjektiv ist.

Wir zeigen nun dass $f_{|U}$ injektiv ist, oder, äquivalent, dass $\operatorname{Ker} f_{|U} = 0$. Sei $u \in U$ so dass $f(u) = 0$, dann $u \in \operatorname{Ker} f$, so dass $u \in U \cap (\operatorname{Ker} f) = \{0\}$.

Das zeigt dass $f_{|U}\colon U \to \mathrm{Imf}$ ein Isomorphismus ist. Insbesondere $\dim \operatorname{Im} f = \dim U = \dim V - \dim \operatorname{Ker} f$. □

Kapitel 4

Lineare Abbildungen und Matrizen

4.1 Matrizenmultiplikation

Definition 4.1.1 (Matrizenmultiplikation)**.** Sei $\mathbb{K}$ ein Körper und seien $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{n\times \ell}$ zwei Matrizen, mit Einträge $A = (a_{ij}), B = (b_{ij})$. Die Matrix $AB \in \mathbb{K}^{m\times \ell}$ ist die Matrix mit einträge

$$(AB)_{ij} = \sum_{h=1}^{n} a_{ih} \cdot b_{hj}$$

Bemerkung 4.1.2. Das Produkt AB ist definiert nur wenn die Anzahl von Spalten von A und die Anzahl von Zeilen von B gleich ist.

Bemerkung 4.1.3. Wenn $B = b \in \mathbb{K}^{n\times 1}$ ein Spaltenvektor ist, das Produkt Ab ist das übliche Matrix-Vektor produkt. Außerdem, wenn $B \in \mathbb{K}^{n\times \ell}$ ist, seien $B^1, \dots, B^\ell \in \mathbb{K}^n$ die Spalten von B, so dass $B = (B^1|B^2|\dots|B^\ell)$. Dann

$$A \cdot B = A \cdot (B^1|B^2|\dots|B^\ell) = (AB^1|AB^2|\dots|AB^\ell)$$

Wir können auch die Zeilen $A_1, \dots, A_m \in \mathbb{K}^{1\times n}$ betrachten und dann

$$A \cdot B = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \cdot B = \begin{pmatrix} A_1 B \\ \vdots \\ A_m B \end{pmatrix}$$

Endlich, können wir auch schreiben

$$A \cdot B = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \cdot (B^1|B^2|\dots|B^\ell) = \begin{pmatrix} A_1B^1 & A_1B^2 & \dots & A_1B^\ell \\ A_2B^1 & A_2B^2 & \dots & A_2B^\ell \\ \vdots & \vdots & \ddots & \vdots \\ A_mB^1 & A_mB^2 & \dots & A_mB^\ell \end{pmatrix}$$

so dass die Eintrage $(AB)_{ij}$ von AB das Produkt von der i-the Zeile von A mit der j-te Spalte von B ist:

$$(AB)_{ij} = A_i B^j.$$

Beispiel 4.1.4. Wir nehmen als Beispiel die Multiplikation zweier Matrizen mit Koeffizienten in $\mathbb{Q}$:

$$\begin{pmatrix}1 & 2 & -1\\ 0 & 3 & 1\end{pmatrix}\cdot\begin{pmatrix}2 & 9\\ 1 & -1\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1\cdot 2+2\cdot 1+(-1)\cdot 0 & 1\cdot 9+2\cdot(-1)+(-1)\cdot 1\\ 0\cdot 2+3\cdot 1+1\cdot 0 & 0\cdot 9+3\cdot(-1)+1\cdot 1\end{pmatrix}$$
$$= \begin{pmatrix}4 & 6\\ 3 & -2\end{pmatrix}$$

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 4.6.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Definition 4.1.5 (Transponierte Matrix)**.** Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix. Die transponierte Matrix $A^t \in \mathbb{K}^{n\times m}$ is die Matrix mit Einträge

$$(A^t)_{ij} = A_{ji}$$

Bemerkung 4.1.6. Die Zeilen bzw. die Spalten von A^t sind die Spalten bzw. die Zeilen von A.

Beispiel 4.1.7. Wir nehmen als Beispiel eine Matrix in $\mathbb{R}^{2\times 3}$

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

Wir führen einige Eigenschaften der Matrixmultiplikation auf

Lemma 4.1.8. *Sei $\mathbb{K}$ ein Körper*

1. *Die Multiplikation von Matrizen ist assoziativ: für alle $A \in \mathbb{K}^{n\times m}, B \in \mathbb{K}^{m\times \ell}, C \in \mathbb{K}^{\ell\times p}$ gilt*
$$(A\cdot B)\cdot C = A\cdot(B\cdot C)$$
2. *Sei $A \in \mathbb{K}^{m\times n}$, dann*
$$I_m\cdot A = A\cdot I_n = A$$
3. *Seien $A, A' \in \mathbb{K}^{n\times m}, B, B' \in \mathbb{K}^{m\times \ell}$ und $\lambda, \lambda', \mu, \mu' \in \mathbb{K}$. Dann*
$$\begin{aligned}(\lambda A+\lambda' A')\cdot B &= \lambda(A\cdot B)+\lambda'(A\cdot B)\\ A\cdot(\mu B+\mu' B') &= \mu(A\cdot B)+\mu'(A\cdot B)\end{aligned}$$
4. *Seien $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{n\times \ell}$. Dann*
$$(AB)^t = B^t A^t$$

Beweis. Einfache Berechnung. □

Bemerkung 4.1.9. Diese Eigenschaften zeigen dass die Menge von Quadratmatrizen $\mathbb{K}^{n\times n}$ mit der Summe und der Multiplikation von Matrizen ein Ring mit Eins ist. Das Eins ist die Identitätsmatrix I_n. Der Ring ist im Allgemein nicht Kommutativ, da die Matrizenmultiplikation nicht kommutativ ist.

Definition 4.1.10 (Invertierbare Matrix)**.** Eine Quadratmatrix $A \in \mathbb{K}^{n\times n}$ heißt invertierbar wenn eine Matrix $B \in \mathbb{K}^{n\times n}$ existiert so dass

$$AB = BA = I_n$$

Die Matrix B heißt die inverse Matrix von A. Die Menge von invertierbare Matrizen ist $GL_n(\mathbb{K})$.

Lemma 4.1.11. *Sei $\mathbb{K}$ ein Körper.*

1. *Sei $A \in GL_n(\mathbb{K})$. Dann hat A eine einzige Inverse. Wir schreiben A^{-1} für die inverse Matrix.*
2. *Sei $A \in GL_n(\mathbb{K})$ invertierbar und seien $B, C \in KK^{n\times m}$ und $D, E \in \mathbb{K}^{m\times n}$ Matrizen. Dann*
$$AB = C \iff B = A^{-1}C, \qquad DA = E \iff D = EA^{-1}$$
3. *Seien $A, B \in GL_n(\mathbb{K})$,dann $AB \in GL_n(\mathbb{K})$, und $(AB)^{-1} = B^{-1}A^{-1}$.*
4. *Sei $A \in GL_n(\mathbb{K})$, dann $A^t \in GL_n(\mathbb{K})$, und $(A^t)^{-1} = (A^{-1})^t$.*

Beweis.

1. Seien B, B' zwei Inversen von A. Dann $B = B \cdot I_n = B \cdot (A \cdot B') = (B \cdot A) \cdot B' = I_n \cdot B' = B'$.
2. Wir sehen dass $AB = C \implies A^{-1}AB = A^{-1}C \implies I_n \cdot B = A^{-1}C \implies B = A^{-1}C$. Ähnliche Berechnungen zeigen die andere Implikationen.
3. Wir berechnen $(AB)B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$, so dass $B^{-1}A^{-1} = (AB)^{-1}$.
4. Wir sehen dass $(A^{-1})^tA^t = (AA^{-1})^t = I_n^t = I_n$, so dass $(A^{-1}) = (A^t)^{-1}$

□

4.2 Die Matrix einer lineare Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$

Sei $\mathbb{K}$ ein Körper. Wir betrachten den Vektorraum $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ von allen lineare Abbildungen $\mathbb{K}^m \to \mathbb{K}^n$. Für jede Matrix $A \in \mathbb{K}^{n\times m}$ haben wir eine lineare Abbildung $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$. Wir erinnern uns dass, wenn $A^1, \ldots, A^n \in \mathbb{K}^m$ die Spalten von A sind, dann

$$L_A(x) = x_1A^1 + \cdots + x_nA^n \qquad \text{für alle } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

Insbesondere, wenn $(e_1, \ldots, e_n)$ die kanonische Basis von $\mathbb{K}^n$ ist, gilt das $L_A(e_i) = A^i$ für alle $i \in \{1, \ldots, n\}$.

Satz 4.2.1. *Sei $\mathbb{K}$ ein Körper*

1. *Die Abbildung*
$$L \colon \mathbb{K}^{m\times n} \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m), \qquad A \mapsto L_A$$
ist ein Isomorphismus von $\mathbb{K}$-Vektorräume.

2. *Seien $A \in \mathbb{K}^{m\times n}, B \in \mathbb{K}^{n\times \ell}$. Dann $L_{AB} = L_A \circ L_B$.*

3. *Sei $A \in \mathbb{K}^{m\times n}$. Dann ist L_A ein Isomorphismus genau dann wenn $m = n$ und A invertierbar ist. In diesem Fall gilt $L_A^{-1} = L_{A^{-1}}$.*

Beweis. 1. Wir beweisen zuerst dass L eine lineare Abbildung ist: seien $A, B \in \mathbb{K}^{m\times n}$ und sei $\lambda \in \mathbb{K}$. Dann für alle $x \in \mathbb{K}^n$ gilt dass

$$L_{A+B}(x) = (A + B)x = Ax + Bx = L_A(x) + L_B(x)$$
$$L_{\lambda A}(x) = (\lambda A)(x) = \lambda \cdot Ax = \lambda \cdot L_A(x)$$

Das zeigt dass $L_{A+B} = L_A + L_B$ und $L_{\lambda \cdot A} = \lambda \cdot L_A$, so dass L linear ist. Um zu zeigen dass L_A injektiv ist, beweisen wir dass $\operatorname{Ker} L = \{0\}$. Sei $A \in \mathbb{K}^{m\times n}$ so dass $L_A = 0$. Dann $A^i = L_A(e_i) = 0$ für alle $i \in \{1, \dots, n\}$ so dass $A = 0$.

Wir müssen am Ende zeigen dass A surjektiv ist. Sei $f : \mathbb{K}^n \to \mathbb{K}^m$ eine lineare Abbildung, sei $(e_1, \dots, e_n)$ die kanonische Basis von $\mathbb{K}^n$ und seien $v_i = f(e_i) \in \mathbb{K}^m$ für $i \in \{1, \dots, n\}$. Sei $A = (v_1|\dots|v_n)$ die Matrix die die v_i als Spaltenvektoren hat. Dann $L_A(e_i) = v_i = f(e_i)$ für alle $i \in \{1, \dots, n\}$. Da diese beiden linearen Abbildungen auf einer Basis übereinstimmen, stimmen sie überall überein: $f = L_A$.

2. Sei $x \in \mathbb{K}^n$. Dann $L_{AB}(x) = (AB)x = A(Bx) = L_A(Bx) = L_A(L_B(x))$. Das zeigt dass $L_{AB} = L_A \circ L_B$.

3. Wenn L_A ein Isomorphismus ist, dann wissen wir dass $n = m$. Die inverse Abbildung L_A^{-1} ist auch eine lineare Abbildung, so dass $L_A^{-1} = L_B$ für eine $B \in \mathbb{K}^{n\times n}$. Außerdem

$$L_{AB} = L_A \circ L_B = L_A \circ L_A^{-1} = \mathrm{Id}_{\mathbb{K}^n} = L_{I_n}$$

so dass $AB = I_n$. Eine ähnliche Begrundung zeigt dass $BA = I_n$ so dass A ist invertierbar, $B = A^{-1}$ und $L_A^{-1} = L_{A^{-1}}$.

Andererseits, wenn A invertierbar ist, dann $L_A \circ L_{A^{-1}} = L_{AA^{-1}} = L_{I_n} = \mathrm{Id}_{\mathbb{K}^n}$, so dass L_A ein Isomorphismus ist, und $L_A^{-1} = L_{A^{-1}}$.

□

Bemerkung 4.2.2. Ein wichtiges Nebenprodukt dieses Beweises ist, dass wir, ausgehend von einer linearen Abbildung $f : \mathbb{K}^n \to \mathbb{K}^m$, eine effektive Methode haben, um die einzige Matrix A zu finden, so dass $L_A = f$. Dies ist die Matrix

$$A = (f(e_1)|f(e_2)|\dots|f(e_n))$$

die die Vektoren $f(e_i) \in \mathbb{K}^m$ als Spalten hat.

4.3 Rang

Definition 4.3.1 (Zeilen/Spaltenraum, Zeilen/Spaltenrang)**.** Seien $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{m\times n}$ eine Matrix. Seien $A^1, \dots, A^n$ die Spalten von A und seien $A_1, \dots, A_m$ die Zeilen von A. Der Zeilenraum und der Spaltenraum von A sind die Vektorräume

$$\langle A_1, \dots, A_m\rangle \subseteq \mathbb{K}^{1\times n}, \qquad \langle A^1, \dots, A^n\rangle \subseteq \mathbb{K}^{m\times 1}$$

Der Zeilenrang und der Spaltenrang von A sind

$$\operatorname{zrk}(A) = \dim\langle A_1, \dots, A_m\rangle, \qquad \operatorname{srk}(A) = \dim\langle A^1, \dots, A^n\rangle$$

Bemerkung 4.3.2. Sei $A \in \mathbb{K}^{m\times n}$ und sei $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$ die entsprechende lineare Abbildung. Wir haben schon gesehen dass

$$L_A(x) = \sum_{i=1}^{n} x_i A^i, \qquad \text{für alle } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

so dass der Spaltenraum von A genau das Bild von L_A ist:

$$\operatorname{Im} L_A = \langle A^1, \dots, A^n\rangle, \qquad \operatorname{srk}(A) = \dim \operatorname{Im} L_A.$$

Eine ähnliche Begrundung zeigt dass der Zeilenraum das Bild von L_{A^t} ist

$$\operatorname{Im} L_{A^t} = \langle A_1, \dots, A_m\rangle, , \qquad \operatorname{zrk}(A) = \dim \operatorname{Im} L_A.$$

Satz 4.3.3. *Sei $\mathbb{A} \in \mathbb{K}^{m\times n}$ eine Matrix.*

1. *Elementare Zeilenumformungen oder Spaltenumformungen auf A lassen den Zeilenrang und den Spaltenrang unverändert.*

2. *Wenn A in Zeilenstufenform ist, dann sind der Zeilenrang und der Spaltenrang gleich zu der Anzahl von Pivots von A.*

3. *Wenn A eine beliebige Matrix ist, dann sind der Zeilenrang und der Spaltenrang von A gleich.*

Beweis. 1. Wir betrachten elementare Zeilenumformungen zuerst. Sei A' eine Matrix, die aus A durch eine Folge von elementaren Zeilenumformungen erhalten wird. Proposition 3.4.24 zeigt dass elementare Zeilenumformungen den Zeilenraum unverändert lassen, so dass $\operatorname{zrk}(A') = \operatorname{zrk}(A)$. Dagegen, ist im Allgemein nicht wahr dass elementare Zeilenumformungen den Spaltenraum unverändert lassen. Was wir aber wissen ist dass elementare Zeilenumformungen die Menge $\operatorname{Los}(A, 0) = \operatorname{Ker} L_A$ unverändert lassen. Dann

$$\operatorname{srk}(A') = \dim \operatorname{Im} L_{A'} = n - \dim \operatorname{Ker} L_{A'} = n - \dim \operatorname{Ker} L_A = \dim \operatorname{Im} L_A = \operatorname{srk}(A)$$

Wir bemerken nun, dass elementare Spaltenumformungen auf A elementaren Zeilenumformungen auf A^t entsprechen, so dass sie $\operatorname{zrk}(A^t)$ und $\operatorname{srk}(A^t)$ nicht verändern. Diese sind aber gleich $\operatorname{srk}(A)$ bzw. $\operatorname{zrk}(A)$.

2. Sei A in Zeilenstufenform mit Pivots $a_{1j_1}, a_{2j_2}, \dots, a_{rj_r}$, mit $j_1 < j_2 < \dots < j_r$. Mit elementaren Zeilenoperationen können wir die Matrix in die reduzierte Zeilen-Echelon-Form bringen. Insbesondere ist für alle $i \in \{1, \dots, m\}$ die Spalte j_i überall Null, mit Ausnahme von 1 an der Stelle (i, j_i). Dann können wir für jedes $i \in \{1, \dots, r\}$ und für jedes $j > j_i$ zur j-ten Spalte das a_{ij}-Vielfache der j_i-ten Spalte subtrahieren. Am Ende haben wir eine Matrix, die überall Null ist, mit Ausnahme von 1 an den Positionen

$(1, j_1), \ldots, (r, j_r)$. Dieser Verfahren hat die Anzahl der Pivots nicht verändert, und da wir nur elementare Zeilen- und Spaltenoperationen verwendet haben, hat es auch den Zeilen- und Spaltenrang nicht verändert. Wir können dann direkt annehmen, dass A diese spezielle Form hat. Dann sind der Spaltenraum und der Zeilenraum wie folgt:

$$\langle A^1, \ldots, A^n \rangle = \langle e_1, \ldots, e_r \rangle \subseteq \mathbb{K}^m, \qquad \langle A_1, \ldots, A_n \rangle = \langle e_{j_1}, \ldots, e_{j_r} \rangle \subseteq \mathbb{K}^n$$

und wir sehen, dass diese beiden Vektorräume Dimension r haben.

3. Wir können A durch elementare Zeilenumformungen in Zeilenstufenform bringen. Da elementare Zeilenumformungen den Spaltenraum und der Zeilenraum nicht verändern, können wir annehmen dass A Zeilenstufenform hat. Dann folgt die Aussage aus Punkt 2.

□

Definition 4.3.4 (Rang einer Matrix)**.** Der Rang $\mathrm{rk}(A)$ einer Matrix $A \in \mathbb{K}^{m\times n}$ ist der Spaltenrang oder, äquivalent dazu, der Zeilenrang:

$$\mathrm{rk}(A) := \mathrm{zrk}(A) = \mathrm{srk}(A)$$

Beispiel 4.3.5. Die Identitätsmatrix $I_n \in \mathbb{K}^{n\times n}$ hat Rang $\mathrm{rk}(I_n) = n$. Die Nullmatrix $0 \in \mathbb{K}^{m\times n}$ hat Rang 0. Wir sehen später mehr komplizierte Beispiele.

Proposition 4.3.6. *Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix.*

1. $\mathrm{rk}(A) \le m, \mathrm{rk}(A) \le n$.
2. $\mathrm{rk}(A) = \mathrm{rk}(A^t)$.
3. *Seien $S \in GL_m(\mathbb{K})$ und $T \in GL_n(\mathbb{K})$. Dann* $\mathrm{rk}(SA) = \mathrm{rk}(AT) = \mathrm{rk}(A)$.
4. $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$ *ist injektiv genau dann, wenn* $\mathrm{rk}(A) = n$.
5. $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$ *ist surjektiv genau dann, wenn* $\mathrm{rk}(A) = m$.
6. *A ist invertierbar, genau dann, wenn $m = n$ und* $\mathrm{rk}(A) = n$.

Beweis.

1. Da $\mathrm{rk}(A)$ die Dimension des Spaltenraums $\operatorname{Im} L_A \subseteq \mathbb{K}^m$ ist, sehen wir dass $\mathrm{rk}(A) \le m$. Da $\mathrm{rk}(A)$ die Dimension des Zeilenraum $\operatorname{Im} L_{A^t}\rangle \subseteq \mathbb{K}^n$ ist, sehen wir dass $\mathrm{rk}(A) \le n$.
2. Wir sehen dass $\mathrm{rk}(A) = \mathrm{zrk}(A) = \mathrm{srk}(A^t) = \mathrm{rk}(A^t)$.
3. Wir wissen dass $L_{AT} = L_A \circ L_T$ und das L_T surjektiv ist, so dass $\operatorname{Im} L_{AT} = \operatorname{Im} L_A$. Dann $\mathrm{rk}(AT) = \dim \operatorname{Im} L_{AT} = \dim \operatorname{Im} L_A = \mathrm{rk}(A)$. Dann sehen wir dass S^t invertierbar ist, so dass $\mathrm{rk}(SA) = \mathrm{rk}((SA)^t) = \mathrm{rk}(A^t S^t) = \mathrm{rk}(A^t) = \mathrm{rk}(A)$.
4. L_A ist injektiv genau dann, wenn $\dim \operatorname{Ker} L_A = 0$. Das bedeutet $n = \dim \operatorname{Im} L_A = \mathrm{rk}(A)$.
5. L_A ist surjektiv genau dann, wenn $\mathrm{rk}\, A = \dim \operatorname{Im} L_A = m$.
6. A ist invertierbar genau dann, wenn L_A ein Isomorphismus ist. Dann müssen wir nur einfach letzten beiden Punkte verwenden.

□

4.4 Algorithmen

Wir stellen nun einige Algorithmen vor, die bei expliziten Berechnungen nützlich sind. Sie sind alle auf der Gaußschen Elimination basiert.

4.4.1 Rang einer Matrix

Sein $A \in \mathbb{K}^{m\times n}$ eine Matrix. Wir wollen den Rang von A berechnen.

- Wir verwenden elementare Zeilenumformungen um A in Zeilenstufenform zu bringen.
- Satz 4.3.3 zeigt dass die Anzahl der Pivots gleich dem Rang von A ist.

Beispiel 4.4.1. Wir wenden den Algorithmus in einem Beispiel mit koeffizienten in $\mathbb{R}$ an:

$$A = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 1 & -1 & 1 & 3 \\ 1 & 5 & 3 & -3 \end{pmatrix} \xrightarrow[Z3\to Z3-Z1]{Z2\to Z2-Z1} \begin{pmatrix} 1 & 2 & 2 & 0 \\ 0 & -3 & -1 & 3 \\ 0 & 3 & 1 & -3 \end{pmatrix} \xrightarrow{Z3\to Z3+Z2} \begin{pmatrix} 1 & 2 & 2 & 0 \\ 0 & -3 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Die letzte Matrix ist in Zeilenstufenform und hat zwei Pivots. Dann hat A Rang 2.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 7.6.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

4.4.2 Lineare Gleichungssysteme

Wir können unsere ersten Ergebnisse über lineare Gleichungssysteme mit unseren neuen Konzepten neu interpretieren

Satz 4.4.2. *Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix und sei $b \in \mathbb{K}^n$ ein Vektor.*

1. *Das LGS $Ax = b$ hat eine Lösung genau dann, wenn $\mathrm{rk}(A) = \mathrm{rk}(A|b)$.*
2. *Die Lösungsmenge $\operatorname{Ker} L_A$ des homogenen LGS $Ax = 0$ hat $\dim \operatorname{Ker} L_A = n - \mathrm{rk}(A)$.*
3. *Wenn x_0 eine Lösung des LGS $Ax = b$ ist, dann ist die Menge aller Lösungen $x_0 + \operatorname{Ker} L_A$.*

Beweis.

1. Das LGS $Ax = b$ hat eine Lösung genau dann, wenn $b \in \operatorname{Im} L_A = \langle A^1, \dots, A^n\rangle$. Das ist äquivalent zu $\dim\langle A_1, \dots, A_n\rangle = \dim\langle A_1, \dots, A_n, b\rangle$, und das ist äquivalent zu $\mathrm{rk}(A) = \mathrm{rk}(A|b)$.
2. $\dim \operatorname{Ker} L_A = n - \dim \operatorname{Im} L_A = n - \mathrm{rk}(A)$.
3. Das haben wir schon in Bemerkung 3.3.13 gezeigt.

□

Damit können wir die Menge der Lösungen eines linearen Systems berechnen: Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix und $b \in \mathbb{K}^m$ ein Vektor.

- Wir bringen die Matrix $(A|b)$ durch elementare Zeilenumformungen in Zeilenstufenform. Wir können also annehmen, dass $(A|b)$ in Zeilenstufenform ist. Dann ist die A in Zeilenstufenform auch.

- Wenn A weniger Pivots als $(A|b)$ hat, dann hat das System keine Lösung.
- Wenn A die gleiche Pivots als $(A|b)$ hat, dann bringen wir die Matrix $(A|b)$ durch elementare Zeilenumformungen in reduziertes Zeilenstufenform. Wir können also annehmen, dass $(A|b)$ in Zeilenstufenform ist. Dann ist A in Zeilenstufenform auch mit Pivots gleich 1 in der Stellen $(1, j_1), \ldots, (r, j_r)$, mit $1 \leq j_1 < \cdots < j_r \leq n$.
- Wir verwenden die Gleichungen von $Ax = b$, um $x_{j_1}, \ldots, x_{j_r}$ in Form der anderen Variablen x_h und von $b_1, \ldots, b_r$ zu schreiben.
- Wir können dann x als Linearkombination von Vektoren mit Koeffizienten x_h und einem Vektor, dessen Einträge die b_i sind, schreiben.
- Die Vektoren mit skalaren Koeffizienten x_h sind ein Erzeugendensystem von $\operatorname{Ker} L_A$ und da sie genau $n - r$ sind, müssen sie eine Basis von $\operatorname{Ker} L_A$ sein. Der andere Vektor ist eine Lösung von $Ax = b$.

Wir können dieses Algorithmus verwenden um eine Basis vom $\operatorname{Ker} L_A$ zu finden. Außerdem, wenn wir anstelle eines bestimmten b einen Vektor von Variablen verwenden, können wir diesen Algorithmus verwenden, um das lineare System für alle Vektoren $b \in \mathbb{K}^m$ gleichzeitig zu lösen. Der Algorithmus liefert dann lineare Gleichungen, die $\operatorname{Im} L_A$ in $\mathbb{K}^m$ definieren

Beispiel 4.4.3. Wir wollen alle $b_1, b_2, b_3 \in \mathbb{Q}^3$ beschreiben, so dass das folgendes LGS eine Lösung hat:

$$\begin{cases} x_1 + x_2 + x_3 - x_4 & = b_1 \\ x_1 - x_2 + x_3 + x_4 & = b_2 \\ 3x_1 + x_2 + 3x_3 - x_4 & = b_3 \end{cases}$$

Außerdem, wenn das System eine Lösung hat. Wollen wir alle Lösungen beschreiben. Anders gesagt, wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 3 & 1 & 3 & -1 \end{pmatrix}$$

wir wollen $\operatorname{Im} L_A$ und $\operatorname{Ker} L_A$ beschreiben, und wenn $b \in \operatorname{Im} L_A$ ist, wollen wir eine Lösung von $Ax = b$ finden. Wir anwenden das Algorithmus

$$(A|b) = \begin{pmatrix} 1 & 1 & 1 & -1 & | & b_1 \\ 1 & -1 & 1 & 1 & | & b_2 \\ 3 & 1 & 3 & -1 & | & b_3 \end{pmatrix} \xrightarrow[Z3 \to Z3 - 3Z1]{Z2 \to Z2 - Z1} \begin{pmatrix} 1 & 1 & 1 & -1 & | & b_1 \\ 0 & -2 & 0 & 2 & | & b_2 - b_1 \\ 0 & -2 & 0 & 2 & | & b_3 - 3b_1 \end{pmatrix}$$
$$\xrightarrow{Z3 \to Z3 - Z2} \begin{pmatrix} 1 & 1 & 1 & -1 & | & b_1 \\ 0 & -2 & 0 & 2 & | & b_2 - b_1 \\ 0 & 0 & 0 & 0 & | & b_3 - 2b_1 - b_2 \end{pmatrix}$$

Die erweiterte Koeffizientenmatrix ist in Zeilenstufenform, und die Koeffizientenmatrix hat 2 Pivots. Die erweiterte Koeffizientenmatrix hat 2 Pivots genau dann, wenn $b_3 - 2b_1 - b_2 = 0$. Das zeigt dass

$$\operatorname{Im} L_A = \{b \in \mathbb{Q}^3 \mid b_3 - 2b_2 - b_2 = 0\}$$

Wenn $b \in \operatorname{Im} L_A$, bringen wir die Matrix in reduzierte Zeilenstufenform:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & b_1 \\ 0 & -2 & 0 & 2 & b_2 - b_1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{Z2 \to -\frac{1}{2}Z2} \left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & b_1 \\ 0 & 1 & 0 & -1 & \frac{1}{2}(b_1 - b_2) \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{Z1 \to Z1 - Z2} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & \frac{1}{2}(b_1 + b_2) \\ 0 & 1 & 0 & -1 & \frac{1}{2}(b_1 - b_2) \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Diese Matrix ist in reduzierter Zeilenstufenform mit Pivots in den Spalten 1 und 2. Wir schreiben das lineare System, das dieser Matrix entspricht, und wir schreiben die Variablen x_1, x_2 durch die anderen

$$\begin{cases} x_1 + x_3 & = \frac{1}{2}(b_1 + b_2) \\ x_2 - x_4 & = \frac{1}{2}(b_1 - b_2) \end{cases} \iff \begin{cases} x_1 & = -x_3 + \frac{1}{2}(b_1 + b_2) \\ x_2 & = x_4 + \frac{1}{2}(b_1 - b_2) \end{cases}$$

So dass $Ax = b$ genau dann, wenn

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -x_3 + \frac{1}{2}(b_1 + b_2) \\ x_4 + \frac{1}{2}(b_1 - b_2) \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{2}(b_1 + b_2) \\ \frac{1}{2}(b_1 - b_2) \\ 0 \\ 0 \end{pmatrix}$$

Insbesondere ist

$$\left(\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

eine Basis vom $\operatorname{Ker} L_A$.

4.4.3 Gleichungen vom Erzeugendensystem und Basis von Gleichungen

Sei $V = \langle v_1, \ldots, v_n \rangle \subseteq \mathbb{K}^m$. Wir wöllen V durch Gleichungen beschreiben.

- Sei $A = (v_1 | \ldots | v_n) \in \mathbb{K}^{m \times n}$ die Matrix, die die v_i als Spaltenvektoren hat. Dann $V = \operatorname{Im} L_A$.
- Wir können das Algorithmus in 4.4.2 nutzen um $\operatorname{Im} L_A$ durch Gleichungen zu beschreiben.

Sei $\operatorname{Ker} L_A \subseteq \mathbb{K}^n$ ein Untervektorraum der durch Gleichungen beschrieben ist. Wir wollen eine Basis von $\operatorname{Ker} L_A$ finden.

- Wir können das Algorithmus in 4.4.2 nutzen, um eine Basis von $\operatorname{Ker} L_A$ zu finden.

4.4.4 Dimension und Basis aus einem Erzeugendensystem

Sei $V = \langle v_1, \dots, v_n \rangle \subseteq \mathbb{K}^m$ ein Untervektorräum. Wir wollen die Dimension von V berechnen und eine Basis von V finden. Wir brauchen ein Lemma

Lemma 4.4.4. *1. Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix und sei $A' \in \mathbb{K}^{m\times n}$ eine Matrix, die aus A durch eine endliche Folge elementarer Zeilenumformungen erhalten wird, seien auch $1 \le j_1 < \dots < j_r \le n$. Die Spaltenvektoren $A^{j_1}, \dots, A^{j_r}$ sind linear unabhängig, genau dann wenn, die Spaltenvektoren $A'^{j_1}, \dots, A'^{j_r}$ linear unabhängig sind.*

2. Sei $A \in \mathbb{K}^{m\times n}$ eine Matrix in Zeilenstufenform mit Pivots in den Spalten $1 \le j_1 < \dots < j_r \le n$. Dann ist $(A^{j_1}, \dots, A^{j_r})$ eine Basis vom Spaltenraum von A.

Beweis. 1. Wenn wir uns auf die Untermatrizen von A, A' beschränken, die nur aus den Spalten $j_1 < \dots < j_n$ bestehen, sehen wir, dass sie durch eine Folge von elementaren Zeilenoperationen aus einander erhalten werden. Wir können dann beweisen, dass alle Spalten von A linear unabhängig sind, genau dann wenn, dass alle Spalten von A' linear unabhängig sind. Aber die Spalten von A (bzw. A') sind linear unabhängig, genau dann, wenn $\operatorname{Ker} L_A = 0$ (bzw. $\operatorname{Ker} L'_A = 0$) und $\operatorname{Ker} L_A = \operatorname{Ker} L_{A'}$.

2. Wir wissen dass $\operatorname{rk}(A) = r$, so dass der Spaltenraum Dimension r hat. Wir müssen dann nur zeigen dass $(A^{j_1}, \dots, A^{j_r})$ linear unabhängig ist. Wir können eine Folge von elementaren Zeilenumformungen verwenden, um die Matrix A in eine Matrix A' in reduzierter Zeilenstufenform zu bringen. Dann wird $(A^{j_1}, \dots, A^{j_r})$ zu $(e_1, \dots, e_r)$, das linear unabhängig ist. Wir sind fertig mit Hilfe von Punkt 1.

□

Wir können nun den Algorithmus angeben

- Sei $A = (v_1|v_2|\dots|v_n) \in \mathbb{K}^{m\times n}$ die Matrix, die die v_i als Spalten hat,so dass V der Spaltenraum von A ist.
- Wir bringen A durch elementare Zeilenumformungen in Zeilenstufenform. Die Matrix in Zeilenstufenform hat Pivots in $1 \le j_1 < \dots < j_r \le n$ so dass $\dim V = r$.
- Das vorheriges Lemma zeigt dass $(v_{j_1}, \dots, v_{j_r})$ linear unabhängig ist und dann auch eine Basis.

Beispiel 4.4.5. Sei

$$V = \left\langle \begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}3\\3\\3\end{pmatrix} \right\rangle \subseteq \mathbb{Q}^3$$

Wir wollen die Dimension von V und eine Basis von V berechnen. Wir anwenden das Algorithmus

$$\begin{pmatrix}1&4&3\\2&5&3\\3&6&3\end{pmatrix} \xrightarrow[Z3\to Z3-3Z1]{Z2\to Z2-2Z1} \begin{pmatrix}1&4&3\\0&-3&-3\\0&-6&-6\end{pmatrix} \xrightarrow{Z3\to Z3-2Z2} \begin{pmatrix}1&4&3\\0&-3&-3\\0&0&0\end{pmatrix}$$

Diese Matrix ist in Zeilenstufenform und hat Pivots in der Spalten 1 und 2. Das zeigt dass $\dim V = 2$ und dass

$$\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right)$$

eine Basis von V ist.

4.4.5 Komplement

Sei $V = \langle v_1, \dots, v_n \rangle \subseteq \mathbb{K}^m$ ein Untervektorräum. Wir wollen ein Komplement $W \subseteq \mathbb{K}^m$ finden, so dass $\mathbb{K}^m = V \oplus W$

- Sei $(w_1, \dots, w_m)$ eine Basis von W, zum Beispiel die kanonische Basis.
- Sei $A = (v_1|v_2|\dots|v_n|w_1|\dots|w_m) \in \mathbb{K}^{m\times(n+m)}$ die Matrix die als Spalten die Vektoren $v_1, \dots, v_n$ und die Vektoren der kanonische Basis von $\mathbb{K}^m$ hat.
- Durch elementare Zeilenumformungen bringen wir in Zeilenstufenform, mit Pivots in den Spalten $1 \leq j_1 < \dots < j_r \leq n < n + j_{r+1} < \dots < n + j_{m-r} \leq n + m$
- $(v_{j_1}, \dots, v_{j_r})$ ist eine Basis von V und $(w_{j_{r+1}}, \dots, w_{j_{m-r}})$ ist eine Basis von einem Komplement W von V.

Beispiel 4.4.6. Sei

$$V = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \right\rangle \subseteq \mathbb{Q}^3$$

Wir wollen die Dimension von V und eine Basis von V berechnen und ein Komplement von W Wir anwenden das Algorithmus

$$\begin{pmatrix} 1 & 4 & 3 & | & 1 & 0 & 0 \\ 2 & 5 & 3 & | & 0 & 1 & 0 \\ 3 & 6 & 3 & | & 0 & 0 & 1 \end{pmatrix} \xrightarrow[Z3 \to Z3 - 3Z1]{Z2 \to Z2 - 2Z1} \begin{pmatrix} 1 & 4 & 3 & | & 1 & 0 & 0 \\ 0 & -3 & -3 & | & -2 & 1 & 0 \\ 0 & -6 & -6 & | & -3 & 0 & 1 \end{pmatrix}$$
$$\xrightarrow{Z3 \to Z3 - 2Z2} \begin{pmatrix} 1 & 4 & 3 & | & 1 & 0 & 0 \\ 0 & -3 & -3 & | & -2 & 1 & 0 \\ 0 & 0 & 0 & | & 1 & -2 & 1 \end{pmatrix}$$

Diese Matrix ist in Zeilenstufenform und hat Pivots in der Spalten $1, 2$ und 4. Das zeigt dass $\dim V = 2$,dass

$$\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right)$$

eine Basis von V ist und dass

$$W = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)$$

ein Komplement von V in $\mathbb{Q}^3$ ist.

4.4.6 Die inverse Matrix

Sei $A \in \mathbb{K}^{n\times n}$ eine Matrix. Wir wöllen überprüfen, ob A invertierbar ist, und wenn A invertierbar ist, wollen wir die Inverse finden.

- Wir betrachten die Matrix $(A|I_n)$, wo $I_n \in \mathbb{K}^n$ die Identitätsmatrix ist, und wir verwenden elementare Zeilenumformungen, bis die erste nxn Matrix in Zeilenstufenform ist.
- Wenn die Matrix in Zeilenstufenform weniger als n Pivots hat, dann ist $\mathrm{rk}(A) < n$ und A ist nicht invertierbar.
- Wenn die Matrix in Zeilenstufenform n Pivot hat, dann verwenden wir elementare Zeilenumformungen, bis die erste $n \times n$ Matrix in reduzierte Zeilenstufenform ist: das bedeutet dass die erste $n \times n$ Matrix die Identitätsmatrix ist.
- Wir haben eine Matrix $(I_n|B)$. Die Matrix B ist die inverse Matrix die wir suchen $B = A^{-1}$.

Um zu beweisen, dass B die inverse Matrix von A ist, müssen wir zeigen, dass $L_B = L_A^{-1}$, das bedeutet $BA^i = e_i$ für alle $i \in \{1, \dots, n\}$. Anders gesagt, müssen wir zeigen dass dass $(I_n|B)\left(\begin{smallmatrix} e_i \\ -A^i \end{smallmatrix}\right) = 0$. Da wir nur elementare Zeilenoperationen angewendet haben, ist dies äquivalent zu $(A|I_n)\left(\begin{smallmatrix} e_i \\ -A^i \end{smallmatrix}\right) = 0$ was wahr ist.

Beispiel 4.4.7. Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$. Diese Matrix ist in Zeilenstufenform und hat 3 Pivots, so dass $\mathrm{rk}(A) = 3$ und A ist invertierbar. Wir können die Inverse matrix von A mit dem Algorithmus finden:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[Z2 \to Z2 - Z3]{Z1 \to Z1 - 3Z2} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & -3 \\ 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$
$$\xrightarrow{Z2 \to \frac{1}{2} Z2} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{Z1 \to Z1 - 2Z2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

Das zeigt dass

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

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 11.6.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

4.5 Die Matrix einer linearen Abbildung zwischen zwei Vektorräume

Ab jetzt sind alle unsere Vektorräume endlich-dimensionale. Seien nun V, W $\mathbb{K}$-Vektorräume, mit $\dim V = n, \dim W = m$, und sei $f\colon V \to W$ eine lineare Abbildung. Seien auch $\mathcal{B}_V = (v_1, \dots, v_n)$

eine Basis von V und $\mathcal{B}_W = (w_1, \dots, w_m)$ eine Basis von W. Seien noch $\mathcal{C}_n$ die kanonische Basis von $\mathbb{K}^n$ und $\mathcal{C}_m$ die kanonische Basis von $\mathbb{K}^m$. Wir betrachten die zwei Isomorphismen

$$\begin{aligned}
\Phi^{\mathcal{C}_n}_{\mathcal{B}_V} &: \mathbb{K}^n \to V, && e_i \mapsto v_i && \text{für alle } i \in \{1, \dots, n\} \\
\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} &: W \to \mathbb{K}^m, && w_j \mapsto e_j && \text{für alle } j \in \{1, \dots, m\}
\end{aligned}$$

und die Komposition $\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} \circ f \circ \Phi^{\mathcal{C}_n}_{\mathcal{B}_V} : \mathbb{K}^n \to \mathbb{K}^m$

$$\mathbb{K}^n \xrightarrow{\Phi^{\mathcal{C}_n}_{\mathcal{B}_V}} V \xrightarrow{f} W \xrightarrow{\Phi^{\mathcal{B}_W}_{\mathcal{C}_m}} \mathbb{K}^m$$

Dies ist eine lineare Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$, und Satz 4.2.1 zeigt dass sie durch eine Matrix dargestellt werden kann:

Definition 4.5.1 (Matrix einer lineare Abbildung bezüglich zwei Basen)**.** Die Matrix $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)$ ist die einzige Matrix $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) \in \mathbb{K}^{m \times n}$ so dass

$$\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} \circ f \circ \Phi^{\mathcal{C}_n}_{\mathcal{B}_V} = L_{M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)}$$

Bemerkung 4.5.2. Bemerkung 4.2.2 zeigt wie man die Matrix $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)$ explizit schreiben kann: die Spalten von $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)$ sind die Vektoren $(\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} \circ f \circ \Phi^{\mathcal{C}_n}_{\mathcal{B}_V})(e_i)$, für $i \in \{1, \dots, n\}$. Um das zu berechnen, betrachten wir für jeder $j \in \{1, \dots, n\}$ die Skalaren $a_{1j}, a_{2j}, \dots, a_{mj} \in \mathbb{K}$ so dass

$$f(v_i) = a_{1j} \cdot w_1 + a_{2j} \cdot w_2 + \dots + a_{mj} \cdot w_m$$

Dann

$$\begin{aligned}
(\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} \circ f \circ \Phi^{\mathcal{C}_n}_{\mathcal{B}_V})(e_i) &= (\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} \circ f)(v_i) \\
&= \Phi^{\mathcal{B}_W}_{\mathcal{C}_m}(a_{1j} \cdot w_1 + a_{2j} \cdot w_2 + \dots + a_{mj} \cdot w_m) \\
&= a_{1j} \cdot \Phi^{\mathcal{B}_W}_{\mathcal{C}_m}(w_1) + a_{2j} \cdot \Phi^{\mathcal{B}_W}_{\mathcal{C}_m}(w_2) + \dots + a_{mj} \cdot \Phi^{\mathcal{B}_W}_{\mathcal{C}_m}(w_m) \\
&= a_{1j} \cdot e_1 + a_{2j} \cdot e_2 + \dots + a_{mj} \cdot e_m
\end{aligned}$$

so dass

$$M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Anders gesagt, die Spalten von A sind durch die Koordinaten der Vektoren $f(v_1), f(v_2), \dots, f(v_m)$ bezüglich der Basis $(w_1, \dots, w_m)$ gegeben

Beispiel 4.5.3. Wir betrachten die lineare Abbildung

$$f : \mathbb{R}[x]_{\leq 2} \to \mathbb{R}[x]_{\leq 3}, \qquad g(x) \mapsto x \cdot g(x)$$

(man kann leicht überprüfen dass diese Abbildung linear ist). Sei $\mathcal{B} = (1, x, x^2)$ die Monomialbasis von $\mathbb{R}[x]_{\leq 2}$ und sei $\mathcal{B}' = (1, x, x^2, x^3)$ die Monomial basis von $\mathbb{R}[x]_{\leq 3}$. Um die Matrix von f

bezüglich die Basen $\mathcal{B}, \mathcal{B}$ zu finden, müssen wir die Elemente $f(1), f(x), f(x^2)$ als Linearkombinationen der Elementen in $\mathcal{B}'$ schreiben:

$$\begin{aligned} f(1) &= x \cdot 1 = 0 \cdot x^0 + 1 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ f(x) &= x \cdot x = 0 \cdot x^0 + 0 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 \\ f(x^2) &= x \cdot x^2 = 0 \cdot x^0 + 0 \cdot x + 0 \cdot x^2 + 1 \cdot x^3 \end{aligned}$$

und die Spalten von $M^{\mathcal{B}}_{\mathcal{B}'}(f)$ sind die Koeffizienten in dieser Linearkombinationen:

$$M^{\mathcal{B}}_{\mathcal{B}'}(f) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wir könnten auch aber die Basis $\widetilde{\mathcal{B}} = (x-1, x+1, x^2+2)$ von $\mathbb{R}[x]_{\leq 2}$ betrachten (überprüfen Sie dass $\widetilde{\mathcal{B}}$ eine Basis von $\mathbb{K}[x]_{\leq 2}$ ist), und dann

$$\begin{aligned} f(x-1) &= x(x-1) = x - x^2 = 0 \cdot x^0 + 1 \cdot x - 1 \cdot x^2 + 0 \cdot x^3 \\ f(x+1) &= x(x+1) = x + x^2 = 0 \cdot x^0 + 1 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 \\ f(x^2+2) &= x(x^2+2) = 2x + x^3 = 0 \cdot x^0 + 2 \cdot x + 1 \cdot x^2 + 1 \cdot x^3 \end{aligned}$$

so dass

$$M^{\widetilde{\mathcal{B}}}_{\mathcal{B}'}(f) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Wir könnten auch die Basis $\widetilde{\mathcal{B}'} = (x+1, x^3-x, x^2+2, x-1)$ von $\mathbb{R}[x]_{\leq 3}$ betrachten und dann

$$\begin{aligned} f(1) &= x = \frac{1}{2} \cdot (x+1) + 0 \cdot (x^3 - x) + 0 \cdot (x^2+2) + \frac{1}{2} \cdot (x-1) \\ f(x) &= x^2 = -1 \cdot (x+1) + 0 \cdot (x^3 - x) + 1 \cdot (x^2+2) + 1 \cdot (x-1) \\ f(x^2) &= x^3 = \frac{1}{2} \cdot (x+1) + 1 \cdot (x^3 - x) + 0 \cdot (x^2+2) + \frac{1}{2} \cdot (x-1) \end{aligned}$$

so dass

$$M^{\mathcal{B}}_{\widetilde{\mathcal{B}'}}(f) = \begin{pmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{pmatrix}$$

Als Aufgabe, kann man die Matrix $M^{\widetilde{\mathcal{B}}}_{\widetilde{\mathcal{B}'}}(f)$ berechnen.

Lemma 4.5.4. *Seien $f \colon V \to W$ eine lineare Abbildung, $\mathcal{B}_V$ eine Basis von V, $\mathcal{B}_W$ eine Basis von W und $M = M^{\mathcal{B}_V}_{\mathcal{B}_W}(f)$ die Matrix von f bezüglich der zwei Basen, so dass*

$$\Phi^{\mathcal{B}_W}_{\mathcal{C}_m} \circ f \circ \Phi^{\mathcal{C}_n}_{\mathcal{B}_V} = L_M$$

Dann

1. $\operatorname{Ker} f = \Phi_{\mathcal{B}_V}^{\mathcal{C}_n}(\operatorname{Ker} L_M)$.

2. $\operatorname{Im} f = \Phi_{\mathcal{B}_W}^{\mathcal{C}_m}(\operatorname{Im} L_M)$

3. $\dim \operatorname{Im} f = \operatorname{rk} M$.

Beweis. Die Abbildungen $\Phi_{\mathcal{C}_m}^{\mathcal{B}_W}$ und $\Phi_{\mathcal{B}_V}^{\mathcal{C}_n}$ sind Isomorphismen, und $(\Phi_{\mathcal{C}_m}^{\mathcal{B}_W})^{-1} = \Phi_{\mathcal{B}_W}^{\mathcal{C}_m}$ und $(\Phi_{\mathcal{B}_V}^{\mathcal{C}_n})^{-1} = \Phi_{\mathcal{C}_n}^{\mathcal{B}_V}$ (Warum?), so dass

$$f = \Phi_{\mathcal{B}_W}^{\mathcal{C}_m} \circ L_M \circ \Phi_{\mathcal{C}_n}^{\mathcal{B}_V}$$

Dann

1. $f(x) = 0$ genau dann wenn $\Phi_{\mathcal{B}_W}^{\mathcal{C}_m}((L_M \circ \Phi_{\mathcal{C}_n}^{\mathcal{B}_V})(x)) = 0$. Da $\Phi_{\mathcal{B}_W}^{\mathcal{C}_m}$ injektiv ist, bedeutet das, dass $(L_M \circ \Phi_{\mathcal{C}_n}^{\mathcal{B}_V})(x) = 0$, d.h. $\Phi_{\mathcal{C}_n}^{\mathcal{B}_V}(x) \in \operatorname{Ker} L_M$. Das bedeutet dass $x \in {\Phi_{\mathcal{C}_n}^{\mathcal{B}_V}}^{-1}(\operatorname{Ker} L_M) = \Phi_{\mathcal{B}_V}^{\mathcal{C}_n}(\operatorname{Ker} L_M)$.

2. Da $\Phi_{\mathcal{C}_n}^{\mathcal{B}_V}$ surjektiv ist, sehen wir dass $\operatorname{Im} f = \Phi_{\mathcal{B}_W}^{\mathcal{C}_m}(\operatorname{Im} L_M)$.

3. Die lineare Abbildung $\Phi_{\mathcal{B}_W \mid \operatorname{Im} L_M}^{\mathcal{C}_m} \colon \operatorname{Im} L_M \to \operatorname{Im} f$ ist ein Isomorphismus, so dass $\dim \operatorname{Im} f = \dim \operatorname{Im} L_M = \operatorname{rk} M$.

□

Definition 4.5.5 (Rang einer lineare Abbildung)**.** Sei $f\colon V \to W$ eine lineare Abbildung. Der Rang von f ist $\operatorname{rk}(f) = \dim \operatorname{Im} f$.

Proposition 4.5.6. *Seien $g\colon U \to V$ und $f\colon V \to W$ lineare Abbildungen und seien $\mathcal{B}_U, \mathcal{B}_V, \mathcal{B}_W$ Basen von U, V und W. Dann*

$$M_{\mathcal{B}_W}^{\mathcal{B}_U}(f \circ g) = M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \cdot M_{\mathcal{B}_V}^{\mathcal{B}_U}(g)$$

Beweis. Seien $\dim U = \ell, \dim V = n, \dim W = m$ und seien $\mathcal{C}_\ell, \mathcal{C}_n, \mathcal{C}_m$ die kanonische Basen von $\mathbb{K}^\ell, \mathbb{K}^n, \mathbb{K}^m$. Dann ist $M = M_{\mathcal{B}_W}^{\mathcal{B}_U}(f \circ g)$ die einzige Matrix so dass $\Phi_{\mathcal{C}_m}^{\mathcal{B}_W} \circ (f \circ g) \circ \Phi_{\mathcal{B}_U}^{\mathcal{C}_\ell} = L_M$. Wir rechnen

$$\begin{aligned}
\Phi_{\mathcal{C}_m}^{\mathcal{B}_W} \circ (f \circ g) \circ \Phi_{\mathcal{B}_U}^{\mathcal{C}_\ell} &= \Phi_{\mathcal{C}_m}^{\mathcal{B}_W} \circ f \circ \operatorname{Id}_V \circ g \circ \Phi_{\mathcal{B}_U}^{\mathcal{C}_\ell} = \Phi_{\mathcal{C}_m}^{\mathcal{B}_W} \circ f \circ (\Phi_{\mathcal{B}_V}^{\mathcal{C}_n} \circ \Phi_{\mathcal{C}_n}^{\mathcal{B}_V}) \circ g \circ \Phi_{\mathcal{B}_U}^{\mathcal{C}_\ell} \\
&= (\Phi_{\mathcal{C}_m}^{\mathcal{B}_W} \circ f \circ \Phi_{\mathcal{B}_V}^{\mathcal{C}_n}) \circ (\Phi_{\mathcal{C}_n}^{\mathcal{B}_V} \circ g \circ \Phi_{\mathcal{B}_U}^{\mathcal{C}_\ell}) \\
&= L_{M_{\mathcal{B}_W}^{\mathcal{B}_V}(f)} \cdot L_{M_{\mathcal{B}_V}^{\mathcal{B}_U}(g)} = L_{M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \cdot M_{\mathcal{B}_V}^{\mathcal{B}_U}(g)}
\end{aligned}$$

so dass $M_{\mathcal{B}_W}^{\mathcal{B}_U}(f \circ g) = M_{\mathcal{B}_W}^{\mathcal{B}_V}(f) \cdot M_{\mathcal{B}_V}^{\mathcal{B}_U}(g)$. □

Korollar 4.5.7. *Seien V, W zwei $\mathbb{K}$-Vektorräume und seien $\mathcal{B}, \widetilde{\mathcal{B}}$ zwei Basen von V und $\mathcal{B}', \widetilde{\mathcal{B}}'$ zwei Basen von W.*

1. $M_{\mathcal{B}}^{\mathcal{B}}(\operatorname{id}_V) = I_n$.

2. $M_{\widetilde{\mathcal{B}}}^{\mathcal{B}}(\operatorname{id}_V) \cdot M_{\mathcal{B}}^{\widetilde{\mathcal{B}}}(\operatorname{id}_V) = I_n$, *so dass*

$$M_{\widetilde{\mathcal{B}}}^{\mathcal{B}}(\operatorname{id}_V) = (M_{\mathcal{B}}^{\widetilde{\mathcal{B}}}(\operatorname{id}_V))^{-1}$$

3. *Wenn $f\colon V \to W$ eine lineare Abbildung ist, dann*

$$M^{\widetilde{\mathcal{B}}}_{\widetilde{\mathcal{B}}'}(f) = M^{\mathcal{B}'}_{\widetilde{\mathcal{B}}'}(\mathrm{id}_W) \cdot M^{\mathcal{B}}_{\mathcal{B}'}(f) \cdot M^{\widetilde{\mathcal{B}}}_{\mathcal{B}}(\mathrm{id}_V)$$

Beweis. 1. Einfache Berechnung.

2. Die vorherige Proposition und Punkt (1) zeigen dass $M^{\mathcal{B}}_{\widetilde{\mathcal{B}}}(\mathrm{id}_V) \cdot M^{\widetilde{\mathcal{B}}}_{\mathcal{B}}(\mathrm{id}_V) = M^{\widetilde{\mathcal{B}}}_{\mathcal{B}'}(\mathrm{id}_V) = I_n$.

3. Die vorherige Proposition zeigt dass

$$M^{\mathcal{B}'}_{\widetilde{\mathcal{B}}'}(\mathrm{id}_W) \cdot M^{\mathcal{B}}_{\mathcal{B}'}(f) \cdot M^{\widetilde{\mathcal{B}}}_{\mathcal{B}}(\mathrm{id}_V) = M^{\mathcal{B}'}_{\widetilde{\mathcal{B}}'}(\mathrm{id}_W) \cdot M^{\widetilde{\mathcal{B}}}_{\mathcal{B}'}(f \circ \mathrm{id}_V) = M^{\widetilde{\mathcal{B}}}_{\widetilde{\mathcal{B}}'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) = M^{\widetilde{\mathcal{B}}}_{\widetilde{\mathcal{B}}'}(f)$$

□

Definition 4.5.8 (Ähnliche Matrizen)**.** Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ sind ähnlich, genau dann wenn eine invertierbare Matrix $N \in GL_n(\mathbb{K})$ existiert so dass

$$B = NAN^{-1}$$

Lemma 4.5.9. *Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ sind ähnlich genau dann, wenn ein n-dimensionaler Vektorraum V und eine lineare Abbildung $f\colon V \to V$ existieren, so dass*

$$A = M^{\mathcal{B}}_{\mathcal{B}}(f), \qquad B = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$$

womit $\mathcal{B}, \mathcal{B}'$ zwei Basen von V sind.

Beweis. Wenn A, B ähnlich sind, sei $N \in GL_n(\mathbb{K})$ invertierbar so dass $A = NBN^{-1}$. Wir betrachten den Vektorraum $V = \mathbb{K}^n$ und die lineare Abbildung $f = L_A\colon \mathbb{K}^n \to \mathbb{K}^n$. Sei $\mathcal{B} = \mathcal{C}_n = (e_1, \dots, e_n)$ die kanonische Basis von $\mathbb{K}^n$. Dann $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$. Wir betrachten dann die Basis $\mathcal{B}' = (Ne_1, \dots, Ne_n)$, so dass die Kolumne von N die Vektoren $\mathrm{id}_{\mathbb{K}^n}(Ne_1), \dots, \mathrm{id}_{\mathbb{K}^n}(Ne_1)$ sind. Das bedeutet genau dass $N = M^{\mathcal{B}'}_{\mathcal{C}_n}(\mathrm{id}_{\mathbb{K}^n}) = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V)$ und dann zeigt die vorherige Proposition dass

$$B = NAN^{-1} = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V)^{-1} = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V) = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$$

Andererseits, wenn A, B die Matrizen von der gleichen Abbildung $f\colon V \to V$ bezüglich zwei Basen von V sind, zeigt die vorherige Proposition dass A, B ähnlich sind. □

Bemerkung 4.5.10. Die Ählichkeit von Matrizen ist eine Äquivalenzrelation auf $\mathbb{K}^{n\times n}$: seien $A, B, C \in \mathbb{K}^{n\times n}$. Wir überprüfen die drei Eigenschaften einer Äquivalenzrelation

- Reflexivität: $A = I_n A I_n^{-1}$.
- Symmetrie: wenn $N \in GL_n(\mathbb{K})$ existiert so dass $B = NAN^{-1}$ dann $A = MBM^{-1}$, womit $M = N^{-1}$.
- Transitivität: wenn $N, M \in GL_n(\mathbb{K})$ existieren so dass $A = NBN^{-1}$ und $B = MCM^{-1}$ dann $A = NBN^{-1} = NMCM^{-1}N^{-1} = (NM)C(NM)^{-1}$.

Bemerkung 4.5.11. Seien A, B zwei ähnliche Matrizen. Dann $\mathrm{rk}(A) = \mathrm{rk}(B)$: sei $f\colon V \to V$ eine lineare Abbildung und seien $\mathcal{B}, \mathcal{B}$ zwei Basen von V so dass $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ und $B = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$. Dann $\mathrm{rk}(A) = \mathrm{rk}(f) = \mathrm{rk}(B)$.

Im Allgemein, wenn wir eine Eigenschaft von Matrizen haben, die bei Ähnlichkeit invariant ist, können wir sie als eine Eigenschaft der durch die Matrix dargestellten linearen Abbildung betrachten.

♣♣♣♣♣♣♣♣♣♣♣♣ **Ende Vorlesung 14.6.2024** ♣♣♣♣♣♣♣♣♣♣♣♣

4.6 Die Determinante

Definition 4.6.1 (Determinantenfunktion)**.** Sei $\mathbb{K}$ ein Körper. Eine Funktion

$$\delta\colon \mathbb{K}^{n\times n} \to \mathbb{K}, \qquad A \mapsto \delta(A)$$

ist eine Determinantenfunktion wenn

1. δ linear in jeder Zeile ist:
$$\delta\begin{pmatrix} A_1 \\ \vdots \\ A_j + A'_j \\ \vdots \\ A_n \end{pmatrix} = \delta\begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix} + \delta\begin{pmatrix} A_1 \\ \vdots \\ A'_j \\ \vdots \\ A_n \end{pmatrix}, \qquad \delta\begin{pmatrix} A_1 \\ \vdots \\ \lambda A_j \\ \vdots \\ A_n \end{pmatrix} = \lambda\delta\begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix}$$

2. δ eine alternierende Funktion der Zeilen ist, d.h. $\delta(A) = 0$ falls A zwei identische Zeile hat.

3. $\delta(I_n) = 1$.

Beispiel 4.6.2. Wenn $n = 1$, dann ist

$$\det\colon \mathbb{K}^{1\times 1} \to \mathbb{K}, \qquad (a) \mapsto a$$

eine Determinantenfunktion. Das ist einfach zu sehen.

Wenn $n = 2$, dann ist

$$\det\colon \mathbb{K}^{2\times 2} \to \mathbb{K}, \qquad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto a_{11}a_{22} - a_{12}a_{21}$$

eine Determinantenfunktion. Wir überprüfen die drei Eigenschaften

- Die Funktion det ist linear in jeder Zeile: wir zeigen dass det linear in der ersten Zeile ist
$$\begin{aligned} \det\begin{pmatrix} a_{11}+a'_{11} & a_{12}+a'_{12} \\ a_{21} & a_{22} \end{pmatrix} &= (a_{11}+a'_{11})a_{22} - (a_{12}+a'_{12})a_{21} = a_{11}a_{22} - a_{12}a_{21} + a'_{11}a_{22} - a'_{12}a_{21} \\ &= \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \det\begin{pmatrix} a'_{11} & a'_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ \det\begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \lambda a_{11}a_{22} - \lambda a_{21}a_{12} = \lambda(a_{11}a_{22} - a_{21}a_{12}) \\ &= \lambda\cdot\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \end{aligned}$$
Eine ähnliche Begrundung zeigt dass det auch linear in der zweiten Zeile ist.

- Falls die Zeilen von A gleich sind:
$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{pmatrix} = a_{11}a_{21} - a_{11}a_{21} = 0$$

- Die Einheitsmatrix
$$\det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = 1\cdot 1 - 0\cdot 0 = 1$$

Proposition 4.6.3. *Sei $\delta\colon \mathbb{K}^{n\times n}\to\mathbb{K}^n$ eine Determinantenfunktion. Die elementare Zeilenumformungen ändern δ wie folgt:*

(GZ1) Vertauschen zweier Zeilen
$$\delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j\\ \vdots\\ A_n\end{pmatrix} = -\delta\begin{pmatrix}A_1\\ \vdots\\ A_j\\ \vdots\\ A_i\\ \vdots\\ A_n\end{pmatrix}$$

(GZ2) Addition des λ-faches einer Zeile zu einer anderen Zeile, $\lambda\in\mathbb{K}$.
$$\delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j+\lambda A_i\\ \vdots\\ A_n\end{pmatrix} = \delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j\\ \vdots\\ A_n\end{pmatrix}$$

(GZ3) Multiplikation einer Zeile mit $\lambda\in\mathbb{K}^\times$
$$\delta\begin{pmatrix}A_1\\ \vdots\\ \lambda\cdot A_i\\ \vdots\\ A_n\end{pmatrix} = \lambda\cdot\delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_n\end{pmatrix}$$

Beweis. :

(GZ3) Das gilt weil eine Determinantenfunktion linear in jeder Zeile ist.

(GZ2) Wir berechnen:
$$\delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j+\lambda A_i\\ \vdots\\ A_n\end{pmatrix} = \delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j\\ \vdots\\ A_n\end{pmatrix} + \lambda\delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_i\\ \vdots\\ A_n\end{pmatrix} = \delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j\\ \vdots\\ A_n\end{pmatrix} + \lambda\cdot 0 = \delta\begin{pmatrix}A_1\\ \vdots\\ A_i\\ \vdots\\ A_j\\ \vdots\\ A_n\end{pmatrix}$$

(GZ1) Wir berechnen:

$$0 = \delta \begin{pmatrix} A_1 \\ \vdots \\ A_i + A_j \\ \vdots \\ A_i + A_j \\ \vdots \\ A_n \end{pmatrix} = \delta \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} + \delta \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix} + \delta \begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} + \delta \begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix} = \delta \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix} + \delta \begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix}$$

□

Satz 4.6.4. *Sei $\delta\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Determinantenfunktion.*

1. *δ ist eindeutig: wenn δ' eine andere Determinantenfunktion ist, dann $\delta = \delta'$.*
2. *$\delta(A) = 0$ genau dann, wenn $\operatorname{rk}(A) < n$.*
3. *$A \in GL_n(\mathbb{K})$ genau dann, wenn $\delta(A) \neq 0$.*
4. *$\delta(AB) = \delta(A) \cdot \delta(B)$ für alle $A, B \in \mathbb{K}^{n\times n}$.*
5. *$\delta(A^{-1}) = \delta(A)^{-1}$ für alle $A \in GL_n(\mathbb{K})$.*
6. *wenn A, B zwei ähnliche Matrizen sind, dann $\delta(A) = \delta(B)$.*

Beweis. 1. Sei $A \in \mathbb{K}^{n\times n}$. Wir wollen zeigen dass $\delta(A) = \delta'(A)$ gleich sind. Elementare Zeilenumformungen auf A haben an beide Seiten die gleiche Wirkung, so dass wir annehmen können, dass A in reduzierte Zeilenstufenform ist. Wenn $\operatorname{rk}(A) < n$ dann ist die letzte Zeile von A gleich Null, so dass $\delta(A) = \delta'(A) = 0$, da eine Determinantenfunktion lineare in der letzten Zeile ist. Wenn $\operatorname{rk}(A) = n$, dann ist A die Identitätsmatrix $A = I_n$, so dass $\delta(I_n) = 1 = \delta'(I_n)$.

2. Wir sehen dass alle elementare Zeilenumformungen auf A bewirken dass $\delta(A)$ mit $1, (-1)$ oder $\lambda \in \mathbb{K}^\times$ multipliziert wird. Insbesondere haben sie keinen Einfluss darauf, ob $\delta(A)$ Null ist oder nicht. Dann können wir annehmen dass A reduzierte Zeilenstufenform hat und wir haben bereits im Beweis für den vorherigen Punkt gezeigt dass $\delta(A) = 0$ genau dann, wenn $\operatorname{rk}(A) < n$.

3. Wir wissen dass A invertierbar ist, genau dann wenn $\operatorname{rk}(A) = n$. Dann folgt die Aussage aus Punkt (2).

4. Falls $\operatorname{rk}(B) < n$, dann gilt $\operatorname{rk}(AB) < n$ auch (Warum?), so dass $\delta(AB) = 0$ und $\delta(A) \cdot \delta(B) = \delta(A) \cdot 0 = 0$. Falls $\operatorname{rk}(B) = n$ dann $\delta(B) \neq 0$ und wir können die Funktion

$$\delta'\colon \mathbb{K}^{n\times n} \to \mathbb{K}, \qquad A \mapsto \frac{\delta(AB)}{\delta(B)}$$

definieren. Als Funktion der Zeilen von A sehen wir dass

$$\delta'\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} = \delta\begin{pmatrix} A_1 \cdot B \\ \vdots \\ A_n \cdot B \end{pmatrix} \cdot \delta(B)^{-1}$$

und man kann leicht zeigen, dass δ' linear in jeder Zeile und alternierend ist. Außerdem $\delta'(I_n) = \delta(I_n B)\cdot\delta(B)^{-1} = \delta(B)\cdot\delta(B)^{-1} = 1$. Dass zeigt dass δ' eine Determinantenfunktion ist, und Punkt (1) zeigt dass $\delta = \delta'$. Das bedeutet genau dass $\delta(AB) = \delta(A) \cdot \delta(B)$.

5. Wir sehen dass $\delta(A) \cdot \delta(A^{-1}) = \delta(AA^{-1}) = \delta(I_n) = 1$.
6. Sei $N \in GL_n(\mathbb{K})$ so dass $A = NBN^{-1}$, dann $\delta(A) = \delta(NBN^{-1}) = \delta(N)\delta(B)\delta(N^{-1}) = \delta(N)\delta(B)\delta(N)^{-1} = \delta(B)$.

□

Wir haben gerade gezeigt dass eine Determinantenfunktion viele tolle Eigenschaften hat, aber wir wissen noch nicht ob eine Determinantenfunktion überhaupt existiert! Wir haben gesehen dass eine Determinantenfunktion existiert für 1×1 und 2×2 Matrizen. Wir zeigen nun, wie wir eine Determinantenfunktion induktiv definieren können.

Sei $A \in \mathbb{K}^{n\times n}$ eine Quadratmatrix und seien $i, j \in \{1, \dots, n\}$. Wir bezeichnen mit $A^{[i,j]} \in \mathbb{K}^{n\times n}$ die Matrix, die man durch Entfernen der i-ten Zeile und der j-ten Spalte von A erhält.

Satz 4.6.5 (Entwicklungssatz von Laplace)**.** *Sei $n \geq 2$ und sei $\delta \colon \mathbb{K}^{(n-1)\times(n-1)} \to \mathbb{K}$ eine Determinantenfunktion. Dann existiert auch eine Determinantenfunktion $\delta \colon \mathbb{K}^{n\times n} \to \mathbb{K}$, die man wie folgt berechnen kann: sei $A \in \mathbb{K}^{n\times n}$ mit Einträge $A = (a_{ij})$. Dann*

- *Entwicklung nach der i-ten Zeile: $\delta(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij} \cdot \delta(A^{[i,j]})$.*
- *Entwicklung nach der j-ten Spalte: $\delta(A) = \sum_{i=1}^{n}(-1)^{i+j}a_{ij} \cdot \delta(A^{[i,j]})$.*

Wir werden den Satz später beweisen, aber zunächst wollen wir einige Konsequenzen betrachten. Wir haben schon gesehen dass eine Determinantenfunktion für 1×1 und 2×2 Matrizen existiert. Dann dank dem Entwicklungssatz von Laplace existiert eine Determinantenfunktion für alle $n \times n$ Matrizen. Außerdem, wissen wir dass diese Determinantenfunktion eindeutig ist:

Definition 4.6.6 (Determinante)**.** Die Determinante ist die einzige Determinantenfunktion

$$\det \colon \mathbb{K}^{n\times n} \to \mathbb{K}$$

Bemerkung 4.6.7. Wir können die Determinante eine Matrix mit dem Entwicklungssatz von Laplace berechnen: Wir wählen eine Zeile oder eine Spalte unserer Matrix und berechnen dann die Entwicklung in Bezug auf diese Zeile oder Spalte. Um sich zu merken, welche Vorzeichen zu verwenden sind, ist es nützlich, das folgende Schema im Kopf zu behalten

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Zum Beispiel kann die Determinante der folgenden Matrix über die erste Spalte berechnet werden:

$$\det\begin{pmatrix}1 & 2 & 1\\ 1 & 1 & 3\\ 0 & 4 & 0\end{pmatrix} = 1{\cdot}\det\begin{pmatrix}1 & 3\\ 4 & 0\end{pmatrix} - 1{\cdot}\det\begin{pmatrix}2 & 1\\ 4 & 0\end{pmatrix} + 0{\cdot}\det\begin{pmatrix}2 & 1\\ 1 & 3\end{pmatrix} = 1{\cdot}(1{\cdot}0-3{\cdot}4)-1{\cdot}(2{\cdot}0-1{\cdot}4)+0 = -8$$

oder auch über die dritte Zeile:

$$\det\begin{pmatrix}1 & 2 & 1\\ 1 & 1 & 3\\ 0 & 4 & 0\end{pmatrix} = 0\cdot\det\begin{pmatrix}2 & 1\\ 1 & 3\end{pmatrix} - 4\cdot\det\begin{pmatrix}1 & 3\\ 1 & 1\end{pmatrix} + 0\cdot\det\begin{pmatrix}1 & 2\\ 1 & 1\end{pmatrix} = 0 - 4\cdot(3-1)+0 = -8$$

Insbesondere ist es einfacher, die Determinante zu berechnen, wenn wir eine Zeile oder Spalte mit vielen Nullen wählen, da wir in diesem Fall weniger Determinanten von Untermatrizen berechnen müssen.

Der Entwicklungssatz von Laplace ist eine Möglichkeit um die Determinante zu berechnen. Eine andere Möglichkeit ist durch elementare Zeilenumformungen. Wir brauchen zuerst eine Definition

Definition 4.6.8 (Dreiecksmatrix)**.** Eine Quadratmatrix $A \in \mathbb{K}^{n\times n}$ ist eine obere Dreiecksmatrix falls $A_{ij} = 0$ für alle $i > j$. Die Matrix heißt untere Dreiecksmatrix falls $A_{ij} = 0$ für alle $i < j$.

Bemerkung 4.6.9. Eine Quadratmatrix $A \in \mathbb{K}^{n\times n}$ in Zeilenstufenform ist eine obere Dreiecksmatrix. Außerdem, A hat Rang $\mathrm{rk}(A) = n$ genau dann, wenn alle Einträge in der Hauptdiagonale nicht null sind.

Proposition 4.6.10. *Sei $A \in \mathbb{K}^{n\times n}$ eine obere oder untere Dreiecksmatrix. Dann ist $\delta(A)$ das Produkt von alle Einträge in der Hauptdiagonale:*

$$\delta(A) = \prod_{i=1}^{n} A_{ii}$$

Beweis. Induktion auf n. Wenn $n = 1$ dann ist die Aussage klar. Nehmen wir an, dass $n \geq 2$ und dass die Aussage für $n-1$ wahr ist. Sei A eine obere Dreiecksmatrix, dann $A_{i1} = 0$ für alle $i > 1$ so dass die Entwicklung der Determinante nach der ersten Spalte zeigt dass

$$\det(A) = A_{11}\cdot\det(A^{[1,1]})$$

Wir sehen dass $A^{[1,1]}$ eine obere Dreiecksmatrix ist, und die Induktionsannahme zeigt dass $\det(A^{[1,1]}) = \prod_{i=2}^{n} A_{ii}$. Dies beendet den Beweis für eine obere Dreiecksmatrix. Der Beweis für eine untere Dreiecksmatrix ist ähnlich: man nutzt die Entwicklung nach der ersten Zeile. $\square$

Bemerkung 4.6.11. Wir können die Determinante eine Matrix $A \in \mathbb{K}^{n\times n}$ wie folgt auch berechnen:

- Wir nützen elementare Zeilenumformungen um eine Matrix A' in Zeilenstufenform zu bekommen. Diese Matrix ist auch eine obere Dreiecksmatrix.
- Die Determinante $\det(A')$ ist das Produkt von alle Einträge in der Hauptdiagonal von A'.

- Wir berechnen $\det(A)$ durch Proposition 4.6.3.

Beispiel 4.6.12. Wir berechnen die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 1 & 4 \\ -1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Wir bringen die Matrix A in Zeilenstufenform durch elementare Zeilenumformungen

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 1 & 4 \\ -1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \xrightarrow[Z4 \to Z4 - Z1]{\substack{Z2 \to Z2 - 2 \cdot Z1 \\ Z3 \to Z3 + Z1}} \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & -3 & -2 \\ 0 & 2 & 3 & 5 \\ 0 & 0 & -1 & -2 \end{pmatrix} \xrightarrow{Z2 \leftrightarrow Z3} \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 2 & 3 & 5 \\ 0 & 0 & -3 & -2 \\ 0 & 0 & -1 & -2 \end{pmatrix}$$

$$\xrightarrow{Z4 \to Z4 - \frac{1}{3} Z3} \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 2 & 3 & 5 \\ 0 & 0 & -3 & -2 \\ 0 & 0 & 0 & -\frac{4}{3} \end{pmatrix} = A'$$

Wir sehen dass $\det(A') = 1 \cdot 2 \cdot (-3) \cdot (-\frac{4}{3}) = 8$ und wenn wir uns die elementare Zeilenumformungen ansehen, die wir verwendet haben, berechnen wir, dass $\det(A) = -\det(A') = -8$.

Eine wichtige Eigenschaft der Determinante ist, dass sie bei Transposition unverändert bleibt

Korollar 4.6.13. *Es gilt dass*

$$\det(A) = \det(A^t) \qquad \textit{für alle } A \in \mathbb{K}^{n \times n}.$$

Insbesondere ist die Determinante die einzige Funktion $\mathbb{K}^{n \times n} \to \mathbb{K}$*, die linear in jeder Spalte von* A *ist, alternierend in der Spalten ist, und so dass* $I_n \mapsto 1$*. Außerdem verändern elementare Spaltenumformungen die Determinante auf die gleiche Weise wie elementare Zeilenumformungen*

Beweis. Wir zeigen die Aussage induktiv auf n. Wenn $n = 1$, dann ist die Aussage klar. Nehmen wir an dass $n \geq 2$ und die Aussage wahr für $n - 1$ ist, und beweisen wir sie für n. F
ür eine Matrix $A \in \mathbb{K}^{n \times n}$, sehen wir dass $(A^t)^{[i,j]} = (A^{[j,i]})^t$, so dass $\det((A^t)^{[i,j]}) = \det(A^{[j,i]})$ dank der Induktionsannahme. Wir berechnen $\det(A^t)$ über die erste Zeile

$$\det(A^t) = \sum_{j=1}^{n} (A^t)_{1j} \cdot \det((A^t)^{[1,j]}) = \sum_{j=1}^{n} (A)_{j1} \det A^{[j,1]} = \det(A)$$

womit die letzte Gleichung die Entwicklung von $\det(A)$ nach der ersten Spalte ist. Die anderen Aussagen folgen aus der Tatsache, dass die Spalten von A die Zeilen von A^t sind und dass elementare Spaltenumformungen auf A elementaren Zeilenumformungen auf A^t entsprechen. □

Das Ergebnis für Dreiecksmatrizen kann auf Blockdreiecksmatrizen verallgemeinert werden:

Definition 4.6.14 (Blockdreiecksmatrix)**.** Eine obere Blockdreiecksmatrix ist eine Matrix mit der Form

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

womit A, D Quadratmatrizen sind, und 0 eine Nullmatrix ist. Eine ähnliche Definition kann für untere Blockdreiecksmatrizen gegeben werden.

Satz 4.6.15. *Die Determinante einer obere Blockdreiecksmatrix ist das Produkt der Determinanten der diagonale Blockmatrizen:*

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A) \cdot \det(D)$$

Das Gleiche gilt für untere Blockdreiecksmatrizen.

Beweis. Wir beweisen die Aussage für obere Blockdreiecksmatrizen. Die Aussage für untere Dreiecksmatrizen ergibt sich aus dieser Aussage, indem man die Transponierte nimmt. Seien $A \in \mathbb{K}^{m\times m}, D \in \mathbb{K}^{n\times n}$ Quadratmatrizen und sei $B \in \mathbb{K}^{m\times n}$. Wenn $\mathrm{rk}(A) < m$ dann L_A ist nicht injektiv, so dass $x \in \mathbb{K}^m, x \neq 0$ existiert mit $Ax = 0$. Dann

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} Ax \\ 0 \end{pmatrix} = 0$$

so dass die ganze $(n+m) \times (n+m)$ Matrix auch Rang kleiner als $n+m$ hat. Das zeigt dass

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

Wenn $\mathrm{rk}(A) = n$, dann $\det(A) \neq 0$ und wir können die folgende Funktion betrachten

$$\delta \colon \mathbb{K}^{n\times n} \to \mathbb{K}, \qquad D \mapsto \det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \cdot \det(A)^{-1}$$

Wir können die Aussage das wir zeigen wollen auch als $\delta(D) = \det(D)$ für alle $D \in \mathbb{K}^{n\times n}$ schreiben. Wir können dann zeigen dass δ eine Determinantenfunktion ist. Man kann leicht zeigen dass δ linear in jeder Spalte ist und auch das δ alternierend in den Spalten ist. Wir müssen zeigen, dass $\delta(I_n) = 1$, d.h.

$$\det \begin{pmatrix} A & B \\ 0 & I_n \end{pmatrix} = \det(A).$$

Elementare Zeileinumformungen auf die erste m Zeilen in der Blockmatrix sind auch elementare Zeileumformungen auf die Zeilen von A. Insbesondere, haben sie die gleiche Wirkung an beide Seiten der Gleichung. Wir können dass annehmen dass A in Zeilenstufenform ist, insbesondere ist A auch eine obere Dreiecksmatrix, und $\det(A) = \prod_{i=1}^m A_{ii}$. Wenn A eine obere Dreiecksmatrix ist, dann ist die Blockmatrix auch eine obere Dreieckmatrix, und

$$\det \begin{pmatrix} A & B \\ 0 & I_n \end{pmatrix} = \prod_{i=1}^m A_{ii} \cdot \prod_{i=1}^n 1 = \det(A).$$

□

Wir können die Determinante auch verwenden, um eine Formel für die Inverse einer invertierbaren Matrix zu finden

Definition 4.6.16 (Adjungierte Matrix)**.** Sei $A \in \mathbb{K}^{n\times n}$. Die adjungierte Matrix $\mathrm{adj}(A) \in \mathbb{K}^{n\times n}$ ist die Matrix mit Einträge

$$\mathrm{adj}(A)_{ij} = (-1)^{i+j} \det(A^{[j,i]})$$

Satz 4.6.17. *Sei* $A \in \mathbb{K}^{n\times n}$ *eine Matrix. Dann*

$$A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) \cdot I_n$$

Insbesondere, wenn $\det(A) \neq 0$ *dann ist* A *invertierbar und*

$$A^{-1} = \frac{1}{\det(A)} \cdot \operatorname{adj}(A)$$

Beweis. Seien $i, j \in \{1, \ldots, n\}$. Dann

$$(A \cdot \operatorname{adj}(A))_{ij} = \sum_{h=1}^{n} A_{ih} \cdot (\operatorname{adj}(A))_{hj} = \sum_{h=1}^{n} (-1)^{h+j} A_{ih} \cdot \det A^{[j,h]}$$

Wenn $i = j$ das ist genau die Entwicklung nach die j-te Zeile der Determinante von A , so dass $(A \cdot \operatorname{adj}(A))_{ii} = \det(A)$. Wenn $i \neq j$, dann ist das die Entwicklung nach die j-te Zeile der Determinante der Matrix A', womit alle Zeilen von A' gleich den Zeilen von A sind, ausser die i-te Zeile, die gleich der j-te Zeile von A ist: $A'_h = A'_h$ für alle $h \neq j$ und $A'_j = A_i$. Insbesondere sind zwei Zeilen von A' gleich, so dass $\operatorname{rk}(A') < n$ und $\det(A') = 0$. Das zeigt dass $(A \cdot \operatorname{adj}(A))_{ij} = 0$, wenn $i \neq j$, so dass $A \cdot \operatorname{adj}(A) = \det(A) \cdot I_n$. Eine ähnliche Begrundung, mit der Entwicklung nach Spalten, zeigt dass $\operatorname{adj}(A) \cdot A = \det(A) \cdot I_n$ Die Aussage über die Inverse von A ist dann eine unmittelbare Folge □

Korollar 4.6.18 (Cramersche Regel)**.** *Sei* $A \in GL_n(\mathbb{K})$ *und sei* $b \in \mathbb{K}^n$ *ein Vektor. Dann hat das LGS* $Ax = b$ *die einzige Lösung*

$$x_i = \frac{1}{\det(A)} \cdot \det(A^1|\ldots|A^{i-1}|b|A^{i+1}|\ldots|A^n) \qquad \textit{für alle } i \in \{1, \ldots, n\}$$

Beweis. Der vorheriger Satz zeigt dass $Ax = b$ genau dann, wenn $x = A^{-1}b = \frac{1}{\det(A)} \cdot \operatorname{adj}(A) \cdot b$. Wir sehen dass

$$(\operatorname{adj}(A) \cdot b)_i = \sum_{j=1}^{n} \operatorname{adj}(A)_{ij} \cdot b_j = \sum_{j=1}^{n} (-1)^{i+j} \det A^{[j,i]} \cdot b_j$$

und das ist genau die Entwicklung der Determinante $\det(A^1|\ldots|A^{i-1}|b|A^{i+1}|\ldots|A^n)$ nach der i-te Spalte. □

Bemerkung 4.6.19. Wir können diesen Satz verwenden um die Inverse einer 2×2 Matrix zu berechnen: eine Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2\times 2}$$

ist invertierbar, genau dann wenn $ad - bc \neq 0$. Dann ist die Inverse

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

und die Cramersche Regel ergibt dass

$$\begin{cases} ax_1 + bx_2 & = y_1 \\ cx_1 + dx_2 & = y_2 \end{cases} \iff \begin{cases} y_1 & = \frac{dx_1 - bx_2}{ad-bc} \\ y_2 & = \frac{-cx_1 + ax_2}{ad-bc} \end{cases}$$

Wenn eine Matrix großer als 2×2 ist, ist es einfacher die inverse Matrix durch elementare Zeilenoperationen zu berechnen. Der gleiche Ratschlag gilt für lineare Systeme: es ist normalerweise einfacher, das Gauß-Verfahren zu nutzen als die Cramersche Regel. Es ist trotzdem gut zu wissen dass eine Formel für die Lösung existiert wenn die Koeffizientenmatrix invertierbar ist.

4.6.1 Die Determinante eines Endomorphismus

Wir haben die Determinante einer Quadratmatrix definiert. Wir können auch die Determinante eines Endomorphismus leicht definieren.

Bemerkung 4.6.20. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorräume und sei $f\colon V \to V$ ein Endomorphismus. Seien auch $\mathcal{B}, \mathcal{B}'$ zwei Basen von V. Dann sind die Matrizen $M^{\mathcal{B}}_{\mathcal{B}}(f), M^{\mathcal{B}'}_{\mathcal{B}'}(f)$ ähnlich, so dass

$$\det M^{\mathcal{B}}_{\mathcal{B}}(f) = \det M^{\mathcal{B}'}_{\mathcal{B}'}(f).$$

Das bringt zu einer Definition

Definition 4.6.21. Sei V ein endlich-dimensionaler Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Die Determinante von f ist

$$\det(f) := \det M^{\mathcal{B}}_{\mathcal{B}}(f)$$

womit $\mathcal{B}$ eine Basis von V ist.

Dann haben wir

Korollar 4.6.22. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Die folgende Aussage sind äquivalent:*

1. *f ist injektiv.*
2. *f ist surjektiv.*
3. *f ist ein Isomorphismus.*
4. $\operatorname{Ker}(f) = \{0\}$.
5. $\operatorname{rk}(f) = \dim V$.
6. $\det(f) \neq 0$.

4.6.2 Beweis des Entwicklungssatzes von Laplace

Wir nehmen an, dass $n \geq 2$ und dass eine Determinantenfunktion $\delta\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ existiert.

Lemma 4.6.23. *Sei $n \geq 1$ und seien $i, j \in \{1, \ldots, n\}$. Die Funktion*

$$\mathbb{K}^{n\times n} \to \mathbb{K}, \qquad A \mapsto a_{ij} \cdot \delta(A^{[i,j]})$$

ist linear in jeder Zeile.

Beweis. Wir betrachten zuerst die Funktion $\mathbb{K}^{n\times n} \to \mathbb{K}, A \mapsto \delta(A^{[i,j]})$. Wenn $h \neq i$ und die h-te Zeile von A eine Linearkombination von Zeilenvektoren ist, dann gilt das auch für die Matrix $A^{[i,j]}$. Da δ linear in jeder Zeile ist, sehen wir dass die Funktion $A \mapsto \delta(A^{[i,j]})$ linear in der h-Zeile ist, falls $h \neq i$. Wenn $h = i$, dann ist die Funktion $A \mapsto \delta(A^{[i,j]})$ konstant in der i-te Zeile. Dagegen, die Funktion $\mathbb{K}^{n\times n} \to \mathbb{K}, A \mapsto a_{ij}$ ist linear in den i-te Zeile und konstant in alle andere Zeilen. Man kann dann leich ubeprüfen dass das Produkt der zwei Funktionen $A \mapsto a_{ij} \cdot \delta(A^{[i,j]})$ linear in jeder Zeile ist. $\square$

Entwicklung nach Spalten

Wir betrachten die Entwicklung nach der ersten Spalte:

$$\delta^1 \colon \mathbb{K}^{n\times n} \to \mathbb{K}, \qquad \delta^1(A) := \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \delta(A^{[i,1]})$$

Wir wollen zeigen, dass δ^1 eine Determinantenfunktion ist. Das vorheriges Lemma zeigt dass alle Funktionen $A \mapsto a_{i1} \cdot \delta(A^{[i,1]})$ linear in jeder Spalte von A sind. Da δ^1 eine Linearkombination dieser Funktionen ist, ist δ^1 auch linear in jeder Zeile.

Wir zeigen nun dass δ^1 alternierend in der Zeilen ist. Sei A eine Matrix mit zwei identische Zeilen $A_i = A_j$ für $i < j$. Wenn $h \neq i, h \neq j$, dann hat die Matrix $A^{[h,1]}$ auch zwei identische Zeilen, so das $\delta(A^{[h,1]}) = 0$. Außerdem, kann man sehen dass $j - i + 1$ Vertauschen von Zeilen die Matrix $A^{[j,1]}$ in der Matrix $A^{[j,1]}$ bringen, so dass $\delta(A^{[j,1]}) = (-1)^{j-i+1}\delta(A^{[i,1]})$. Dann

$$\begin{aligned}\delta^1(A) &= (-1)^{i+1} a_{i1} \cdot \delta(A^{[i,1]}) + (-1)^{j+1} a_{j1} \cdot \delta(A^{[j,1]}) = (-1)^{i+1} a_{i1} \cdot \delta(A^{[i,1]}) + (-1)^{2j+i} a_{i1} \cdot \delta(A^{[i,1]}) \\ &= a_{i1} \cdot \delta(A^{[i,1]}) \cdot ((-1)^{i+1} + (-1)^{i}) = a_{i1} \cdot \delta(A^{[i,1]}) \cdot 0 = 0.\end{aligned}$$

Das zeigt dass δ^1 eine alternierende Funktion der Zeilen ist. Endlich müssen wir beweisen dass $\delta^1(I_n) = 1$, aber man sieht leicht dass $\delta^1(I_n) = 1 \cdot \delta(I_{n-1}) = 1 \cdot 1 = 1$. Insgesamt, haben wir gezeigt dass die Entwicklung nach der ersten Spalte eine Determinantenfunktion ist, und eine ähnliche Begrundung zeigt dass die Entwicklung nach der j-ten Spalte eine Determinantenfunktion ist.

Entwicklung nach Zeilen

Wir betrachten die Entwicklung nach der ersten Zeile

$$\delta_1 \colon \mathbb{K}^{n\times n} \to \mathbb{K}, \qquad \delta_1(A) := \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \delta(A^{[1,j]})$$

Wir zeigen dass δ_1 eine lineare Abbildung in jeder Zeile ist. Das vorheriges Lemma zeigt dass die Funktionen $A \mapsto a_{1j} \cdot \delta(A^{[1,j]})$ linear in jeder Zeile sind, und da δ_1 eine Linearkombination dieser Funktionen ist, ist δ_1 auch linear in jeder Zeile. wir müssen nun zeigen dass δ_1 alternierend ist: nehmen wir an, dass A zwei identische Zeilen $A_i = A_h$, mit $i, h \geq 2$. Dann haben alle die Matrizen $A^{[1,j]}$ auch zwei identische Zeilen, so dass $\delta(A^{[1,j]}) = 0$ und $\delta_1(A) = 0$ auch. Nehmen wir nun an, dass $A_1 = A_i$ für ein $i \geq 2$, zum Beispiel $A_1 = A_2$. Dann können wir durch Induktion auf n annehmen dass $\delta(A^{[1,j]})$ durch die Laplace-Entwicklung nach die erste Zeile berechnet werden kann:

$$\delta(A^{[1,j]}) = \sum_{1\leq h<j\leq n} (-1)^{1+h} a_{2,h} \delta(A^{[1,j],[2,h]}) - \sum_{n\geq h>j\geq 1} (-1)^{i+h} a_{2,h} \delta(A^{[1,j],[2,h]})$$

womit $A^{[1,j],[2,h]}$ die Matrix ist, die wir bekommen wenn wir die erste zwei Zeilen von A und die j-de und h-te Spalten von A weglassen. Insbesondere $A^{[1,j],[2,h]} = A^{[1,h],[2,j]}$ Da die erste zwei

Zeilen von A gleich sind, sehen wir dass $a_{2,h} = a_{1,h}$, so dass

$$\begin{aligned}
\delta_1(A) &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \delta(A^{[1,j]}) \\
&= \sum_{1 \le h < j \le n} (-1)^{1+j+h} a_{1j} a_{1h} \delta(A^{[1,j],[2,h]}) - \sum_{n \ge h > j \ge 1} (-1)^{1+j+h} a_{1j} a_{1h} \delta(A^{[1,j],[2,h]}) \\
&= \sum_{1 \le h < j \le n} \left((-1)^{1+j+h} a_{1j} a_{1h} \delta(A^{[1,j],[2,h]}) - (-1)^{1+h+j} a_{1h} a_{1j} \delta(A^{[1,h],[2,j]}) \right) = 0.
\end{aligned}$$

Eine ähnliche Begrundung zeigt dass $\delta_1(A) = 0$ falls die erste Zeile von A und die i-te Zeile mit $i \ge 2$ gleich sind. Am Ende, sehen wir dass $\delta_1(I_n) = 1 \cdot \delta(I_{n-1}) = 1 \cdot 1 = 1$. Das zeigt dass die Entwicklung nach der ersten Zeile eine Determinantenfunktion ist, und eine ähnliche Begrundung zeigt dass die Entwicklung nach der i-te Zeile auch eine Determinantenfunktion ist.

Kapitel 5

Eigenwerte und Eigenvektoren

Alle Vektorräume hier haben endliche Dimension.

5.1 Diagonalisierbare Matrizen und Endomorphismen

Die beste Matrizen sind Diagonalmatrizen:

Definition 5.1.1 (Diagonalmatrix)**.** Sei $\mathbb{K}$ ein Körper. Eine Diagonalmatrix ist eine Quadratmatrix $D \in \mathbb{K}^{n\times n}$ deren Einträge außer die Hauptdiagonale Null sind

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Wir können zum Beispiel sehr einfach den Rang oder die Determinante einer Diagonalmatrix berechnen. Wir können auch sehr einfach die Potenzen einer Diagonalmatrix berechnen:

$$D^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix} \qquad \text{für alle } k \geq 0$$

und auch die Inverse (falls $\lambda_i \neq 0$ für alle i):

$$D^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & \dots & 0 \\ 0 & \lambda_2^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^{-1} \end{pmatrix}$$

Bemerkung 5.1.2. Was bedeutet dass eine Matrix diagonal ist? Eine Matrix A ist diagonal genau dann, wenn die i-te Spalte A^i ein λ_i-faches des kanonisches Basisvektors e_i ist. Wir wissen aber dass $A^i = Ae_i$, so dass A diagonal ist, genau dann, wenn

$$Ae_i = L_A(e_i) = \lambda_i \cdot e_i, \qquad \text{für } \lambda_i \in \mathbb{K}$$

Dies fuhrt zu eine wichtige Definition

Definition 5.1.3 (Eigenvektoren und Eigenwerte)**.** Sei $A \in \mathbb{K}^{n\times n}$ eine Matrix. Ein Eigenvektor von A ist ein Vektor $v \in \mathbb{K}^n, v \neq 0$ so dass ein $\lambda \in \mathbb{K}$, genannt Eigenwert, existiert, mit

$$A \cdot v = \lambda \cdot v$$

Sei V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Ein Eigenvektor von f ist ein Vektor $v \in V, v \neq 0$ so dass ein $\lambda \in \mathbb{K}$, genannt Eigenwert, existiert, mit

$$f(v) = \lambda \cdot v$$

Wir sagen dass $\lambda \in \mathbb{K}$ ein Eigenwert von der Matrix A oder von dem Endomorphismus f ist, falls ein Eigenvektor mit Eigenwert λ existiert.

Bemerkung 5.1.4. Die Eigenvektoren und Eigenwerte einer Matrix $A \in \mathbb{K}^{n\times n}$ sind genau die Eigenvektoren und Eigenwerte des Endomorphismus $L_A\colon \mathbb{K}^n \to \mathbb{K}^n$.

Bemerkung 5.1.5. In der Definition von Eigenvektoren, ist wichtig dass der Vektor nicht Null ist, da $f(0) = \lambda \cdot 0$ gilt für alle $\lambda \in \mathbb{K}$.

Definition 5.1.6 (Diagonalisierbare Abbildung)**.** Sei V ein $\mathbb{K}$-Vektorraum. Ein Endomorphismus $f\colon V \to V$ ist diagonalisierbar wenn V eine Basis von Eigenvektoren von f.

Bemerkung 5.1.7. Sei $f\colon V \to V$ diagonalisierbar und sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V wo alle die v_i Eigenvektoren von V sind. Sei λ_i der Eigenwert von v_i so dass $f(v_i) = \lambda_i \cdot v_i$. Dann sehen wir dass die Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ diagonal ist:

$$M^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Andererseits, wenn $\mathcal{B}$ eine Basis von V ist, so dass die Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ diagonal ist, dann ist $\mathcal{B}$ eine Basis von Eigenvektoren von f und f ist dann diagonalisierbar. Das zeigt dass f diagonalisierbar ist, genau dann, wenn eine Basis $\mathcal{B}$ von V existiert so dass $D = M^{\mathcal{B}}_{\mathcal{B}}(f)$ diagonal ist.

In diesem Fall, sei $\mathcal{B}'$ eine andere beliebige Basis, so dass die Matrix $A = M^{\mathcal{B}'}_{\mathcal{B}'}(f)$ nicht unbedingt diagonal ist. Dann

$$D = M^{\mathcal{B}}_{\mathcal{B}}(f) = M^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_V) \cdot M^{\mathcal{B}'}_{\mathcal{B}'}(f) \cdot M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V) = N \cdot A \cdot N^{-1}$$

Das führt zu eine andere Definition:

Definition 5.1.8 (Diagonalisierbare Matrix)**.** Eine Matrix $A \in \mathbb{K}^{n\times n}$ ist diagonalisierbar wenn A ähnlich zu einer diagonalen Matrix ist.

Bemerkung 5.1.9. Die vorherige Diskussion zeigt dass $A \in \mathbb{K}^{n\times n}$ ist diagonalisierbar genau dann, wenn die entsprechende lineare Abbildung $L_A\colon \mathbb{K}^n \to \mathbb{K}^n$ diagonalisierbar ist.

Diagonalisierbare lineare Abbildungen oder Matrizen sind prinzipiell so gut wie diagonale Matrizen. Leider sind nicht alle Matrizen diagonalisierbar:

Beispiel 5.1.10. Zum Beispiel, ist die Matrix

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

nicht diagonalisierbar: nehmen wir an dass $N \in GL_2(\mathbb{K})$ existiert so dass $NAN^{-1} = D$ diagonal ist. Dann $D^2 = NA^2N^{-1} = N \cdot 0 \cdot N^{-1} = 0$, so dass $D = 0$ auch. Aber dann $A = N^{-1}DN = 0$ auch, Widerspruch.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 21.6.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

5.1.1 Eigenräume und das Charakteristisches Polynom

Seien V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Die erste einfache aber nützliche Bemerkung ist das, für jedes $v \in V$ und $\lambda \in \mathbb{K}$

$$f(v) = \lambda \cdot v \iff (f - \lambda \cdot \mathrm{id}_V)(v) = 0$$

Lemma 5.1.11. *Seien V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, und $f\colon V \to V$ ein Endomorphismus.*

1. *Sei $v \in V, v \neq 0$.*

$$v \text{ ist ein Eigenvektor von } f \text{ mit Eigenwert } \lambda \iff v \in \mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V)$$

.

2. *Sei $\lambda \in \mathbb{K}$.*

$$\lambda \text{ ist ein Eigenwert von } f \iff \det(f - \lambda \cdot \mathrm{id}_V) = 0.$$

Beweis. 1. Das folgt aus unseren einfachen Bemerkung.

2. λ ist ein Eigenwert genau dann, wenn eine Eigenvektor mit Eigenwert λ existiert, was bedeuted $\mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V)$.

□

Definition 5.1.12 (Eigenraum)**.** Sei $f\colon V \to V$ ein Endomorphismus eines $\mathbb{K}$-Vektorraums V und sei $\lambda \in \mathbb{K}$. Der Eigenraum von λ ist

$$\mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V)$$

Definition 5.1.13 (Charakteristisches Polynom)**.** Sei $A \in \mathbb{K}^{n\times n}$ eine Matrix und sei t eine Variabel. Das Charakteristisches Polynom von A ist

$$\chi_A(t) = \det(A - tI) \in \mathbb{K}[t]$$

Seien V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Das charakteristisches Polynom von f ist

$$\chi_f(t) = \chi_{M^{\mathcal{B}}_{\mathcal{B}}(f)}(t)$$

wo $\mathcal{B}$ eine Basis von V ist.

Bemerkung 5.1.14. Wenn A, B zwei ähnliche Matrizen sind, dann $A = NBN^{-1}$ für eine $N \in GL_n(\mathbb{K})$. Dann

$$\begin{aligned}\chi_A(t) &= \det(tI - A) = \det(tNIN^{-1} - NBN^{-1}) \\ &= \det(N(tI - A)N^{-1}) = \det(N)\det(tI - B)\det(N)^{-1} = \chi_B(t)\end{aligned}$$

Insbesondere ist das Polynom $\chi_{M^{\mathcal{B}}_{\mathcal{B}}(f)}$ unabhängig von der Basis $\mathcal{B}$. Das zeigt dass das charakteristisches Polynom eines Endomorphismus wohldefiniert ist.

Lemma 5.1.15. *Sei V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Dann $\lambda \in \mathbb{K}$ ist ein Eigenwert von f genau dann, wenn λ eine Nullstelle des charakteristisches Polynom ist: $\chi_f(\lambda) = 0$.*

Beweis. Wir wissen dass $\lambda \in \mathbb{K}$ ein Eigenwert ist, genau dann wenn $\det(f - \lambda I) = 0$, und das bedeutet genau dass $\chi_f(\lambda) = 0$. $\square$

Beispiel 5.1.16. Wir betrachten die Matrix $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$ und das Entsprechendes Endomorphismus $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$. Das charakteristisches Polynom ist

$$\chi_A(t) = \det(tI - A) = \det\begin{pmatrix} 1-t & 1 \\ 4 & 1-t \end{pmatrix} = (t-1)^2 - 4 = t^2 + 1 - 2t - 4 = t^2 - 2t - 3$$

Wir sehen dass dieses Polynom zwei Nullstellen hat: $t^2 - 2t - 3 = (t+1)(t-3)$ so dass die Matrix A zwei Eigenwerte in $\mathbb{R}$ hat: -1 und 3. Wir müssen dann mindestens ein Eigenvektor mit Eigenwert -1 und ein Eigenvektor mit Eigenwert 3 haben. Wir können Basen der entsprechenden Eigenräume berechnen:

$$\operatorname{Ker}\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\rangle$$
$$\operatorname{Ker}\begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle$$

Dann ist $\left(\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right)$ ein Eigenvektor mit Eigenwert -1 und $\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$ ist ein Eigenvektor mit Eigenwert 3. Da diese zwei Vektoren offensichtlich linear unabhängig sind, ist $\mathcal{B} = (\left(\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right))$eine Basis von $\mathbb{R}^2$. Das bedeutet insbesondere dass die Matrix A und das Endomorphismus L_A diagonalisierbar sind. Wir können die Diagonalisierung explizit durchfuhren: sei $\mathcal{C}$ die kanonische Basis von $\mathbb{R}^2$, dann

$$M^{\mathcal{C}}_{\mathcal{C}}(L_A) = A, \qquad M^{\mathcal{B}}_{\mathcal{B}}(L_A) = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} = D$$

so dass

$$D = M^{\mathcal{B}}_{\mathcal{B}}(L_A) = M^{\mathcal{C}}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}) M^{\mathcal{C}}_{\mathcal{C}}(L_A) M^{\mathcal{B}}_{\mathcal{C}}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}$$

Lemma 5.1.17. *Sei V ein $\mathbb{K}$-Vektorraum mit $\dim V = n$ und sei $f\colon V \to V$ ein Endomorphismus. Das charakteristisches Polynom von f hat Grad n. Insbesondere hat f maximal n verschiedene Eigenwerte in $\mathbb{K}$.*

Beweis. Wir zeigen die Aussage für das charakteristisches Polynom einer Matrix $A \in \mathbb{K}^{n\times n}$. Der Beweis ist induktiv auf n: wenn $n = 1$ dann ist die Aussage klar. Wenn $n > 1$, nehmen wir an dass die Aussage wahr für $(n-1) \times (n-1)$-Matrizen ist. Dann

$$\chi_A(t) = \det(A - tI) = \det \begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{pmatrix}$$

Die Laplace-Entwicklung nach der ersten Zeile zeigt dass

$$\begin{aligned} \chi_A(t) &= (a_{11} - t) \cdot \det(A - tI)^{[1,1]} + \sum_{j=2}^{n} (-1)^{1+j} a_{1j} \cdot \det(A - tI)^{[1,j]} \\ &= (a_{11} - t) \cdot \det \begin{pmatrix} a_{22} - t & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} - t \end{pmatrix} + \sum_{j=2}^{n} (-1)^{1+j} a_{1j} \cdot \det(A - tI)^{[1,j]} \end{aligned}$$

Wir sehen dass $\det(A - tI)^{[1,1]}$ das charakteristisches Polynom der Matrix $A^{[1,1]}$, ist, so dass die Induktionsannahme zeigt das $\det(A - tI)^{[1,1]}$ Grad $n - 1$ hat, so dass $(a_{11} - t) \cdot \det(A - tI)^{[1,1]}$ Grad n hat. Wir betrachten den die Summanden $\det(A - tI)^{[1,j]}$ für $j \geq 2$: in der Einträge dieser Matrix gibt es $n - 2$ Polynome von Grad 1 und alle andere Einträge sind Skalaren. Die Laplace Entwicklung zeigt dass die Determinante eine Summe von Produkten von verschiedene Einträge ist, und so ein Produkt kann ein Polynom von Grad höchstens $n - 2$ sein. Das zeigt dass $\chi_A(t)$ Grad n hat.Da alle Eigenwerte $\lambda \in \mathbb{K}$ von A Nullstellen von $\chi_A(t)$ sind, sehen wir dass A kann höchstens n verschiedene Eigenwerte in $\mathbb{K}$ haben. □

Proposition 5.1.18. *Seien V ein $\mathbb{K}$-Vektorraum, $f : V \to V$ ein Endomorphismus und $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ die paarweise verschiedene Eigenwerte von f. Dann sind die entsprechende Eigenräume in direkte Summe:*

$$\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V) \oplus \dots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_V) \subseteq V$$

Insbesondere, nehmen wir für jedes $i \in \{1, \dots, r\}$ eine Basis $\mathcal{B}_i = (v_{i,1}, \dots, v_{i,n_i})$ von $\operatorname{Ker}(A - \lambda_i I)$. *Dann ist*

$$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_r = (v_{1,1}, \dots, v_{r,n_r})$$

eine Basis

Beweis. Der Beweis ist induktiv auf r. Wenn $r = 1$ es gibt nichts zu zeigen. Wenn $r > 1$, nehmen wir an dass $\operatorname{Ker}(A - \lambda_i \operatorname{id}_V)$ für $i \in \{1, \dots, r - 1\}$ in direkte Summe sind. Dann müssen wir zeigen dass

$$(\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)) \oplus \dots \oplus \operatorname{Ker}(f - \lambda_{r-1} \operatorname{id}_V)) \cap \operatorname{Ker}(f - \lambda_r \operatorname{id}_V) = \{0\}$$

Sei v in diesem Schnitt. Da $v \in (\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)) \oplus \dots \oplus \operatorname{Ker}(f - \lambda_{r-1} \operatorname{id}_V))$ existieren eindeutige $v_i \in \operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ für $i \in \{1, \dots, r - 1\}$ so dass

$$v = v_1 + \dots + v_{n-1}$$

Da $v \in \operatorname{Ker}(f - \lambda_r \operatorname{id}_V)$, wissen wir dass $f(v) = \lambda_r v$, so dass

$$\begin{aligned} f(v) &= f(v_1) + \cdots + f(v_{r-1}) \\ \iff \lambda_r v &= \lambda_1 v_1 + \ldots \lambda_{r-1} v_{r-1} \end{aligned}$$

Wir wissen aber auch dass

$$\lambda_r v = \lambda_r v_1 + \cdots + \lambda_r v_{r-1}$$

so dass

$$\lambda_1 v_1 + \ldots \lambda_{r-1} v_{r-1} = \lambda_r v_1 + \cdots + \lambda_r v_{r-1}$$

. Da die $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ für $i \in \{1, \ldots, r-1\}$ in direkte Summe sind, sehen wir dass $\lambda_i v_i = \lambda_r v_i$ oder, äquivalent, dass $(\lambda_i - \lambda_r) v_i = 0$ für alle $i \in \{1, \ldots, r-1\}$. Da $\lambda_i - \lambda_r \neq 0$,es muss sein dass $v_i = 0$ für alle i. Das zeigt dass $v = 0$ auch. □

Satz 5.1.19. *Seien V ein $\mathbb{K}$-Vektorraum, $f : V \to V$ ein Endomorphismus und $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ die paarweise verschiedene Eigenwerte von f. Dann*

$$\sum_{i=1}^{r} \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \leq \dim V$$

und f is diagonalisierbar, genau dann, wenn die Gleichung gilt. In diesem Fall, sei $\mathcal{B}_i = (v_{i,1}, \ldots, v_{i,n_i})$ eine Basis von $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ für jedes $i \in \{1, \ldots, r\}$. Dann ist $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ eine Basis von Eigenvektoren von f.

Beweis. Seien zuerst $\mathcal{B}_i = (v_{i,1}, \ldots, v_{i,n_i})$ Basen von $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ für jedes i. Da diese Untervektorräume in direkte Summe sind, ist $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ eine Basis vom Untervektorraum $\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_V) \subseteq V$, so dass

$$\sum_{i=1}^{r} \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) = \dim \left(\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_V)\right) \leq \dim V$$

Wenn die Gleichung gilt, dann $(\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_V)) = V$ so dass $\mathcal{B}$ eine Basis von V ist. Da alle Vektoren in $\mathcal{B}$ Eigenvektoren von f sind, ist f diagonalisierbar. Andererseits, wenn f diagonalisierbar ist, sei $\mathcal{B}'$ eine Basis von Eigenvektoren von f. Wir betrachten für jedes $i \in \{1, \ldots, r\}$ die Subfamilie $\mathcal{B}'_i \subseteq \mathcal{B}'$ von Eigenvektoren in $\mathcal{B}$ mit Eigenwert λ_i. Die Vektoren in $\mathcal{B}'_i$ sind linear unabhängige Vektoren in $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$, so dass $|\mathcal{B}'_i| \leq \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$. Dann

$$\dim V = |\mathcal{B}'| = \sum_{i=1}^{r} |\mathcal{B}'_i| \leq \sum_{i=1}^{r} \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \leq \dim V$$

so dass $|\mathcal{B}'_i| = \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ für alle i und $\sum_{i=1}^{r} \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) = \dim V$. □

5.1.2 Algebraische und geometrische Vielfachheit

Lemma 5.1.20. *Sei $g(t) \in \mathbb{K}[t], g(t) \neq 0$ ein Polynom und sei $\lambda \in \mathbb{K}$ ein Skalar. Es gibt ein einziges $m \in \mathbb{N}$ gibt so dass*

$$g(t) = (t - \lambda)^m \cdot h(t)$$

mit $h(\lambda) \neq 0$.

Beweis. Wir zeigen das induktiv auf den Grad von g. Wenn $g(\lambda) \neq 0$, dann $m = 0$ und $h(t) = g(t)$. Dieses deckt auch den Fall $\deg(g) = 0$. Wenn $\deg(g) > 0$ und $g(\lambda) = 0$, dann $g(t) = (t-\lambda) \cdot g_1(t)$ für $g_1(t) \in \mathbb{K}[t]$. Wir sehen dass $\deg(g_1) = \deg(g) - 1$ und die Induktionsannahme zeigt dass $g_1(t) = (t-\lambda)^{m_1} \cdot h(t)$ für ein $m_1 \in \mathbb{N}$ und ein $h(t) \in \mathbb{K}[t]$ mit $h(\lambda) \neq 0$. Dann $g(t) = (t-\lambda) \cdot g_1(t) = (t-\lambda)^{m_1+1} \cdot h(t)$. Das zeigt dass m existiert. Wir müssen zeigen dass m eindeutig ist. Seien $m, m' \in \mathbb{N}$ und $h(t), h'(t) \in \mathbb{K}[t]$ mit $h(\lambda) \neq 0, h'(\lambda) \neq 0$ so dass

$$g(t) = (t-\lambda)^m \cdot h(t) = (t-\lambda)^{m'} \cdot h'(t)$$

Nehmen wir an dass $m > m'$, dann

$$0 = (t-\lambda)^m \cdot h(t) - (t-\lambda)^{m'} \cdot h'(t) = (t-\lambda)^{m'} \cdot ((t-\lambda)^{m-m'} \cdot h(t) - h'(t))$$

und dann $(t-\lambda)^{m-m'} \cdot h(t) - h'(t) = 0$. Aber das bedeutet $h'(t) = (t-\lambda)^{m-m'} \cdot h(t)$ so dass $h'(\lambda) = 0$, Widerspruch. Das zeigt dass $m = m'$. □

Definition 5.1.21 (Vielfachheit einer Nullstelle)**.** Sei $g(t) \in \mathbb{K}[t]$ ein Polynom und sei $\lambda \in \mathbb{K}$. Die Vielfachheit von λ als Nullstelle von $g(t)$ ist das einziges $m \in \mathbb{N}$ so dass

$$g(t) = (t-\lambda)^m \cdot h(t)$$

mit $h(\lambda) \neq 0$. Insbesondere ist λ eine Nullstelle von f genau dann, wenn die Vielfachheit positiv ist.

Bemerkung 5.1.22. Sei $g(t) \in \mathbb{K}[t], g(t) \neq 0$ und seien $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$. Für jedes $i \in \{1, \ldots, r\}$, sei m_i die Vielfachheit von λ_i in $g(t)$. Dann $g(t) = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r} \cdot h(t)$, mit $h(\lambda_i) = 0$ für alle i. Man kann das leicht induktiv auf r zeigen.

Definition 5.1.23 (Algebraische und geometrische Vielfachheit eines Eigenwerts)**.** Seien V ein $\mathbb{K}$-Vektorraum, $f \colon V \to V$ ein Endomorphismus und $\lambda \in \mathbb{K}$ ein Eigenwert von f. Die algebraische Vielfachheit von λ ist die Vielfachheit von λ als Nullstelle von $\chi_f(t)$. Die geometrische Vielfachheit von λ ist die Dimension $\dim \operatorname{Ker}(f - \lambda \cdot \mathrm{id}_V)$.

Lemma 5.1.24. *Die geometrische Vielfachheit ist immer kleiner oder gleich als die algebraische Vielfachheit.*

Beweis. Sei $\mathcal{B}_\lambda = (v_1, \ldots, v_e)$ eine Basis von $\operatorname{Ker}(f - \lambda \,\mathrm{id}_V)$ so dass e die geometrische Vielfachheit ist. Wir können $\mathcal{B}_\lambda$ zu einer Basis $\mathcal{B} = (v_1, \ldots, v_e, v_{e+1}, \ldots, v_n)$ ergänzen und dann ist die Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ eine obere Blockdreiecksmatrix:

$$A = M^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda \cdot I_e & B \\ 0 & D \end{pmatrix}$$

und

$$\chi_f(t) = \chi_A(t) = \det \begin{pmatrix} \lambda \cdot I_e - tI_e & B \\ 0 & D - tI_{n-e} \end{pmatrix} = \det((t-\lambda)I_e) \cdot \det(D - tI_{n-e}) = (t-\lambda)^e \cdot \chi_D(t)$$

Sei dann m' die Vielfachheit von λ als Nullstelle von D, so dass $\chi_D(t) = (t-\lambda)^{m'} \cdot h(t)$, mit $h(\lambda) \neq 0$. Dann $\chi_f(t) = (t-\lambda)^{e+m'} \cdot h(t)$ so dass die algebraische Vielfachheit von λ genau $e + m'$ ist, und $e + m' \geq e$. □

Proposition 5.1.25. *Sei V ein $\mathbb{K}$-Vektorraum. Ein Endomorphismus $f\colon V \to V$ ist diagonalisierbar genau dann, wenn das charakteristisches Polynom als Produkt von lineare Faktoren sich zerlegt:*

$$\chi_f(t) = a_n(t-\lambda_1)^{m_1} \cdot \dots \cdot (t-\lambda_r)^{m_r}, \qquad \text{mit } a_n \in \mathbb{K}, a_n \neq 0$$

und die algebraische Vielfachheit und geometrische Vielfachheit jedes Eigenwerts λ_i gleich sind.

Beweis. Sei $n = \dim V$ und seien $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ die paarweise verschiedene Eigenwerte von f und sei e_i (bzw. m_i) die geometrische (bzw. die algebraische) Vielfachheit von λ_i. Dann $\chi_f(t) = (t-\lambda_1)^{m_1} \cdot \dots \cdot (t-\lambda_r)^{m_r} \cdot h(t)$ mit $h(\lambda_i) \neq 0$ für alle i. Wir sehen dass $\deg(h(t)) = \deg(\chi(f)) - \sum_{i=1}^r m_i = n - \sum_{i=1}^n m_i$. Wir wissen auch dass $e_i \leq m_i$ für alle i, so dass

$$\sum_{i=1}^n e_i \leq \sum_{i=1}^n m_i \leq n$$

Wir haben im Satz 4.2.1 gezeigt dass f diagonalisierbar ist, genau dann, wenn $\sum_{i=1}^n e_i = n$. Die obige Ungleichung zeigt dann dass f diagonalisierbar ist genau dann, wenn $e_i = m_i$ für alle i und $\sum_{i=1}^n m_i = n$. Die letzte Bedingung bedeutet genau dass $h(t)$ ein konstantes Polynom ist, d.h. $\chi_f(t)$ ist als Produkt von lineare Faktoren zerlegbar. □

Beispiel 5.1.26. Wir betrachten die Matrix

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

Wir wollen wissen ob A diagonalisierbar ist, und wenn A diagonalisierbar ist wollen wir eine Basis von Eigenvektoren finden. Wir berechnen zuerst das charakteristisches Polynom von A:

$$\begin{aligned}
\chi_A(t) &= \det \begin{pmatrix} -t & 1 & 1 \\ -1 & 2-t & 1 \\ -1 & 1 & 2-t \end{pmatrix} \\
&= -t \cdot \det \begin{pmatrix} 2-t & 1 \\ 1 & 2-t \end{pmatrix} - 1 \cdot \det \begin{pmatrix} -1 & 1 \\ -1 & 2-t \end{pmatrix} + 1 \cdot \det \begin{pmatrix} -1 & 2-t \\ -1 & 1 \end{pmatrix} \\
&= -t((2-t)^2-1) - 1(t-2+1) + (-1+2-t) = -t(t^2+4-4t-1) - t + 1 + 1 - t \\
&= -t^3 + 4t^2 - 5t + 2
\end{aligned}$$

Wir sehen das $\chi_A(1) = -1+4-5+2 = 0$ und wir zerlegen $\chi_A(t) = (t-1)(-t^2+3t-2)$. Wir können die quadratische Gleichung $t^2 - 3t + 2 = 0$ lösen und wir finden dass $(t^2 - 3t + 2) = (t-1)(t-2)$. Dann

$$\chi_A(t) = -(t-1)^2(t-2)$$

Insbesondere ist $\chi_A(t)$ ein Produkt von lineare Faktoren. Wir sehen auch dass A zwei verschiedene Eigenwerte hat: $\lambda_1 = 1$ mit algebraische Vielfachheit 2 und $\lambda_2 = 2$ mit algebraische Vielfachheit 1. Wir wollen die geometrische Vielfachheit auch bestimmen so dass wir betrachten die Eigenräume :

- $\mathrm{Ker}(A-I)$: Das ist

$$\mathrm{Ker} \begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

und man kann berechnen dass eine Basis von $\operatorname{Ker}(A - I)$

$$\mathcal{B}_1 = \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$$

ist. Das zeigt dass $\dim \operatorname{Ker}(A-I) = 2$ so dass die geometrische Vielfachheit vom Eigenwert 1 genau die algebraische Vielfachheit ist.

Wir bemerken dass die algebraische Vielfachheit von $\lambda_2 = 2$ einfach 1 ist, und wir wissen auch dass $\dim \operatorname{Ker}(A - 2I) \geq 1$, weil $\operatorname{Ker}(A - 2I) \neq \{0\}$. Da die geometrische Vielfachheit immer kleiner oder gleich als die algebraische Vielfachheit ist, es muss sein dass $\dim \operatorname{Ker}(A - 2I) = 1$. Das zeigt dass die algebraische und die geometrische Vielfachheiten von beide Eigenwerte gleich sind, so dass wir wissen dass A diagonalisierbar ist auch ohne die Berechnung einer Basis von $\operatorname{Ker}(A - 2I)$. Um eine Basis von Eigenvektoren zu finden, müssen wir aber diese Berechnung durchfuhren:

- $\operatorname{Ker}(A - 2I)$: Das ist

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

  und wir können berechnen dass

$$\mathcal{B}_2 = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

  eine Basis vom $\operatorname{Ker}(A - 2I)$ ist.

Am Ende haben wir dass

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

eine Basis von Eigenvektoren von A ist.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 25.6.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Beispiel 5.1.27. Sei $\lambda \in \mathbb{K}$ und sei

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

Das charakteristisches Polynom von A ist

$$\chi_A(t) = \det \begin{pmatrix} \lambda - t & 1 \\ 0 & \lambda - t \end{pmatrix} = (\lambda - t)^2 = (t - \lambda)^2$$

Das einziges Eigenwert von A ist λ, mit algebraische Vielfachheit 2. Die geometrische Vielfachheit ist

$$\dim \operatorname{Ker}(A - \lambda I) = \dim \operatorname{Ker} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$$

weil die Matrix
$$(A - \lambda I) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$
in Zeilenstufenform ist und mit 1 Pivots, so dass $\mathrm{rk}(A - \lambda I) = 1$ und $\dim \mathrm{Ker}(A - \lambda I) = 2 - \mathrm{rk}(A - \lambda I) = 2 - 1 = 1$. Das zeigt dass die geometrische Vielfachheit von λ kleiner als die algebraische Vielfaccheit von λ ist, so dass die Matrix A nicht diagonalisierbar ist.

Korollar 5.1.28. *Sei V ein $\mathbb{K}$-Vektorraum mit $\dim V = n$ und sei $f\colon V \to V$ ein Endomorphismus mit n paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. Dann ist f diagonalisierbar.*

Beweis. Das charakteristisches Polynom hat Grad n und hat n paarweise verschiedene Nullstellen, so dass $\chi_f(t) = a_n \cdot (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$, mit $a_n \in \mathbb{K}, a_n \neq 0$. Das zeigt dass $\chi_f(t)$ ein Produkt von linearen Faktoren ist und dass die algebraische Vielfachheit jedes Eigenwerts 1 ist. Da die geometrische Vielfachheit mindestens 1 ist, die muss genau 1 sein. Das zeigt dass f diagonalisierbar ist. □

5.2 Polynome und Endomorphismen

Sei V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Sei auch $P(t) = \sum_{h=0}^{n} a_n t^n \in \mathbb{K}[t]$ ein Polynom. Wir können das Polynom in f evaluieren:
$$P(f) \coloneqq \sum_{i=0}^{n} a_n \cdot f^n, \text{ womit } f^n = \underbrace{f \circ f \circ \ldots \circ f}_{n \text{ mal}}, \; f^0 = \mathrm{id}_V$$
so dass $P(f)$ auch ein Endomorphismus $P(f)\colon V \to V$ ist: $P(f) \in \mathrm{End}(V)$. Das definiert eine Abbildung:
$$\mathrm{ev}_f\colon \mathbb{K}[t] \to \mathrm{End}(V), \quad P(t) \mapsto P(f)$$
Diese Abbildung ist linear und multiplikativ: für alle $P, Q \in \mathbb{K}[t], \lambda, \mu \in \mathbb{K}$ gilt dass
$$\mathrm{ev}_f(\lambda \cdot P(t) + \mu \cdot Q(t)) = \lambda P(f) + \mu Q(f), \qquad \mathrm{ev}_f(P(t) \cdot Q(t)) = P(f) \cdot Q(f)$$
und $\mathrm{ev}_f(1) = \mathrm{id}_V$. Insbesondere sehen wir dass für alle $P(t), Q(t) \in \mathrm{End}(V)$ gilt dass
$$P(f) \cdot Q(f) = \mathrm{ev}_f(P(t)Q(t)) = \mathrm{ev}_f(Q(t)P(t)) = Q(f) \cdot P(f)$$

Bemerkung 5.2.1. Wenn $A \in \mathbb{K}^{n \times n}$ eine Matrix ist, können wir ein ähnliches Endomorphismus
$$\mathrm{ev}_A\colon \mathbb{K}[t] \to \mathbb{K}^{n \times n}, \qquad P(t) \mapsto P(A)$$
definieren. Wenn $N \in GL_n(\mathbb{K})$ invertierbar ist, dann gilt (siehe Übungsblatt 8) für jedes $P(t) \in \mathbb{K}[t]$ dass
$$P(NAN^{-1}) = N \cdot P(A) \cdot N^{-1}$$
Seien auch V ein $\mathbb{K}$-Vektorraum, $f\colon V \to V$ ein Endomorphismus, $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und $M^{\mathcal{B}_\mathcal{B}(f)}$ die entsprechende Matrix die f darstellt. Dann gilt dass
$$M^{\mathcal{B}}_{\mathcal{B}}(P(f)) = P(M^{\mathcal{B}}_{\mathcal{B}}(f))$$
So dass Polynome von Matrizen oder Polynome von Endomorphismen im Wesentlichen äquivalent sind.

Insbesondere, können wir die Menge aller Polynome betrachten die in f Null sind.

$$I_f = \mathrm{Ker}(\mathrm{ev}_f) = \{P(t) \in \mathbb{K}[t] \mid P(f) = 0\}$$

Dies ist ein Untervektorraum von $\mathbb{K}[t]$, aber es ist mehr als dass: es gilt dass

$$A(t) \in \mathbb{K}[t], P(t) \in I_f \implies A(t)P(t) \in I_f$$

Das ist einfach zu sehen: $A(f) \cdot P(f) = A(f) \cdot 0 = 0$.

Definition 5.2.2 (Ideal)**.** Sei R ein kommutativer Ring mit 1. Eine Untergruppe $I \subseteq A$ ist ein Ideal wenn

$$a \in R, x \in I \implies ax \in I$$

Wenn $A \in \mathbb{K}^{n \times n}$ eine Quadratmatrix ist, dann können wir auch das Ideal

$$I_A = \{P(t) \in \mathbb{K}[t] \mid P(A) = 0\}$$

betrachten.

Bemerkung 5.2.3. Man kann zeigen (siehe Übungsblatt 8) dass wenn A, B ähnliche Matrizen sind, dann $I_A = I_B$. Außerdem, wenn $f\colon V \to V$ ein Endomorphismus ist, und wenn $\mathcal{B}$ eine Basis von V ist, dann $I_f = I_{M^{\mathcal{B}}_{\mathcal{B}}(f)}$.

5.2.1 Der Satz von Hamilton-Cayley

Wir haben gesehen dass I_f ein Ideal von $\mathbb{K}[t]$ ist. Wir zeigen nun, dass dieses Ideal nicht Null ist (wenn V endlich-dimensionaler ist):

Satz 5.2.4 (Hamilton-Cayley)**.** *Sei V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Sei $\chi_f(t) \in \mathbb{K}[t]$ das charakteristisches Polynom von f. Dann*

$$\chi_f(f) = 0$$

Äquivalent, wenn $A \in \mathbb{K}^{n \times n}$ eine Matrix ist, und $\chi_A(t)$ das charakteristisches Polynom ist, dann

$$\chi_A(A) = 0.$$

Wir brauchen ein bisschen Vorbereitung:

Definition 5.2.5 (Invariante Vektorraum)**.** Seien V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Ein Untervektorraum $W \subseteq V$ ist f-invariant, oder invariant für f falls $f(W) \subseteq W$. In diesem Fall ist die Einschränkung $f_{|W}\colon W \to W$ auch ein Endomorphismus.

Bemerkung 5.2.6. Sei $f\colon V \to V$ ein Endomorphismus und sei $W \subseteq V$ ein f-invariante Untervektorraum. Wenn $P(t) \in \mathbb{K}[t]$ ein Polynom ist, dann ist W auch $P(f)$-invariant und außerdem

$$P(f)_{|W} = P(f_{|W})$$

Definition 5.2.7 (Teilbarkeit)**.** Seien $P(t), Q(t) \in \mathbb{K}[t]$ Polynome. Wir sagen dass $P(t)$ teilt $Q(t)$ und wir schreiben $P(t)|Q(t)$ falls $H(t) \in \mathbb{K}[t]$ existiert so dass

$$Q(t) = P(t) \cdot H(t)$$

Bemerkung 5.2.8. Seien $P(t), Q(t) \in \mathbb{K}[t]$ zwei Polynome mit $P(t)|Q(t)$ und $\deg(Q(t)) \neq 0$. Dann $Q(t) = P(t)H(t)$ für $P(t), H(t) \neq 0$ so dass $\deg(Q(t)) = \deg(P(t)) + \deg(H(t)) \geq \deg(P(t))$.

Lemma 5.2.9. *Seien V ein $\mathbb{K}$-Vektorraum, $f\colon V \to V$ ein Endomorphismus und $W \subseteq V$ ein f-invariante UnterTvektorraum $f(W) \subseteq W$. Das charakteristisches Polynom $\chi_{f_{|W}}(t)$ teilt das charakteristisches Polynom $\chi_f(t)$.*

Beweis. Sei $\mathcal{B}' = (w_1, \dots, w_m)$ eine Basis von W. Wir ergänzen $\mathcal{B}'$ zu einer Basis $\mathcal{B} = (w_1, \dots, w_m, v_{m+1}, \dots, v_n)$ von V. Die Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ ist eine obere Blockdreiecksmatrix mit der Form

$$M^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} M^{\mathcal{B}'}_{\mathcal{B}'}(f_{|W}) & B \\ 0 & D \end{pmatrix}$$

und dann

$$\begin{aligned} \chi_f(t) = \det(M^{\mathcal{B}}_{\mathcal{B}}(f) - tI_n) &= \det \begin{pmatrix} M^{\mathcal{B}'}_{\mathcal{B}'}(f_{|W}) - tI_m & B \\ 0 & D - tI_{n-m} \end{pmatrix} \\ &= \det(M^{\mathcal{B}'}_{\mathcal{B}'}(f_{|W}) - tI_m) \cdot \det(D - tI_{n-m}) = \chi_{f_{|W}}(t) \cdot \chi_D(t) \end{aligned}$$

□

Definition 5.2.10 (Begleitmatrix)**.** Sei $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in \mathbb{K}[t]$ ein normiertes Polynom mit Grad n. Die Begleitmatrix von $P(t)$ ist die $n \times n$-Matrix

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix} \in \mathbb{K}^{n \times n}.$$

Lemma 5.2.11. *Sei $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in \mathbb{K}[t]$ ein normiertes Polynom mit Grad n und sei A die Begleitmatrix von $P(t)$. Das charakteristisches Polynom von A ist $\chi_A(t) = (-1)^n P(t)$.*

Beweis. Der Beweis ist induktiv auf n. Wenn $n = 1$, dann $P(t) = t$ und $A = (0)$, so dass

$\chi_A(t) = t$. Sei $n > 1$. Dann zeigen die Laplace-Entwicklung und die Induktionsannahme dass

$$\chi_A(t) = \det \begin{pmatrix} -t & 0 & 0 & \dots & 0 & -a_0 \\ 1 & -t & 0 & \dots & 0 & -a_1 \\ 0 & 1 & -t & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -t & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} - t \end{pmatrix}$$

$$= -t \det \begin{pmatrix} -t & 0 & 0 & \dots & 0 & -a_1 \\ 1 & -t & 0 & \dots & 0 & -a_2 \\ 0 & 1 & -t & \dots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -t & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} - t \end{pmatrix} + (-1)^{n+1}(-a_0) \begin{pmatrix} 1 & -t & 0 & \dots & 0 \\ 0 & 1 & -t & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$= -t \cdot (-1)^{n-1}(a_1 + a_2 t + \dots + a_{n-1}t^{n-2} + t^{n-1}) + (-1)^{n+1}(-a_0)$$
$$= (-1)^n (a_1 t + \dots + a_{n-1}t^{n-1} + a_0) = (-1)^n P(t).$$

□

Nun können wir den Satz von Hamilton-Cayley beweisen:

Beweis. Wir müssen beweisen dass $\chi_f(f)(v) = 0$ für alle $v \in V$. Wenn $v = 0$, ist das klar. Wenn $v \neq 0$, betrachten wir die Vektoren $v, f(v), f^2(v), \dots$ und die Untervektorräume

$$\langle v \rangle \subseteq \langle v, f(v) \rangle \subseteq \langle v, f(v), f^2(v) \rangle \subseteq \dots$$

Da V endlich-dimensional ist, können diese Inklusionen nicht alle echt sein, sonst haben wir eine unendliche steigende Kette von ganze Zahlen mit eine obere Schranke

$$\dim\langle v \rangle < \dim\langle v, f(v) \rangle < \dim\langle v, f(v), f^2(v) \rangle < \dots \leq \dim V$$

was unmöglich ist. Dann sei m die kleinste positive ganze Zahl, so dass

$$\langle v, f(v), \dots, f^m(v) \rangle = \langle v, f(v), \dots, f^{m-1}(v) \rangle$$

Das bedeutet dass $(v, f(v), \dots, f^{m-1}(v))$ linear unabhängig sind und dass $f^m(v) \in \langle v, \dots, f^{m-1}(v) \rangle$, so dass

$$f^m(v) = a_0 \cdot v + a_1 \cdot f(v) + \dots + a_{m-1} \cdot f^{m-1}(v)$$

für manche $a_i \in \mathbb{K}$. Sei $W = \langle v, f(v), \dots, f^{m-1}(v) \rangle$. Dann $f(W) \subseteq \langle f(v), f^2(v), \dots, f^m(v) \rangle \subseteq W$, so dass W ein f-invariante Untervektorraum ist, und $\chi_{f_{|W}}(t) | \chi_f(t)$. Das bedeutet dass $\chi_f(t) = H(t) \cdot \chi_{f_{|W}}(t)$ für ein $H(t) \in \mathbb{K}[t]$ und dann $\chi_f(f) = H(f) \cdot \chi_{f_{|W}}(f)$. Wenn wir zeigen dass $\chi_{f_{|W}}(f) \cdot v = 0$ dann gilt $\chi_f(f) \cdot v = 0$ auch. Da $v \in W$, müssen wir zeigen dass $\chi_{f_{|W}}(f_{|W}) \cdot v = 0$.

Wir betrachten die Basis $\mathcal{B} = (v, f(v), \dots, f^m(v))$ von W. Die Matrix von $f_{|W}$ hat die Form

$$M_{\mathcal{B}}^{\mathcal{B}}(f_{|W}) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & a_{m-2} \\ 0 & 0 & 0 & \dots & 1 & a_{m-1} \end{pmatrix}$$

Diese Matrix ist die Begleitmatrix des Polynoms $P(t) = t^m - a_{m-1}t^{m-1} - \cdots - a_1 t - a_0$, so dass $\chi_{f_{|W}}(t) = t^m - a_{m-1}t^{m-1} - \cdots - a_0$. Endlich sehen wir dass

$$\chi_{f_{|W}}(f)(v) = f^m(v) - a_{m-1}f^{m-1}(v) - \cdots - a_1 f(v) - a_0 \cdot v = 0$$

□

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 28.6.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

5.2.2 Das Minimalpolynom

Sei V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Der Satz von Hamilton-Cayley zeigt dass das Ideal

$$I_f = \{P(t) \in \mathbb{K}[t] \mid P(f) = 0\}$$

nicht Null ist $I_f \neq \{0\}$. Wir können dann ein wichtiges Ergebnis über die Struktur der Ideale im Polynomring $\mathbb{K}[t]$ verwenden:

Satz 5.2.12. *Sei $I \subseteq \mathbb{K}[t]$ ein Ideal, $I \neq \{0\}$. Dann existiert ein einziges normiertes Polynom $\mu(t) \in I$ von minimalem Grad. Ausserdem gilt dass $P(t) \in I$ genau dann, wenn $\mu(t)|P(t)$. Anders gesagt:*

$$I = \{H(t)\mu(t) \mid H(t) \in \mathbb{K}[t]\}$$

Beweis. In der Vorlesung *Algebraische Strukturen.* □

Definition 5.2.13 (Minimalpolynom)**.** Sei V ein $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Das einziges normiertes Polynom $\mu_f(t) \in I_f$ von minimalem Grad heißt das Minimalpolynom von f.

Sei $A \in \mathbb{K}^{n\times n}$ eine Quadratmatrix. Das einziges normiertes Polynom $\mu_A(t) \in I_A$ von minimalem Grad heißt das Minimalpolynom von f.

Die Eigenschaften das Minimalpolynom zeigen dass für jedes $P(t) \in \mathbb{K}[t]$ folgendes gilt:

$$P(f) = 0 \iff \mu_f(t)|P(t)$$

und μ_f ist das einziges normiertes Polynom mit dieser Eigenschaft (warum?). Ähnliches gilt für das Minimalpolynom einer Matrix.

Bemerkung 5.2.14. Bemerkung 5.2.3 zeigt dass wenn A, B ähnliche Matrizen sind, dann $\mu_A(t) = \mu_B(t)$. Außerdem, wenn $f\colon V \to V$ ein Endomorphismus ist, und wenn $\mathcal{B}$ eine Basis von V ist, dann $\mu_f(t) = \mu_{M^{\mathcal{B}}_{\mathcal{B}}(t)}$.

Beispiel 5.2.15. Sei $D = \lambda I$ eine Diagonalmatrix mit alle Einträge in der Hauptdiagonale gleich λ. Wir betrachten das Polynom $(t - \lambda)$. Wir sehen dass $(t - \lambda I)(D) = (D - \lambda I) = 0$ so dass $t - \lambda \in I_{\lambda I}$. Da es kein Polynom vom Grad Null geben kann, das auf D Null ist (warum?), ist dies das minimale Polynom von D: $\mu_{\lambda I}(t) = t - \lambda$. Insbesondere sehen wir dass für ein Polynom $P(t) \in \mathbb{K}[t]$ gilt dass

$$P(\lambda I) = 0 \iff (t - \lambda)|P(t) \iff P(\lambda) = 0$$

Beispiel 5.2.16. Sei D eine Diagonalmatrix mit paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_r$: das bedeutet dass D eine Blockmatrix ist

$$D = \begin{pmatrix} \lambda_1 I_{n_1} & 0 & \ldots & 0 \\ 0 & \lambda_2 I_{n_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_r I_{n_r} \end{pmatrix}$$

Sei $P(t) \in \mathbb{K}[t]$ ein Polynom. Man kann leicht sehen dass

$$P(D) = \begin{pmatrix} P(\lambda_1 I_{n_1}) & 0 & \ldots & 0 \\ 0 & P(\lambda_2 I_{n_2}) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & P(\lambda_r I_{n_r}) \end{pmatrix}$$

Insbesondere $P(D) = 0$ genau dann, wenn $P(\lambda_1 I) = \cdots = P(\lambda_r I) = 0$ und das ist äquivalent zu $P(\lambda_1) = \cdots = P(\lambda_r) = 0$. Korollar 3.1.32 zeigt dann, dass

$$P(D) = 0 \iff (t - \lambda_1) \cdot \cdots \cdot (t - \lambda_r) | P(t)$$

Insbesondere sehen wir dass

$$\mu_D = (t - \lambda_1) \cdot \cdots \cdot (t - \lambda_r)$$

Beispiel 5.2.17. Wir betrachten die Matrix

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

Wir sehen dass $A^2 = 0$ so dass t^2 ein normiertes Polynom von Grad 2 in I_A ist. Wir sehen auch dass $A - \lambda I \neq 0$ für jedes $\lambda \in \mathbb{K}[t]$ so dass I_A kein Polynom von Grad 1 enthält. Das zeigt dass $\mu_A(t) = t^2$.

Proposition 5.2.18. *Sei V ein $\mathbb{K}$-Vektorraum mit $\dim V = n$ und sei $f\colon V \to V$ ein Endomorphismus.*

1. *Das Minimalpolynom $\mu_f(t)$ hat Grad $\deg(\mu_f(t)) \leq n$.*
2. *Die Nullstellen $\lambda \in \mathbb{K}$ des Minimalpolynoms sind genau die Eigenwerte von f.*
3. *Wenn $W \subseteq V$ ein f-invariante UnterVektorraum ist, dann $\mu_{f_{|W}}(t) | \mu_f(t)$.*

Beweis. 1. Das charakteristisches Polynom $\chi_f(t)$ hat Grad n und $\chi_f(f) = 0$, so dass $\mu_f(t) | \chi_f(f)$. Dann $\deg(\mu_f(t)) \leq \deg(\chi_f(t)) = n$.

2. Wir wissen dass $\chi_f(t) = H(t)\mu_f(t)$ für ein $H(t) \in \mathbb{K}[t]$. Sei $\lambda \in \mathbb{K}$ eine Nullstelle von μ_f: dann $\chi(f)(\lambda) = H(\lambda)\mu_f(\lambda) = 0$ so dass λ auch eine Nullstelle von $\chi_f(t)$ ist. Das bedeutet dass λ ein Eigenwert von f ist. Andererseits, sei $\lambda \in \mathbb{K}$ ein Eigenwert von f. Dann existiert $v \in V, v \neq 0$ mit $f(v) = \lambda \cdot v$. Sei $P(t) = \sum_{h=0}^n a_h t^h$ ein Polynom, dann $P(f)(v) = \sum_{h=0}^n a_h f^n(v) = \sum_{h=0}^n a_h \lambda^n v = P(\lambda)v$. Insbesondere $\mu_f(\lambda) \cdot v = \mu_f(f)(v) = 0(v) = 0$. Da $v \neq 0$, bedeutet dass das $\mu_f(\lambda) = 0$.

3. Wir sehen dass $\mu_f(f_{|W}) = \mu_f(f)_{|W} = 0_{|W} = 0$, so dass $\mu_f \in I_{f_{|W}}$. Das bedeutet dass $\mu_{f_{|W}}(t)|\mu_f(t)$.

$\square$

Satz 5.2.19. *Sei V ein $\mathbb{K}$-Vektorraum. Ein Endomorphismus $f\colon V \to V$ ist diagonalisierbar genau dann, wenn dass Minimalpolynom $\mu_f(t)$ die Form*

$$\mu_f(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_r)$$

hat, mit $\lambda_1, \dots, \lambda_r$ paarweise Verschiedene. In diesem Fall sind die $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ die Eigenwerte von f.

Beweis. Nehmen wir an dass f diagonalisierbar ist, und sei $\mathcal{B}$ eine Basis von Eigenvektoren von f. Dann $D = M^{\mathcal{B}}_{\mathcal{B}}(f)$ ist diagonal mit Eigenwerte $\lambda_1, \dots, \lambda_r$ und wir haben schon gesehen dass $\mu_f(t) = \mu_D(t) = (t - \lambda_1) \dots (t - \lambda_r)$.

Andererseits, nehmen wir an dass $\mu_f(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_r)$ mit λ_i paarweise Verschiedene. Dann zeigt die vorherige Proposition dass die λ_i genau die Eigenwerte von f sind. Wir definieren die Polynome

$$P_i(t) = (t - \lambda_1) \cdot \dots \cdot \widehat{(t - \lambda_i)} \cdot \dots \cdot (t - \lambda_r), \quad \text{für } i \in \{1, \dots, r\}$$

Wir sehen dass $P_i(\lambda_j) = 0$ für $i \neq j$ und $P_i(\lambda_i) \neq 0$. Wir betrachten dann das Polynom

$$P(t) = \sum_{i=1}^{r} \frac{1}{P_i(\lambda_i)} \cdot P_i(t)$$

Dies hat Grad $\deg(P(t)) \leq r - 1$, weil jedes Polynom $P_i(t)$ Grad $\deg(P_i(t)) \leq r - 1$ hat. Außerdem, gilt dass $P(\lambda_i) = \frac{1}{P_i(\lambda_i)} P_i(\lambda_i) = 1$. Dann hat das Polynom $P(t) - 1$ Grad kleiner als r und gleichzeitig r verschiedene Nullstellen. Der Korollar 3.1.32 zeigt dass $P(t) = 1$, so dass $P(f) = \mathrm{id}_V$.

Wir betrachten nun $P_i(f)$ als Endomorphismus $P_i(f)\colon V \to V$. Wir sehen dass $(f - \lambda_i \,\mathrm{id}_V) P_i(f) = (f - \lambda_i \,\mathrm{id}_V) \prod_{j \neq i} (f - \lambda_j \,\mathrm{id}_V) = \prod_{i=1}^{r} (f - \lambda_r \,\mathrm{id}_V) = \mu_f(f) = 0$. Das bedeutet dass $\operatorname{Im} P_i(f) \subseteq \operatorname{Ker}(f - \lambda_i \,\mathrm{id}_V)$: sei dann $v \in V$. Wir sehen dass

$$v = \mathrm{id}_V(v) = P(f)(v) = \sum_{i=1}^{r} \frac{1}{P_i(\lambda_i)} P_i(f)(v), \quad \text{mit} \quad \frac{1}{P_i(\lambda_i)} P_i(f)(v) \in \operatorname{Ker}(f - \lambda_i \,\mathrm{id}_V)$$

Das zeigt dass $V = \operatorname{Ker}(f - \lambda_1 \,\mathrm{id}_V) \oplus \dots \oplus \operatorname{Ker}(f - \lambda_r \,\mathrm{id}_V)$ so dass f diagonalisierbar ist. $\square$

Korollar 5.2.20. *Seien V ein $\mathbb{K}$-Vektorraum, $f\colon V \to V$ ein Endomorphismus und $W \subseteq V$ ein f-invariante Vektorraum. Wenn f diagonalisierbar ist, dann ist $f_{|W}$ auch diagonalisierbar.*

Beweis. Da f diagonalisierbar ist, wissen wir dass $\mu_f(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_r)$ mit λ_i paarweise verschiedene. Da W invariant für f ist, wissen wir dass $\chi_{f_{|W}}(t)|\chi_f(t)$. Eine Tatsache aus der Algebra besagt, dass $\mu_{f_{|W}}(t)$ auch ein Produkt von paarweise verschiedenen normierte Polynome vom Grad 1 sein muss. Dies zeigt, dass $f_{|W}$ auch diagonalisierbar ist.

Wenn wir diese Tatsache aus der Algebra nicht verwenden wollen, dann sei $P_i(f)$ wie im Beweis des vorherigen Satzes. Wir haben gezeigt dass $\operatorname{Im} P_i(f) \subseteq \operatorname{Ker}(f - \lambda_i \,\mathrm{id}_V)$. Außerdem, sehen wir dass $P_i(f)(W) \subseteq W$ so dass für alle $w \in W$ gilt dass

$$w = P(f)w = \sum_{i=1}^{r} \frac{1}{P_i(\lambda_i)} P_i(f)w$$

mit $P_i(f)w \in \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \cap W = \operatorname{Ker}(f_{|W} - \lambda_i \operatorname{id}_W)$. Dann sehen wir wie fruher dass $f_{|W}$ diagonalisierbar ist. □

5.3 Die Jordansche Normalform

Sei V ein $\mathbb{K}$-Vektorraum und sei $f: V \to V$ ein Endomorphismus. Wir haben gesehen, dass f nicht diagonalisierbar sein muss. Ein Problem könnte sein, dass es einfach nicht genug Eigenwerte in $\mathbb{K}$ gibt, was bedeutet, dass sich das charakteristische Polynom $\chi_f(t)$ nicht als Produkt von Linearfaktoren zerlegt.

Ein Problem könnte sein, dass es einfach nicht genug Eigenwerte in $\mathbb{K}$ gibt, was bedeutet, dass sich das charakteristische Polynom $\chi_f(t)$ nicht als Produkt von Linearfaktoren zerlegt.

Ein anderes Problem ist, dass das charakteristische Polynom ein Produkt von Linearfaktoren sein könnte:

$$\chi_f(t) = (t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_r)^{m_r}$$

aber manche geometrische Vielfachheite können kleiner sein als die algebraischen Vielfachheite. In diesem zweiten Fall ist der Endomorphismus nicht diagonalisierbar, aber er kann dennoch in eine spezielle Form gebracht werden.

Definition 5.3.1 (Jordansche Normalform)**.** Ein Jordanblock mit Eigenwert λ ist eine Quadratmatrix mit der Form

$$J = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ 0 & 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda \end{pmatrix}$$

Eine Quadratmatrix A ist in Jordansche Normalform wenn A eine Blockmatrix ist mit der Form

$$A = \begin{pmatrix} J_1 & 0 & 0 & \ldots & 0 \\ 0 & J_2 & 0 & \ldots & 0 \\ 0 & 0 & J_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & J_s \end{pmatrix}$$

ist, wobei die J_i Jordanblöcke sind.

Satz 5.3.2 (Jordansche Normalform für Endomorphismen)**.** *Seien V ein $\mathbb{K}$-Vektorraum und $f: V \to V$ ein Endomorphismus. Wenn dass charakteristisches Polynom $\chi_f(t)$ als Produkt von Linearenfaktoren zerlegt dann existiert eine Basis $\mathcal{B}$ von V so dass die entsprechende Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ in Jordansche Normalform ist. Außerdem ist die Jordansche Normalform bis auf eine mögliche Umordnung der Jordanblöcke nicht von der Basis abhängig*

Als Korollar bekommen wir

Satz 5.3.3 (Jordan Normalform für Matrizen)**.** *Sei $A \in \mathbb{K}^{n \times n}$ eine Matrix so dass das charakteristisches Polynom $\chi_A(t)$ als Produkt von Linearfaktoren zerlegt. Dann ist A ähnlich zu eine Matrix in Jordansche Normalform. Diese heißt die Jordansche Normalform von A.*

Eine andere Matrix $B \in \mathbb{K}^{n\times n}$ ist ähnlich zu A genau dann, wenn das charakteristische Polynom von B in ein Produkt von Linearfaktoren zerfällt und die beiden Jordansche Normalformen von A und B bis auf eine Umordnung der Jordanblöcke gleich sind.

Insbesondere teilt sich über einem algebraisch geschlossenen Feld jedes nicht konstante Polynom in ein Produkt von Linearfaktoren auf. Da wir wissen, dass $\mathbb{C}$ algebraisch geschlossen ist, erhalten wir

Korollar 5.3.4. *Sei V ein $\mathbb{C}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus. Dann existiert eine Basis $\mathcal{B}$ von V so dass $M^{\mathcal{B}}_{\mathcal{B}}(f)$ in Jordansche Normalform ist.*

Sei $A \in \mathbb{C}^{n\times n}$ eine Matrix. Dann ist A ähnlich zu einer Matrix in Jordansche Normalform.

Wir schreiben hier einen Beweis, den wir am Ende des Kurses diskutieren werden, falls wir genug Zeit haben

5.3.1 Nilpotente Endomorphismen

Wir betrachten zuerst den Fall wo das einziges Eigenwert $\lambda = 0$ ist. Ein Jordanblock mit Eigenwert $\lambda = 0$ hat die Form

$$J = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Bemerkung 5.3.5. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, sei $f\colon V \to V$ ein Endomorphismus und sei $\mathcal{B}$ eine Basis von V. Die Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ ist ein Jordanblock mit Eigenwert $\lambda = 0$ genau dann wenn die Basis $\mathcal{B}$ die Form

$$\mathcal{B} = (f^{n-1}(v), f^{n-2}(v), \dots, f(v), v)$$

hat, für ein Vektor v mit $f^n(v) = 0$. Insbesondere gilt dass $f^n = 0$.

Definition 5.3.6 (Nilpotente Endomorphismen und Matrizen)**.** Sei $\mathbb{K}$ ein Körper und sei V ein $\mathbb{K}$-Vektorraum. Ein Endomorphismus $f\colon V \to V$ ist nilpotent, falls $f^m = 0$ für ein $m \in \mathbb{N}$. Eine Matrix $A \in \mathbb{K}^{n\times n}$ ist nilpotent, falls $A^m = 0$ für ein $m \in \mathbb{N}$.

Satz 5.3.7 (Existenz der Jordan-Normalform für nilpotente Endomorphismen)**.** *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ nilpotent.*

Es gibt eine Basis $\mathcal{B}$ von V so dass $M^{\mathcal{B}}_{\mathcal{B}}(f)$ in Jordansche Normalform ist mit Jordanblöcken mit Eigenwert 0.

Um diesen Satz zu beweisen, folgen wir Mark Wildon in "*A short proof of the existence of Jordan normal form*" (man kann das einfach online finden). Wir brauchen zwei Lemmata

Lemma 5.3.8. *Sei $f\colon V \to V$ ein Endomorphismus und sei $P(t) \in \mathbb{K}[t]$ ein Polynom. Die Untervektorraume $\operatorname{Ker} P(f)$ und $\operatorname{Im} P(f)$ sind f-invariant*

Beweis. Wir sehen dass $P(f) \circ f = f \circ P(f)$. Sei nun $v \in \operatorname{Ker} P(f)$. Dann $P(f)(f(v)) = f(P(f)(v)) = f(0) = 0$ so dass $f(v) \in \operatorname{Ker} P(f)$. Sei auch $v = P(f)(w) \in \operatorname{Im} P(f)$, wobei $w \in V$ ein Vektor ist. Dann ist $f(v) = f(P(f)(w)) = P(f)(f(w))$ ein Vektor in $\operatorname{Im} P(f)$. □

Beweis der Existenz der Jordansche Normalform für Nilpotente Endomorphismen. Wir zeigen dass es eine Basis von V gibt, die die Form

$$\mathcal{B} = (f^{k_1-1}(v_1), f^{k_1-2}(v_1), \dots, f(v_1), v_1, \dots, f^{k_s-1}(v_s), f^{k_s-2}(v_s), \dots, f(v_s), v_s)$$

hat, mit $v_i \in V$ und $f^{k_i}(v_i) = 0$. Der Beweis erfolgt durch Induktion auf $\dim V$. Wenn $\dim V = 0$ ist, gibt es nichts zu beweisen. Wenn $\dim V \geq 1$,können wir auch annehmen dass $f \neq 0$, denn in diesem Fall ist das Ergebnis offensichtlich. Wenn $\operatorname{Im} f = V$ dann ist f ein Isomorphismus, und das ist unmöglich, da f nilpotent ist. Denn gilt dass $\{0\} \subsetneq \operatorname{Im} f \subsetneq V$. Der Untervektorraum $f(V)$ ist f-invariant, das Endomorphismus $f_{|\operatorname{Im} f}$ ist nilpotent, und $\dim \operatorname{Im} f < \dim V$. Dann zeigt die Induktionsannahme das wir eine Basis $\mathcal{B}'$ von $\operatorname{Im} f$ mit der folgenden Form haben:

$$\mathcal{B}' = (f^{h_1-1}(w_1), f^{h_1-2}(w_1), \dots, f(w_1), w_1, \dots, f^{h_r-1}(w_r), f^{h_r-2}(w_r), \dots, f(w_r), w_r)$$

womit $w_i \in f(V)$ und $f^{h_i}(w_i) = 0$. Da $w_i \in \operatorname{Im} f$ existieren $v_i \in V$ so dass $w_i = f(v_i)$. Dann sind $f^{h_1}(v_1), \dots, f^{h_r}(v_r) \in \operatorname{Ker} f$ linear unabhängig. Wir können sie zu einer Basis $(f^{h_1}(v_1), \dots, f^{h_r}(v_r), u_1, \dots, u_t)$ von $\operatorname{Ker}(f)$ erweitern. Wir wollen nun zeigen dass

$$\mathcal{B} = (f^{h_1}(v_1), f^{h_1-1}(v_1), \dots, f^2(v_1), f(v_1), v_1, \dots, f^{h_r}(v_r), f^{h_r-1}(v_r), \dots, f^2(v_r), f(v_r), v_r, u_1, \dots, u_t)$$

eine Basis von V ist. Die Anzahl der Vektoren ist

$$\sum_{i=1}^{r}(h_i + 1) + t = \sum_{i=1} h_i + r + t = \dim \operatorname{Im} f + \dim \operatorname{Ker} f = \dim V$$

Wir müssen nun zeigen dass die Vektoren linear unabhängig sind. Nehmen wir an dass

$$\sum_{i=1}^{r}\sum_{j=0}^{h_i} \lambda_{i,j} f^j(v_i) + \sum_{\ell=1}^{t} \mu_\ell u_\ell = 0 \qquad \text{für } \lambda_{i,j}, \mu_\ell \in \mathbb{K}$$

Wenn wir f anwenden, sehen wir, dass

$$\sum_{i=1}^{r}\sum_{j=0}^{h_i-1} \lambda_{i,j} f^j(w_i) = 0$$

und da $\mathcal{B}'$ eine Basis ist, sehen wir dass $\lambda_{i,j} = 0$ für alle $1 \leq i \leq r, 0 \leq j \leq h_i - 1$. Dann bleiben wir mit

$$\sum_{i=1}^{r} \lambda_{i,h_i} f^{h_i}(v_i) + \sum_{\ell=1}^{t} \mu_\ell u_\ell = 0 \qquad \text{für } \lambda_{i,j}, \mu_\ell \in \mathbb{K}$$

und da die Vektoren in dieser Linearkombination linear unabhängig in $\operatorname{Ker} f$ sind, sehen wir dass $\lambda_{i,h_i} = 0$ für alle i und $\mu_\ell = 0$ für alle ℓ. □

5.3.2 Die Jordansche Normalform: Existenz

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

Lemma 5.3.9. *Sei $g\colon V \to V$ ein Endomorphismus. Dann existiert $n \in \mathbb{N}$ so dass*

$$\operatorname{Ker} g^n = \operatorname{Ker} g^{n+1}$$

In diesem Fall, $\operatorname{Ker} g^n = \operatorname{Ker} g^{n+m}$ *für alle $m \geq 0$, und*

$$V = \operatorname{Ker} g^n \oplus \operatorname{Im} g^n$$

Beweis. Wie sehen dass $\operatorname{Ker} g^k \subseteq \operatorname{Ker} g^{k+1}$ für alle $k \in \mathbb{N}$. Da V endlich-dimensionaler ist, die Inklusionen

$$\operatorname{Ker} g \subseteq \operatorname{Ker} g^2 \subseteq \operatorname{Ker} g^3 \subseteq \dots$$

können nicht alle echt sein. Dann existiert n so dass $\operatorname{Ker} g^n = \operatorname{Ker} g^{n+1}$. Wir zeigen durch Induktion auf m dass $\operatorname{Ker} g^n = \operatorname{Ker} g^{n+m}$ für alle $m \in \mathbb{N}$. Wenn $m = 0, 1$, dann ist das wahr. Wenn $m > 1$, sei $x \in \operatorname{Ker} g^{n+m}$: dann $g^{n+1}(g^{m-1}(x)) = g^{n+m}(x) = 0$. Das bedeutet dass $g^{m-1}(x) \in \operatorname{Ker} g^{n+1} = \operatorname{Ker} g^n$, so dass $g^n(g^{m-1}(x)) = g^{n+m-1}(x) = 0$. Das bedeutet dass $x \in \operatorname{Ker} g^{n+m-1}$, und die Induktionsannahme zeigt dass $\operatorname{Ker} g^{n+m-1} = \operatorname{Ker} g^n$. Das zeigt dass $\operatorname{Ker} g^{n+m} \subseteq \operatorname{Ker} g^n$, und die andere Inklusion ist immer wahr.

Endlich, sehen wir dass $\dim V = \dim \operatorname{Ker} g^n + \dim \operatorname{Im} g^n$. Wir müssen zeigen dass $(\operatorname{Ker} g^n) \cap (\operatorname{Im} g^n) = 0$. Sei $v \in (\operatorname{Ker} g^n) \cap (\operatorname{Im} g^n)$: da $v \in (\operatorname{Im} g^n)$, wissen wir das $v = g^n(w)$ für ein $w \in V$. Da $v \in \operatorname{Ker} g^n$, wissen wir dass $g^{2n}(w) = g^n(v) = 0$. Dann $w \in \operatorname{Ker} g^{2n} = \operatorname{Ker} g^n$, so dass $v = g^n(w) = 0$. □

Seien V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus so dass

$$\chi_f(t) = a_n(t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_r}$$

in Linearfaktoren zerfallt, mit $\lambda_i \in \mathbb{K}$ die paarweise verschiedene Eigenwerte. Die m_i sind die algebraische Vielfachheiten.

Definition 5.3.10 (Hauptraum)**.** Für jedes Eigenwert λ_i, sei $n_i \in \mathbb{N}$ so dass $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)^{n_i} = \operatorname{Ker}(f - \lambda_i \operatorname{id}_V)^{n_i+1}$. Der Untervektorraum $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)^{n_i}$ ist der Hauptraum der Eigenwert λ_i.

Bemerkung 5.3.11. Das vorheriges Lemma zeigt dass der Hauptraum existiert, dass er unabhängig von den gewählten n_i ist und dass er positive Dimension hat, da er mindestens den Eigenraum $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ enthält.

Satz 5.3.12 (Existenz der Jordansche Normalform)**.** *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus so dass*

$$\chi_f(t) = a_n(t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_r}$$

in Linearfaktoren mit $\lambda_i \in \mathbb{K}$ zerfallt. Dann existiert eine Basis $\mathcal{B}$ von V so dass $M^{\mathcal{B}}_{\mathcal{B}}(f)$ Jordansche Normalform hat.

Beweis. Der Beweis ist durch Induktion auf $n = \dim V$. Wenn $n = 0$ es gibt nichts zu zeigen. Nehmen wir an dass $n > 1$ und dass die Aussage für alle Vektorräume von Dimension kleiner als n wahr ist. Sei λ_1 ein Eigenwert von f und sei $\operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)^{n_1}$ der entsprechende Hauptraum. Dann wissen wir dass

$$V = \operatorname{Ker}(f - \lambda_1 \operatorname{id}_V)^{n_1} \oplus \operatorname{Im}(f - \lambda_1 \operatorname{id}_V)^{n_1}$$

und außerdem sind beide Untervektorräume f-invariant. Wir wissen dass $\dim \operatorname{Im}(f-\lambda_1 \operatorname{id}_V)^{n_1} < \dim V$ und die Induktionsannahme zeigt dass eine Basis $\mathcal{B}''$ von $\operatorname{Im}(f-\lambda_1 \operatorname{id}_V)^{n_1}$ existiert so dass $M_{\mathcal{B}''}^{\mathcal{B}''}(f_{|\operatorname{Im}(f-\lambda_1 \operatorname{id}_V)^{n_1}})$ Jordansche Normalform hat. Wenn eine Basis $\mathcal{B}'$ von $\operatorname{Ker}(f-\lambda_1 \operatorname{id}_V)^{n_1}$ existiert so dass $M_{\mathcal{B}'}^{\mathcal{B}'}(f_{|\ker(f-\lambda_1 \operatorname{id}_V)^{n_1}})$ Jordansche Normalform hat, dann ist $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$ eine Basis von V und man kann leicht sehen dass $M_{\mathcal{B}}^{\mathcal{B}}(f)$ Jordansche Normalform hat.

Um zu zeigen dass $\mathcal{B}'$ existiert, betrachten wir die Einschränkung $g = (f-\lambda_1 \operatorname{id}_V)_{|\operatorname{Ker}(f-\lambda_1 \operatorname{id}_V)^{n_1})}$. Wir sehen dass $g^{n_1} = 0$ so dass g nilpotent ist und dann existiert eine Basis $\mathcal{B}'$ so dass $M_{\mathcal{B}'}^{\mathcal{B}'}(g)$ Jordansche Normalform hat, mit alle Jordanblöcke mit Eigenwert 0. Dann sehen wir dass $f_{|\operatorname{Ker}(f-\lambda_1 \operatorname{id}_V)^{n_1})} = g + \lambda_1 \cdot \operatorname{id}$, so dass $M_{\mathcal{B}'}^{\mathcal{B}'}(f_{|\operatorname{Ker}(f-\lambda_1 \operatorname{id}_V)^{n_1})}) = M_{\mathcal{B}'}^{\mathcal{B}'}(g) + \lambda_1 I$ Jordansche normalform hat, mit alle Jordanblöcke mit Eigenwert λ_1. $\square$

Kapitel 6

Euklidische und Unitäre Räume

6.1 Euklidische Räume

Definition 6.1.1 (Skalarprodukt)**.** Sei V ein $\mathbb{R}$-Vektorraum. Ein Skalarprodukt auf $\mathbb{R}$ ist eine Abbildung

$$\langle\ ,\ \rangle\colon V \times V \to \mathbb{R}, \qquad (v, w) \mapsto \langle v, w\rangle$$

mit den folgende Eigenschäften:

- **Bilinear** (linear in jeder Komponente): für alle $v, v', w, w' \in V$ und $\lambda, \lambda' \in \mathbb{R}$ gilt dass

$$\langle \lambda v + \lambda' v', w\rangle = \lambda\langle v, w\rangle + \lambda'\langle v', w\rangle$$
$$\langle v, \mu w + \mu' w'\rangle = \mu\langle v, w\rangle + \mu'\langle v, w'\rangle$$

- **Symmetrisch**: für alle $v, w \in V$ gilt dass

$$\langle v, w\rangle = \langle w, v\rangle.$$

- **Positiv definit**: für alle $v \in V$ gilt dass

$$\langle v, v\rangle \geq 0, \quad \text{ und } \langle v, v\rangle = 0 \iff v = 0$$

Ein $\mathbb{R}$-Vektorraum V mit einem Skalarprodukt $\langle\rangle$ heißt ein euklidischer Raum.

Definition 6.1.2 (Bilinearform)**.** Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$b\colon V \times V \to \mathbb{K}, \qquad (v, w) \mapsto b(v, w)$$

heißt Bilinearform, falls sie bilinear ist. Sie heißt symmetrische Bilinearform, falls sie bilinear und symmetrisch ist.

Bemerkung 6.1.3. Ein Skalarprodukt ist eine reelle Bilinearform die symmetrisch und positiv-definit ist. Wir sehen, dass die Definition von Bilinearität und Symmetrie für jedes Körper sinnvoll ist, während die Definition von positiv-definit nur für die reellen Zahlen sinnvoll ist. Wir werden später sehen, wie man sie auf die komplexen Zahlen generalisieren kann.

Beispiel 6.1.4. Sei $n \geq 1$. Das Standard-Skalarprodukt auf $\mathbb{R}^n$ ist

$$\langle x, y\rangle = x^t \cdot y = \sum_{i=1}^{n} x_i y_i, \qquad \text{für} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Man kann leicht zeigen dass dieses ein Skalarprodukt ist: die Formel mit den Koordinaten von x zeigt dass das Produkt symmetrisch ist. Da es symmetrisch ist, reicht es für den Beweis, dass es bilinear ist, zu beweisen, dass es in der ersten Komponente linear ist: seien $x, x', y \in \mathbb{R}^n$ und $\lambda, \lambda' \in \mathbb{R}$. Dann

$$\langle \lambda x + \lambda' x', y\rangle = (\lambda x + \lambda' x')^t \cdot y = \lambda x^t \cdot y + \lambda' x'^t \cdot y = \lambda\langle x, y\rangle + \lambda'\langle x', y\rangle$$

Wir zeigen nun dass es positiv-definit ist: sei $x \in \mathbb{R}^n$. Dann

$$\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 \geq 0$$

und $\langle x, x\rangle = 0$ genau dann, wenn $x_i^2 = 0$ für alle i, was bedeutet $x_i = 0$ für alle i, d.h. $x = 0$.

Beispiel 6.1.5. Sei $V = C^0([0,1], \mathbb{R})$ die Menge von alle stetige Funktionen

$$f\colon [0,1] \to \mathbb{R}$$

Diese Menge ist ein $\mathbb{R}$-Vektorraum mit den übliche Summe von Funktionen und Multiplikation einer Funktion mit einer $\lambda \in \mathbb{R}$. Man definiert dass L^2-Skalarprodukt auf V als

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx$$

Das ist wohldefiniert weil das Produkt von stetige Funktionen eine stetige Funktion ist. Die Definition zeigt dass $\langle\, ,\, \rangle$ symmetrisch ist und die Bilinearität ist eine Konsequenz der Linearität des Integrals. Wir zeigen dass $\langle\, ,\, \rangle$ positiv definit ist: sei $f \in C^0([0,1])$, dann

$$\langle f, f\rangle = \int_0^1 f^2(x)dx$$

und $\langle f, f\rangle \geq 0$, weil $f^2(x) \geq 0$. Außerdem wenn $x_0 \in [0,1]$ existiert so dass $f(x_0) \neq 0$, dann, da f stetig ist, können wir annehmen dass $x_0 \in (0,1)$. Wir wissen dass $|f(x_0)| > 0$, und, da f stetig ist, existiert $\delta > 0$ und $m > 0$ so dass $|f(x)| > m$ für alle $x \in (x_0 - \delta, x_0 + \delta)$. Dann

$$\langle f, f\rangle = \int_0^1 f^2 dx \geq \int_{x_0-\delta}^{x_0+\delta} f^2(x)dx \geq \int_{x_0-\delta}^{x_0+\delta} m^2 dx = 2\delta \cdot m^2 > 0$$

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 2.7.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

6.1.1 Die Gramsche Matrix

Seien $n \geq 1$ und sei $A \in \mathbb{R}^{n\times n}$ eine Matrix. Wir definieren die Abbildung

$$\langle\, ,\, \rangle_A \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (x,y) \mapsto x^t A y = \sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i y_j$$

Lemma 6.1.6. *1.* $A_{ij} = \langle e_i, e_j\rangle_A$.

2. $\langle\, ,\, \rangle_A$ *ist bilinear.*

3. $\langle\, ,\, \rangle_A$ *ist symmetrisch genau dann, wenn* $A = A^t$.

Beweis. 1. Die Formel $\langle x,y\rangle_A = \sum_{i,j} A_{ij} x_i x_j$ zeigt dass $\langle e_i, e_j\rangle_A = A_{ij}$.

2. Seien $x, x', y, y' \in \mathbb{R}^n$ und seien $\lambda, \lambda', \mu, \mu' \in \mathbb{R}$. Dann

$$\begin{aligned}\langle \lambda x + \lambda' x', y\rangle_A &= (\lambda x + \lambda' x')^t A y = \lambda x^t A y + \lambda'(x')^t A y = \lambda\langle x,y\rangle_A + \lambda'\langle x', y\rangle_A \\ \langle x, \mu y + \mu' y'\rangle_A &= x^t A(\mu y + \mu' y') = \mu x^t A y + \mu' x^t A y' = \mu\langle x,y\rangle_A + \mu'\langle x, y'\rangle_A\end{aligned}$$

3. Wenn $A = A^t$, dann $\langle x,y\rangle_A = x^t A y = x^t A^t y = (Ax)^t y = y^t A x = \langle y,x\rangle_A$. Andererseits, wenn $\langle\, ,\, \rangle_A$ symmetrisch ist, dann $A_{ij} = \langle e_i, e_j\rangle_A = \langle e_j, e_i\rangle_A = A_{ji}$, so dass $A = A^t$.

□

Definition 6.1.7 (Symmetrische Matrix)**.** Sei $\mathbb{K}$ ein beliebiger Körper. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt symmetrisch, falls $A^t = A$.

Definition 6.1.8 (Positiv definite Matrix)**.** Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist positiv definit, falls für alle $x \in \mathbb{R}^n$ gilt dass

$$x^t A x \geq 0, \qquad \text{und} \qquad x^t A x = 0 \iff x = 0$$

Bemerkung 6.1.9. Sei $A \in \mathbb{R}^{n\times n}$ eine Matrix. Die bilineare Abbildung $\langle\, ,\, \rangle_A$ ist ein Skalarprodukt genau dann, wenn A symmetrisch und positiv definit ist.

Beispiel 6.1.10. Sei $V = \mathbb{R}^n$ und sei $A = I_n$. Dann ist $\langle\, ,\, \rangle$ das Standard-Skalarprodukt auf $\mathbb{R}^n$.

Andererseits, sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und sei $\langle\, ,\, \rangle$ ein Skalarprodukt auf V. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Wir definieren die Gramsche Matrix von $\langle\, ,\, \rangle$ bezüglich $\mathcal{B}$ als:

$$M_{\mathcal{B}}(\langle\, ,\, \rangle) = \begin{pmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \ldots & \langle v_1, v_n\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \ldots & \langle v_2, v_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1\rangle & \langle v_n, v_2\rangle & \ldots & \langle v_n, v_n\rangle \end{pmatrix}$$

Beispiel 6.1.11. Sei $V = \mathbb{R}^n$ und sei $\mathcal{C}$ die kanonische Basis und sei $\langle\, ,\, \rangle$ das Standard-Skalarprodukt. Dann $M_{\mathcal{C}}(\langle\, ,\, \rangle) = I_n$. Im Allgemein, wenn $A \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit ist, dann $M_{\mathcal{C}}(\langle\, ,\, \rangle_A) = A$.

Lemma 6.1.12. *1. Sei $\mathcal{C}_n$ die kanonische Basis von $\mathbb{R}^n$ und sei $\Phi_{\mathcal{C}}^{\mathcal{B}}\colon V \to \mathbb{R}^n$ das Isomorphismus so dass $v_i \mapsto e_i$. Dann ist die Gramsche Matrix $M_{\mathcal{B}}(\langle\,,\,\rangle)$ die einzige Matrix $M \in \mathbb{R}^{n\times n}$ so dass*

$$\langle v, w\rangle = \langle \Phi_{\mathcal{C}}^{\mathcal{B}}(v), \Phi_{\mathcal{C}}^{\mathcal{B}}(w)\rangle_M$$

für alle $v, w \in V$.

2. Die Gramsche Matrix $M_{\mathcal{B}}(\langle\,,\,\rangle)$ is symmetrisch und positiv definit.

3. Sei $\mathcal{B}' = (v_1', \ldots, v_n')$ eine andere Basis von V und sei $N = M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V)$. Dann

$$M_{\mathcal{B}'}(\langle\,,\,\rangle) = N^t \cdot M_{\mathcal{B}}(\langle\,,\,\rangle) \cdot N$$

Beweis. Sei $M = M_{\mathcal{B}}(\langle\,,\,\rangle)$.

1. Seien $v, w \in V$ und $x = \Phi_{\mathcal{C}}^{\mathcal{B}}(v)$ und $y = \Phi_{\mathcal{C}}^{\mathcal{B}}(w)$, so dass $v = \sum_{i=1}^n x_i v_i, w = \sum_{j=1}^n y_j v_j$. Dann

$$\langle v, w\rangle = \langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle v_i, v_j\rangle = x^t M y = \langle x, y\rangle_M.$$

Sei nun $M' \in \mathbb{R}^{n\times n}$ auch eine Matrix so dass $\langle v, w\rangle = \langle \Phi_{\mathcal{C}}^{\mathcal{B}}(v), \Phi_{\mathcal{C}}^{\mathcal{B}}(w)\rangle_M'$ für alle $v, w \in V$. Dann $M_{ij}' = \langle e_i, e_j\rangle_{M'} = \langle v_i, v_j\rangle = M_{ij}$, so dass $M = M'$.

2. Punkt 1 zeigt dass $\langle\,,\,\rangle_M$ ein Skalarprodukt auf $\mathbb{R}^n$ ist (Warum?). Dann ist M symmetrisch und positiv definit.

3. Wir schreiben $N = (a_{ij})$ so dass $v_i' = \sum_{h=1}^n a_{hi} v_h$ und $v_j' = \sum_{k=1}^n a_{kj} v_k$. Dann

$$M_{\mathcal{B}'}(\langle\,,\,\rangle)_{ij} = \langle v_i', v_j'\rangle = \langle \sum_{h=1}^n a_{hi} v_h, \sum_{k=1}^n a_{kj} v_k\rangle = \sum_{h=1}^n \sum_{k=1}^n a_{hi} a_{kj} \langle v_h, v_k\rangle = (N^t M N)_{ij}$$

□

6.1.2 Norm

Definition 6.1.13. Sei V ein $\mathbb{R}$-Vektorraum. Eine Norm ist eine Funktion

$$\|\cdot\|\colon V \to \mathbb{R}$$

so dass

- **Positiv definit**: $\|v\| \geq 0$ für alle $v \in V$ und $\|v\| = 0$ genau dann, wenn $v = 0$.
- **Homogen**: $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$ für alle $v \in V, \lambda \in \mathbb{R}$.
- **Dreiecksungleichung**: $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$.

Beispiel 6.1.14. Auf $\mathbb{R}^n$ haben wir die euklidische Norm

$$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$$

Wir werden später beweisen dass diese eine Norm ist. Dies kann auf die p-Norm für jedes $p \geq 1$ verallgemeinert werden

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

Zum Beispiel

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

Den Beweis, dass die allgemeine p-Norm eine Norm ist, überlassen wir dem Analysis-Kurs. Eine andere Norm ist auch die Unendlich-Norm:

$$\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

Den Beweis dass diese eine Norm ist, überlassen wir dem Analysis-Kurs.

Beispiel 6.1.15. Wir betrachten den Vektorraum $C^0([0,1])$. Für jede $p \geq 1$ gibt es die L^p-Norm

$$\|f\|_p = \left(\int_0^1 |f(x)|^p dx\right)^{\frac{1}{p}}$$

und wir haben auch die L^∞-Norm

$$\|f\|_\infty = \max_{x \in [0,1]} |f(x)|$$

Wir werden später beweisen dass die L^2-Norm eine Norm ist. Die andere verlassen wir dem Analysis-Kurs.

Proposition 6.1.16. *Sei $(V, \langle\, , \,\rangle)$ ein euklidischer Raum. Wir definieren die Abbildung*

$$\|\cdot\| \colon V \to \mathbb{R}, \qquad v \mapsto \sqrt{\langle v, v\rangle}$$

Folgendes gilt

1. *(Cauchy-Schwartz Ungleichung) Für alle $v, w \in V$*

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$$

2. *Die Abbildung $\|\cdot\|$ ist eine Norm. Sie heißt die induzierte Norm vom Skalarprodukt.*

Beweis. 1. Seien $v, w \in V$. Wenn $v = 0$ oder $w = 0$ sind beide Seite der Ungleichung gleich Null. Wenn $v \neq 0, w \neq 0$, dann $\langle v - \lambda w, v - \lambda w\rangle \geq 0$ für alle $\lambda \in \mathbb{R}$, so dass

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w\rangle = \langle v, v - \lambda w\rangle - \lambda\langle w, v - \lambda w\rangle \\ &= \langle v, v\rangle - \lambda\langle v, w\rangle - \lambda\langle w, v\rangle + \lambda^2\langle w, w\rangle \\ &= \langle v, v\rangle - 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle \end{aligned}$$

Insbesondere wenn $\lambda = \frac{\langle v,w\rangle}{\langle w,w\rangle}$, sehen wir dass

$$\begin{aligned} 0 &\leq \langle v,v\rangle - 2\frac{\langle v,w\rangle}{\langle w,w\rangle}\langle v,w\rangle + \frac{\langle v,w\rangle^2}{\langle w,w\rangle^2}\langle w,w\rangle = \langle v,v\rangle - 2\frac{\langle v,w\rangle^2}{\langle w,w\rangle} + \frac{\langle v,w\rangle^2}{\langle w,w\rangle} \\ &= \langle v,v\rangle - \frac{\langle v,w\rangle^2}{\langle w,w\rangle} \end{aligned}$$

Das bedeutet dass

$$\langle v,v\rangle \geq \frac{\langle v,w\rangle^2}{\langle w,w\rangle}$$

und, da $\langle w,w,\rangle > 0$ (Warum?) das ist auch äquivalent zu

$$\langle v,v\rangle\langle w,w\rangle \geq \langle v,w\rangle^2$$

Durch Ziehen der Quadratwurzeln auf beiden Seiten ergibt sich

$$\|v\| \cdot \|w\| \geq |\langle v,w\rangle|$$

2. Da das Skalarprodukt positiv definit ist, sehen wir dass $\|\cdot\|$ auch positiv definit ist. Für alle $v \in V, \lambda \in \mathbb{R}$ gilt dass

$$\|\lambda \cdot v\| = \sqrt{\langle \lambda \cdot v, \lambda \cdot v\rangle} = \sqrt{\lambda^2 \cdot \langle v,v\rangle} = |\lambda| \cdot \sqrt{\langle v,v\rangle} = |\lambda| \cdot \|v\|$$

so dass $\|\cdot\|$ auch homogen ist. Wir zeigen das die Dreiecksungleichung gilt: seien $v,w \in V$. Wir wollen zeigen dass $\|v+w\| \leq \|v\| + \|w\|$. Da beide Seiten positiv sind, ist diese Ungleichung äquivalent zu $\|v+w\|^2 \leq (\|v\|+\|w\|)^2$. Dann

$$\begin{aligned} \|v+w\|^2 \leq (\|v\|+\|w\|)^2 &\iff \langle v+w, v+w\rangle \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\ &\iff \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle \leq \langle v,v\rangle + 2\|v\|\|w\| + \langle w,w\rangle \\ &\iff 2\langle v,w\rangle \leq 2\|v\| \cdot \|w\| \end{aligned}$$

Für diese letzte Ungleichung zeigt die Cauchy-Schwartz Ungleichung dass

$$\langle v,w\rangle \leq |\langle v,w\rangle| \leq \|v\| \cdot \|w\|$$

□

Bemerkung 6.1.17. Wenn wir ab jetzt einen euklidischen Vektorraum $(V, \langle\, ,\, \rangle)$ haben und eine Norm $\|\|$ schreiben, meinen wir immer die durch das Skalarprodukt induzierte Norm.

Beispiel 6.1.18. Das zeigt dass die euklidische Norm $\|\cdot\|_2$ eine Norm auf $\mathbb{R}^n$ ist und dass die L^2-Norm $\|\cdot\|_{L^2}$ eine Norm auf $C^0([0,1])$ eine Norm ist. Sie sind durch Skalarprodukte induzierte Normen.

Bemerkung 6.1.19. Sei $(V, \langle\, ,\, \rangle)$ ein euklidischer Vektorraum. Wir können die Länge eines Vektors v als $\|v\|$ und auch den Abstand zwischen zwei Elementen $v,w \in \mathbb{R}^n$ als $\|v-w\|$ definieren. Wir können auch den Winkel zwischen zwei Vektoren v,w, die nicht Null sind, als den einzigen $\theta \in [0,\pi]$ definieren, so dass $\cos\theta = \frac{\langle v,w\rangle}{\|v\|\|w\|}$. Dann gilt $\langle v,w\rangle = \|v\| \cdot \|w\|\rangle \cdot \cos\theta$. Dies sind genau die üblichen Begriffe für Abstand und Winkel auf der üblichen euklidischen Ebene $\mathbb{R}^2$. Man beachte, dass wir einen Abstand auch nur mit einer Norm definieren können, aber zur Definition des Winkels benötigen wir das Skalarprodukt.

6.1.3 Orthogonalität

Definition 6.1.20 (Orthogonale Vektoren)**.** Sei $(V, \langle\,,\rangle)$ ein euklidischer Vektorraum. Zwei Vektoren $v, w \in V$ sind orthogonal, falls

$$\langle v, w\rangle = 0$$

Bemerkung 6.1.21. Der Nullvektor ist orthogonal zu allen anderen, da das Skalarprodukt in jedem Faktor linear ist: $\langle v, 0\rangle = 0$ für alle $v \in V$. Außerdem sehen wir dass zwei Vektoren v, w, beide nicht Null, orthogonal sind genau dann wenn der Winkel zwischen den zwei $\frac{\pi}{2}$ ist.

Proposition 6.1.22. *Sei $(V, \langle\,,\,\rangle)$ ein euklidischer Vektorraum und seien $v_1, \ldots, v_r \in V, v_i \neq 0$ paarweise orthogonale Vektoren. Dann sind $v_1, \ldots, v_n$ linear unabhängig.*

Beweis. Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ so dass

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

Dann für jedes $j \in \{1, \ldots, n\}$ gilt dass

$$0 = \langle 0, v_j\rangle = \sum_{i=1}^{n} \lambda_j \langle v_i, v_j\rangle = \lambda_j \cdot \langle v_j, v_j\rangle$$

Da $v_j \neq 0$, sehen wir dass $\langle v_j, v_j\rangle \neq 0$ so dass $\lambda_j = 0$. Das zeigt dass $\lambda_1 = \cdots = \lambda_n = 0$ und das bedeutet dass die $v_1, \ldots, v_n$ linear unabhängig sind. □

Definition 6.1.23 (Orthogonale Basis)**.** Sei $(V, \langle\,,\,\rangle)$ ein endlich-dimensionaler euklidischer Vektorraum. Eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ heißt orthogonal falls die Vektoren paarweise orthogonal sind:

$$\langle v_i, v_j\rangle = 0 \qquad \text{falls } i \neq j$$

Die Basis $\mathcal{B}$ heißt orthonormal wenn sie orthogonal ist und wenn alle Vektoren Norm 1 haben: $\|v_i\| = 1$.

Beispiel 6.1.24. Die kanonische Basis $\mathcal{C} = (e_1, \ldots, e_n)$ ist eine orthonornmale Basis für das Standardskalarprodukt von $\mathbb{R}^n$.

Bemerkung 6.1.25. Eine Basis $\mathcal{B}$ ist orthogonal, genau dann, wenn die Gramsche Matrix $M_{\mathcal{B}}(\langle\,,\rangle)$ diagonal ist. Eine Basis $\mathcal{B}$ ist orthonormal, genau dann, wenn $M_{\mathcal{B}}(\langle\,,\rangle)$ die Identitätsmatrix ist.

Bemerkung 6.1.26. Sei $(V, \langle\rangle)$ ein euklidischer Vektorraum und seien $v_1, \ldots, v_n \in V$ Vektoren so dass

$$\langle v_i, v_j\rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Dann ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine orthonormale Basis: die Vektoren sind alle nicht Null, da sie Norm 1 haben, und sie sind auch paarweise orthogonal. Da $\dim V = n$, sind diese Vektoren eine Basis.

6.1.4 Das Gram-Schmidtsches Verfahren

Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Das Gram-Schmitdsches Verfahren liefert eine orthogonale Basis von V wie folgt. Wir definieren:

$$\begin{aligned}
v_1' &:= v_1 \\
v_2' &:= v_2 - \frac{\langle v_2, v_1'\rangle}{\langle v_1', v_1'\rangle} v_1' \\
v_3' &:= v_3 - \frac{\langle v_3, v_1'\rangle}{\langle v_1', v_1'\rangle} v_1' - \frac{\langle v_3, v_2'\rangle}{\langle v_2', v_2'\rangle} v_2' \\
&\ \vdots \\
v_k' &:= v_k - \frac{\langle v_k, v_1'\rangle}{\langle v_1', v_1'\rangle} v_1' - \frac{\langle v_k, v_2'\rangle}{\langle v_2', v_2'\rangle} v_2' \ \ldots\ - \frac{\langle v_k, v_{k-1}'\rangle}{\langle v_{k-1}', v_{k-1}'\rangle} v_{k-1}' \\
&\ \vdots \\
v_n' &:= v_n - \frac{\langle v_n, v_1'\rangle}{\langle v_1', v_1'\rangle} v_1' - \frac{\langle v_n, v_2'\rangle}{\langle v_2', v_2'\rangle} v_2' \ \ldots\ \ldots\ - \frac{\langle v_n, v_{n-1}'\rangle}{\langle v_{n-1}', v_{n-1}'\rangle} v_{n-1}'
\end{aligned}$$

Satz 6.1.27 (Gram-Schmidt-Verfahren)**.** *Das Gram-Schmidt-Verfahren ist wohldefiniert und* $(v_1', \ldots, v_k')$ *ist eine orthogonale Basis des Unterverktorraums* $\langle v_1, \ldots, v_k\rangle$.

Beweis. Die Tatsache, dass der Algorithmus wohldefiniert ist, bedeutet, dass $\langle v_i', v_i'\rangle \neq 0$ für alle $2 \leq i \leq n-1$, da die Division durch Null die einzige verbotene Operation im Algorithmus ist. Da $\langle\ ,\ \rangle$ positiv definit ist, bedeutet dies, dass $v_i' \neq 0$ für alle $i \geq 2$.

Wir beweisen die Aussage durch Induktion auf n. Wenn $n = 1$ dann brauchen wir nicht zu beweisen, dass der Algorithmus wohldefiniert ist, und $(v_1') = (v_1)$ ist eine orthogonale Basis von $\langle v_1\rangle$. Wenn $n > 1$, nehmen wir an dass die Aussage für $n-1$ wahr ist. Wenn wir die Induktionshypothese auf die Basis $\mathcal{B}_W = (v_1, \ldots, v_{n-1})$ des Vektorraums $W = \langle v_1, \ldots, v_{n-1}\rangle$ anwenden (Warum ist $\mathcal{B}_W$ eine Basis von W?), sehen wir, dass $(v_1', \ldots, v_k')$ eine orthogonale Basis von $\langle v_1, \ldots, v_k\rangle$ für alle $1 \leq k \leq n-1$ ist. Insbesondere $v_i' \neq 0$ für alle $i \in \{1, \ldots, n-1\}$ so dass

$$v_n' := v_n - \frac{\langle v_n, v_1'\rangle}{\langle v_1', v_1'\rangle} v_1' - \frac{\langle v_n, v_2'\rangle}{\langle v_2', v_2'\rangle} v_2' \ \ldots\ \ldots\ - \frac{\langle v_n, v_{n-1}'\rangle}{\langle v_{n-1}', v_{n-1}'\rangle} v_{n-1}'$$

wohldefiniert ist. Wir zeigen dass $v_n \neq 0$. Nehmen wir an, dass $v_n = 0$. Dann $v_n \in \langle v_1', \ldots, v_n'\rangle = \langle v_1, \ldots, v_{n-1}\rangle$, und das ist unmöglich, weil $(v_1, \ldots, v_{n-1}, v_n)$ linear unabhängig ist. Wir zeigen nun, dass $\langle v_n', v_k'\rangle = 0$ für alle $k \leq n-1$. Die Induktionsannahme zeigt dass $\langle v_h', v_k'\rangle = 0$ für alle $h, k \in \{1, \ldots, n-1\}$ mit $h \neq k$. Dann für $1 \leq k \leq n-1$ gilt dass

$$\langle v_n', v_k'\rangle = \langle v_n, v_k'\rangle - \frac{\langle v_n, v_k'\rangle}{\langle v_k', v_k'\rangle}\langle v_k', v_k'\rangle = \langle v_n, v_k'\rangle - \langle v_n, v_k'\rangle = 0$$

Das zeigt dass die $v_1', \ldots, v_n'$ paarweise orthogonal sind und alle nicht Null sind. Insbesondere sind sie auch linear unabhängig und deswegen eine Basis des Vektorraums $\langle v_1, \ldots, v_n\rangle$. $\square$

Korollar 6.1.28. *Sei* $(V, \langle\ ,\ \rangle)$ *ein endlich-dimensionaler euklidischer Vektorraum. Dann hat* V *eine orthonormale Basis.*

Beweis. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Das Gram-Schmidt-Verfahren liefert eine orthogonale Basis $\mathcal{B}' = (v_1', \ldots, v_n')$ von $V = \langle v_1, \ldots, v_n \rangle$. Dann ist

$$\mathcal{B}'' = \left(\frac{1}{\|v_1'\|} v_1', \ldots, \frac{1}{\|v_n'\|} v_n' \right)$$

eine orthonormale Basis. □

Beispiel 6.1.29. Wir anwenden das Gram-Schmidt-Verfahren für die Basis

$$\mathcal{B} = (v_1, v_2, v_3) = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix} \right)$$

des Vektorraums $\mathbb{R}^3$ mit dem Standardskalarprodukt. Dann

$$v_1' = v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Wir sehen dass $\langle v_2, v_1' \rangle = 2$ und dass $\langle v_1', v_1' \rangle = 2$ so dass

$$v_2' = v_2 - \frac{2}{2} v_1' = v_2 - v_1' = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Endlich, sehen wir dass $\langle v_3, v_1' \rangle = 4, \langle v_3, v_2' \rangle = 0, \langle v_2', v_2' \rangle = 1$ so dass

$$v_3' = v_3 - \frac{4}{2} v_1' - \frac{0}{2} v_2' = v_3 - 2v_1' = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}$$

Das Gram-Schmidt-Verfahren liefert die orthonormale Basis

$$(v_1', v_2', v_3') = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix} \right)$$

von $\mathbb{R}^3$. Die Normen dieser Vektoren sind $\|v_1\| = \sqrt{2}, \|v_2\| = 1, \|v_3\| = \sqrt{8}$ so dass

$$\left(\frac{1}{\sqrt{2}} v_1', v_2', \frac{1}{\sqrt{8}} v_3' \right) = \left(\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right)$$

eine orthonormale Basis ist.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 2.7.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Korollar 6.1.30. *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und positiv definit. Dann existiert $N \in GL_n(\mathbb{R})$ so dass*

$$N^t A N = I_n$$

Beweis. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine orthonormale Basis fürs Skalarprodukt $\langle\, , \rangle_A$ auf $\mathbb{R}^n$, sei $\mathcal{C}$ die kanonische Basis von $\mathbb{R}^n$ und sei $N = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_V)$. Dann $N \in GL_n(\mathbb{R})$ und

$$I_n = M_{\mathcal{B}}(\langle\, , \rangle_A) = N^t M_{\mathcal{C}}(\langle\, , \rangle_A) N = NAN^t.$$

□

6.1.5 Orthogonale Untervektorraum

Definition 6.1.31. Sei $(V, \langle\, , \rangle)$ ein euklidischer Vektorraum. Für $u \in V$ definieren wir

$$u^\perp = \{v \in V \mid \langle v, u\rangle = 0\}$$

Für ein Untervektorraum $U \subseteq V$, definieren wir

$$U^\perp = \{v \in V \mid \langle v, u\rangle \ = 0 \text{ für alle } u \in U\}$$

Lemma 6.1.32. *1. Seu $u \in V$ ein Vektor. Dann ist $u^\perp \subseteq V$ ein Untervektorraum.*

2. Sei $U \subseteq V$ ein Untervektorraum. Dann $U = \cap_{u\in U} u^\perp$. Insbesondere ist U ein Untervektorraum.

3. Wenn $U = \langle u_1, \ldots, u_n\rangle$ endlich erzeugt ist, dann

$$U^\perp = \bigcap_{i=1}^{n} u_i^\perp = \{v \in V \mid \langle v, u_i\rangle = 0 \textit{ für alle } i\}$$

Beweis. 1. Wir betrachten die Abbildung $\phi_u\colon V \to \mathbb{R}, v \mapsto \langle v, u\rangle$. Diese Abbildung ist linear weil $\langle\, , \rangle$ linear in der ersten Komponente ist, und $u^\perp = \operatorname{Ker} \phi_u$. Insbesondere ist $u^\perp$ ein Untervektorraum.

2. Die Definition von $U^\perp$ zeigt dass $U = \cap_{u\in U} u^\perp$. Da ein Schnitt von beliebige viele Untervektorräumen wiederum ein Untervektorraum ist, ist $U^\perp$ ein Untervektorraum.

3. Punkt 2 zeigt dass $U^\perp \subseteq \cap_{i=1}^n u_i^\perp$. Fir die andere Inklusion sei $v \in \cap_{i=1}^n u_i^\perp$ und sei $u \in U$. Dann $u = \lambda_1 u_1 + \cdots + \lambda_n u_n$ für $\lambda_i \in \mathbb{R}$, so dass

$$\langle u, v\rangle = \langle \sum_{i=1} \lambda_i u_i, v\rangle = \sum_{i=1}^{n} \lambda_i \langle u_i, v\rangle = \sum_{i=1}^{n} \lambda_i \cdot 0 = 0$$

Das zeigt dass $v \in U^\perp$.

$\square$

Proposition 6.1.33 (Orthogonales Komplement)**.** *Sei $(V, \langle\, , \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum und sei $U \subseteq V$ ein Untervektorraum. Dann*

$$V = U \oplus U^\perp$$

Insbesondere $\dim U^\perp = \dim V - \dim U$.

Beweis. Wir zeigen zuerst dass $U, U^\perp$ in direkte Summe sind: wenn $u \in U \cap U^\perp$, dann $\langle u, u\rangle = 0$, so dass $u = 0$. Wir müssen nun zeigen dass $V = U \oplus U^\perp$, und das bedeutet $\dim U + \dim U^\perp \geq \dim V$, oder $\dim U^\perp \geq \dim V - \dim U$.

Sei $\mathcal{B}_U = (v_1, \ldots, v_r)$ eine Basis von U. Wir können diese Basis zu einer Basis $\mathcal{B} = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ von V ergänzen. Das Gram-Schmidt-Verfahren liefert eine orthogonale Basis $\mathcal{B}' = (v'_1, \ldots, v'_r, v'_{r+1}, \ldots, v'_n)$ so dass $(v'_1, \ldots, v'_n)$ eine orthogonale Basis von $\langle v_1, \ldots, v_r\rangle = U$ ist. Da $\mathcal{B}'$ eine orthogonale Basis ist, das vorheriges Lemma zeigt dass $v'_i \in U^\perp$ für alle $i = r + 1, \ldots, n$, so dass $U^\perp \supseteq \langle v'_{r+1}, \ldots, v'_n\rangle$. Da diese Vektoren linear unabhängig sind, sehen wir dass $\dim U^\perp \geq n - r = \dim V - \dim U$. $\square$

Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum und sei $U \subseteq V$ ein Untervektorraum. Wir wissen dass $V = U \oplus U^\perp$ so dass jeder $v \in V$ eine Summe

$$v = u + u', \qquad u \in U, u' \in U^\perp$$

ist, mit u, u' eindeutig bestimmt. Wir nennen u, u' die orthogonale Projektionen von v auf $U, U^\perp$ und wir schreiben $u = \mathrm{pr}_U(v), u' = \mathrm{pr}_{U^\perp}(v)$.

Proposition 6.1.34. *1. Sei $\mathcal{B}_U = (u_1, \dots, u_r)$ eine orthogonale Basis von U. Dann*

$$\begin{aligned} \mathrm{pr}_U(v) &= \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \cdots + \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 \\ \mathrm{pr}_{U^\perp}(v) &= v - \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 - \cdots - \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 \end{aligned}$$

2. Es gilt dass

$$\|v - \mathrm{pr}_U(v)\| = \min_{\widetilde{u} \in U}\{\|v - \widetilde{u}\|\}$$

und $\mathrm{pr}_U(v)$ ist der einziger Vektor in U mit dieser Eigenschaft.

Beweis. 1. Sei $u = \frac{\langle v,u_1\rangle}{\langle u_1,u_1\rangle}u_1 + \cdots + \frac{\langle v,u_1\rangle}{\langle u_1,u_1\rangle}u_1$ und sei $u' = v - u$. Es ist klar dass $u \in U$, da u eine Linearkombination von $u_1, \dots, u_n$ ist. Da die $u_1, \dots, u_n$ paarweise orthogonal sind, sehen wir dass

$$\langle u', u_i\rangle = \langle v, u_i\rangle - \sum_{j=1}^{n} \frac{\langle v, u_j\rangle}{\langle u_j, u_j\rangle}\langle u_i, u_j\rangle = \langle v, u_i\rangle - \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}\langle u_i, u_i\rangle = \langle v, u_i\rangle - \langle v, u_i\rangle = 0 \qquad \text{für alle } i$$

Das zeigt dass $u' \in U^\perp$ und $v = u + u'$ so dass $u = \mathrm{pr}_U(v), u' = \mathrm{pr}_{U^\perp}(v)$.

2. Seien $u = \mathrm{pr}_U(v)$ und $u' = \mathrm{pr}_{U^\perp}(v) = v - u$. Sei auch $\widetilde{u} \in U$. Wir müssen zeigen dass $\|v - \widetilde{u}\| \geq \|v - u\|$, und um dies zu beweisen, können wir beide Seiten quadrieren: dann

$$\begin{aligned} \|v - \widetilde{u}\|^2 &= \|v - u + u - \widetilde{u}\|^2 = \|u' + u - \widetilde{u}\|^2 \\ &= \|u'\|^2 + 2\langle u', u - \widetilde{u}\rangle + \|u - \widetilde{u}\|^2 = \|u'\|^2 + \|u - \widetilde{u}\|^2 \end{aligned}$$

Wir haben verwenden dass $\langle u', u - \widetilde{u}\rangle = 0$, weil $u' \in U^\perp, u - \widetilde{u} \in U$. Dann

$$\|v - \widetilde{u}\|^2 = \|u'\|^2 + \|u - \widetilde{u}\|^2 \geq \|u'\|^2$$

und die Gleichung gilt, genau dann, wenn $\|u - \widetilde{u}\|^2 = 0$, was bedeutet $\widetilde{u} = u$. Das ist was wir zeigen wollten.

□

Bemerkung 6.1.35. Wenn wir wieder das Gram-Schmitd-Verfahren betrachten und von einer Basis $\mathcal{B} = (v_1, \dots, v_n)$ ausgehen, sehen wir, dass wir bei jedem Schritt die Projektion von v_k auf den Raum $\langle v_1, \dots, v_{k-1}\rangle^\perp$ vornehmen.

Bemerkung 6.1.36. Im zweiten Punkt des vorigen Lemmas haben wir nur verwendet, dass der Vektorraum V sich als direkte Summe $V = U \oplus U^\perp$ zerlegt, so dass das Ergebnis immer gilt, wenn diese Zerlegung gilt, auch wenn der Raum V unendlich dimensional sein könnte.

6.1.6 Orthogonale Abbildungen und Matrizen

Definition 6.1.37 (Orthogonale Abbildung)**.** Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum, eine lineare Abbildung $f\colon V \to V$ ist orthogonal wenn

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \qquad \text{für alle } v, w \in V$$

Beispiel 6.1.38. Die Abbildungen id_V und $-\mathrm{id}_V$ sind orthogonal.

Lemma 6.1.39. *Sei $f\colon V \to V$ eine orthogonale Abbildung.*

1. *f erhält Normen von Vektoren, Abstände zwischen Vektoren und Winkeln zwischen Vektoren.*
2. *$\langle f(v), f(w)\rangle = 0 \iff \langle v, w\rangle = 0$.*
3. *wenn $g\colon V \to V$ auch orthogonal ist, dann ist $f \circ g$ orthogonal.*
4. *f is injektiv.*
5. *wenn V endlich dimensionaler ist, dann ist f ein Isomorphismus und die inverse Abbildung $f\colon W \to V$ ist auch orthogonal.*

Beweis.
1. Wir müssen zeigen dass $\|f(v)\| = \|v\|$, $\|f(v)-f(w)\| = \|v-w\|$ und $\frac{\langle f(v), f(w)\rangle}{\|f(v)\|\|f(w)\|} = \frac{\langle v,w\rangle}{\|v\|\|w\|}$. Alles ist eine direkte Konsequenz der Definition, z.B.:
$$\|f(v)\| = \sqrt{\langle f(v), f(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|.$$
2. Direkte Konsequenz der Definition.
3. Seien $v, w \in V$. Dann $\langle (f\circ g)(v), (f \circ g)(w)\rangle = \langle f(g(v)), f(g(w))\rangle = \langle g(v), g(w)\rangle = \langle v, w\rangle$.
4. Sei $v \in \operatorname{Ker} f$. Dann $\|x\| = \|f(x)\| = \|0\| = 0$, so dass $x = 0$.
5. Wenn V endlich-dimensionaler ist, dann ist jede injektive lineare Abbildung $f\colon V \to V$ ein Isomorphismus. Sei $f^{-1}\colon V \to v$ die inverse Abbildung. Dann $\langle f^{-1}(v), f^{-1}(w)\rangle = \langle f(f^{-1}(v)), f(f^{-1}(w))\rangle = \langle v, w\rangle$ für alle $v, w \in V$. Das zeigt dass f^{-1} orthogonal ist.

□

Lemma 6.1.40. *Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine orthonormale Basis von V. Ein Endomorphismus $f\colon V \to V$ ist orthogonal, genau dann, wenn $(f(v_1), \ldots, f(v_n))$ eine orthonormale Basis von V ist.*

Beweis. Wenn f orthogonal ist, dann ist f ein Isomorphismus so dass $(f(v_1), \ldots, f(v_n))$ eine Basis von V ist. Außerdem $\langle f(v_i), f(v_j)\rangle = \langle v_i, v_j\rangle = 0$ falls $i \neq j$ und $\langle f(v_i), f(v_i)\rangle = \langle v_i, v_i\rangle = 1$. Andererseits, wenn $(f(v_1), \ldots, f(v_n))$ eine orthonormale Basis ist, bedeutet dass, dass $\langle f(v_i), f(v_j)\rangle = \langle v_i, v_j\rangle$ für alle i, j. Dann seien $v = \sum_{i=1}^n x_i v_i$ und $w = \sum_{i=1}^n y_i v_i$ zwei beliebige Vektoren in V, dass wir als Linearkombinationen von Vektoren in $\mathcal{B}$ geschrieben haben. Dann

$$\langle f(v), f(w)\rangle = \langle \sum_{i=1}^n x_i f(v_i), \sum_{j=1}^n y_j f(v_j)\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle v_i, v_j\rangle = \langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\rangle = \langle v, w\rangle$$

□

Definition 6.1.41 (Orthogonale matrix)**.** Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt orthogonal, falls

$$A^t A = I_n$$

Die Menge $O_n(\mathbb{R}) = \{A \in \mathbb{R}^{n\times n} \mid A \text{ orthogonal } \}$ heißt die orthogonale Gruppe

Beispiel 6.1.42. Die Identitätsmatrix I_n ist orthogonal, und die Matrix $-I_n$ ist auch orthogonal.

Beispiel 6.1.43. Sei $\theta \in [0, 2\pi)$ die Matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

ist orthogonal:

$$\begin{aligned} R_\theta^t R_\theta &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta \\ -\cos\theta\sin\theta + \sin\theta\cos\theta & \sin^2\theta + \cos^2\theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \end{aligned}$$

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 9.7.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Lemma 6.1.44. *Sei $A \in \mathbb{R}^{n\times n}$ eine Matrix. Wir betrachten auch $\mathbb{R}^n$ mit dem Standardskalarprodukt $\langle\, ,\, \rangle$. Die folgende Aussäge sind äquivalent:*

1. *A is orthogonal.*
2. *A^t ist orthogonal.*
3. *A ist invertierbar und $A^{-1} = A^t$.*
4. *Die Spalten von A sind eine orthonormale Basis von $\mathbb{R}^n$.*
5. *Die Zeilen von A sind eine orthonormale Basis von $\mathbb{R}^n$.*
6. *Die lineare Abbildung $L_A \colon \mathbb{R}^n \to \mathbb{R}^n$ ist orthogonal.*

Beweis. (1) $\iff$ (3): A ist orthogonal, genau dann, wenn $A^t A = I$ und das bedeutet dass A invertierbar ist und dass $A^{-1} = A^t$.

(2) $\iff$ (3): A^t ist orthogonal, genau dann, wenn $AA^t = I$ und das bedeutet dass A invertierbar ist und dass $A^{-1} = A^t$.

(1) $\iff$ (4): wir sehen dass $(A^t A)_{ij} = (A^i)^t A^j = \langle A^i, A^j\rangle$ so dass wenn $A^t A = I$, dann sind die Spaltenvektoren $(A^1, \dots, A^n)$ paarweise orthogonal und mit Norm $\|A^i\| = 1$. Insbesondere sind die Spaltenvektoren linear unabhängig und da $\dim \mathbb{R}^n = n$ sind sie auch eine Basis von $\mathbb{R}^n$. Andererseits, wenn die Spaltenvektoren eine orthonormale Basis sind, dann $\langle A^i, A^j\rangle = 0$ falls $i \neq j$ und $\langle A^i, A^i\rangle = 1$, so dass $A^t A = I$.

(2) $\iff$ (5): wir sehen dass $(AA^t)_{ij} = (A_i) \cdot A^t_j = \langle A^t_i, A^t_j\rangle$ und Dann können wir den gleichen Beweis wie für die vorherige Äquivalenz führen.

(4) $\iff$ (6): die kanonische Basis $\mathcal{C} = (e_1, \dots, e_n)$ ist orthonormal für das Standardskalarprodukt. Dann zeigt das vorheriges Lemma dass L_A orthogonal ist, genau dann wenn $(L_A(e_1), \dots, L_A(e_n)) = (A^1, \dots, A^n)$ eine orthonormale Basis ist. $\square$

Korollar 6.1.45. *Seien$A, B \in O_n(\mathbb{R})$ zwei orthogonale Matrizen. Dann ist $AB \in O_n(\mathbb{R})$ auch orthogonal.*

Beweis. Wir sehen dass $(AB)^t(AB) = B^t A^t AB = B^t \cdot I_n \cdot B = B^t B = I_n$. Wir können auch feststellen dass $L_{AB} = L_A \circ L_B$ und dass die Komposition von orthogonale Endomorphismen noch orthogonal ist. □

Proposition 6.1.46. *Sei $(V, \langle\, , \rangle)$ ein endlich dimensionaler euklidischer Vektorraum, sei $f : V \to V$ ein Endomorphismus und sei $\mathcal{B}$ eine orthonormale Basis von V. Das Endomorphismus $f : V \to V$ ist orthogonal, genau dann, wenn die Matrix $M^{\mathcal{B}}_{\mathcal{B}}(f)$ orthogonal ist.*

Beweis. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ und sei $\Phi^{\mathcal{B}}_{\mathcal{C}} : V \to \mathbb{R}^n$ das Isomorphismus so dass $\Phi^{\mathcal{B}}_{\mathcal{C}}(v_i) = e_i$. Da die Basis $\mathcal{B}$ orthonormal ist, ist die Gramsche matrix $M_{\mathcal{B}}(\langle\, , \rangle) = I_n$ die Identitätsmatrix. Das bedeutet dass $\langle v, w\rangle = \langle \Phi^{\mathcal{B}}_{\mathcal{C}}(v), \Phi^{\mathcal{B}}_{\mathcal{C}}(w)\rangle_{I_n}$ wobei $\langle\, , \rangle_{I_n}$ das Standardskalarprodukt auf $\mathbb{R}^n$ ist. Insbesondere, sei $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$: dann $\Phi^{\mathcal{B}}_{\mathcal{C}}(f(v_i)) = A^i$, und dann

$$\langle f(v_i), f(v_j)\rangle = \langle A^i, A^j\rangle_{I_n}$$

Wir wissen dass f orthogonal ist, genau dann, wenn $(f(v_1), \ldots, f(v_n))$ eine orthogonale Basis ist. Die vorherige Gleichung zeigt dass $(f(v_1), \ldots, f(v_n)$ eine orthonormale Basis ist, genau dann, wenn $(A^1, \ldots, A^n)$ eine orthonormale Basis ist, und das bedeutet dass A orthogonal ist. □

6.2 Unitäre Räume

Definition 6.2.1 (Skalarprodukte für komplexe Vektorräume)**.** Sei V ein $\mathbb{C}$-Vektorraum. Ein Skalarprodukt auf V ist eine Abbildung

$$\langle\, , \rangle : V \times V \to \mathbb{C}$$

mit folgenden Eigenschaften:

- **Sesquilinear**: für alle $v, v', w, w' \in V$ und $\lambda, \lambda' \in \mathbb{R}$ gilt dass

$$\langle \lambda v + \lambda' v', w\rangle = \overline{\lambda}\langle v, w\rangle + \overline{\lambda'}\langle v', w\rangle$$
$$\langle v, \mu w + \mu' w'\rangle = \mu\langle v, w\rangle + \mu'\langle v, w'\rangle$$

- **Hermitesch**: für alle $v, w \in V$ gilt dass

$$\langle v, w\rangle = \overline{\langle w, v\rangle}.$$

- **Positiv definit**: für alle $v \in V$ gilt dass

$$\langle v, v\rangle \geq 0, \quad \text{und } \langle v, v\rangle = 0 \iff v = 0$$

Ein $\mathbb{C}$-Vektorraum V mit einem Skalarprodukt $\langle\, , \rangle$ heißt ein unitärer Raum.

Bemerkung 6.2.2. Wenn $\langle\, , \rangle$ Hermitesch ist, dann $\langle v, v\rangle = \overline{\langle v, v\rangle}$ für alle $v \in V$, so dass $\langle v, v\rangle \in \mathbb{R}$.

Beispiel 6.2.3. Das Standardskalarprodukt auf $\mathbb{C}^n$ ist

$$\langle x, y\rangle = \overline{x}^t \cdot y = \sum_{i=1}^{n} \overline{x}_i \cdot y_i.$$

Beispiel 6.2.4. Wir betrachten den $\mathbb{C}$-Vektorraum $C^0([0,1],\mathbb{C}) = \{f\colon [0,1] \to \mathbb{C} \mid f \text{ stetig } \}$. Diese ist ein unitärer Raum mit dem Skalarprodukt

$$\langle f, g\rangle = \int_0^1 \overline{f(x)} g(x) dx$$

Definition 6.2.5 (Hermitsche Matrix, Positiv definite Matrix)**.** Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt Hermitesch, falls $A = \overline{A}^t$. Eine Hermitesch Matrix $A \in \mathbb{C}^{n\times n}$ ist positiv definit, falls $\overline{x}^t Ax \geq 0$ für alle $x \in \mathbb{R}^n$, und $\overline{x}^t Ax = 0$, genau dann, wenn $x = 0$.

Wir können die Ergebnisse für euklidische Vektorräume zu unitäre Vektorräume verallgemeinern. Wir geben die Beweise nicht, weil sie ähnlich zu den Beweise für euklidische Vektorräume sind.

Lemma 6.2.6. *Sei $A \in \mathbb{C}^{n\times n}$ eine Matrix. Dann ist die Abbildung*

$$\langle\ ,\ \rangle_A \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \qquad (x, y) \mapsto \overline{x}^t Ay$$

ein Skalarprodukt, genau dann, wenn A Hermitesch und positiv definit ist.

Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum. Wir können die Gramsche Matrix von $\langle\ ,\ \rangle$ bezüglich eine Basis $\mathcal{B}$ wie für euklidische Vektorräume definieren. Dann

Lemma 6.2.7. *1. Die Gramsche Matrix $M_\mathcal{B}(\langle\ ,\ \rangle)$ ist Hermitesch und positiv definit.*

2. Die Gramsche Matrix $M_\mathcal{B}(\langle\ ,\ \rangle)$ ist die einzige Matrix $M \in \mathbb{C}^{n\times n}$ so dass

$$\langle v, w\rangle = \langle \Phi^\mathcal{B}_\mathcal{C}(v), \Phi^\mathcal{B}_\mathcal{C}(w)\rangle_M$$

für alle $v, w \in v$.

3. Sei $\mathcal{B}' = (v'_1, \ldots, v'_n)$ eine andere Basis von V und sei $N = M^{\mathcal{B}'}_\mathcal{B}(\mathrm{id}_V)$. Dann

$$M_{\mathcal{B}'}(\langle\ ,\ \rangle) = \overline{N}^t \cdot M_\mathcal{B}(\langle\ ,\ \rangle) \cdot N$$

Wenn $(V, \langle\rangle)$ ein unitärer Vektorraum ist, wir können die Norm $\|v\| = \sqrt{\langle v, v\rangle}$ wie für euklidische Vektorräume definieren. Die Cauchy-Schwarzt Ungleichung gilt auch hier.

Wir können auch Orthogonalität wie für euklidische Vektorräume definieren, und das Gram-Schmidt-Verfahren funktioniert für unitäre Vektorräume auch. Insbesondere, hat jeder endlich-dimensionaler unitärer Vektorraum V eine orthonormale Basis und wenn $U \subseteq V$ ein Untervektorraum ist, dann $V = U \oplus U^\perp$.

Definition 6.2.8 (Unitäre Abbildungen und Matrizen)**.** Sei $(V, \langle\ ,\ \rangle)$ ein unitärer Vektorraum. Ein Endomorphismus $f\colon V \to V$ ist unitär, falls

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \qquad \text{für alle } v, w \in V$$

Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt unitär, falls

$$\overline{A}^t A = I_n$$

Die Menge von unitäre $n \times n$ Matrizen ist die unitäre Gruppe $U_n(\mathbb{C})$.

Alle Aussagen über orthogonale Abbildungen und Matrizen gelten auch für unitäre Abbildungen und Matrizen, nur die transponierte Matrix A^t soll durch die Matrix $\overline{A}^t$ in alle Aussage ersetzt werden.

6.3 Normale und Selbstadjungierte Endomorphismen

Ein Vektorraum mit Skalarprodukt $(V, \langle\, , \rangle)$ is entweder ein euklidischer Vektorraum über $\mathbb{R}$ oder ein hermitescher Vektorraum über $\mathbb{C}$. Wir bezeichnen den Basiskörper mit $\mathbb{K}$, und für jedes $\lambda \in \mathbb{K}$ schreiben wir $\overline{\lambda}$ für das komplex Konjugierte (was über $\mathbb{R}$ keine Wirkung hat).

Lemma 6.3.1. *Sei $(V, \langle\, , \rangle)$ ein Vektorraum mit einem Skalarprodukt und sei $\mathcal{B} = (v_1, \dots, v_n)$ eine orthonormale Basis. Dann*

$$v = \sum_{i=1}^{n} \langle v_i, v \rangle \cdot v_i$$

Beweis. Wir schreiben $v = \sum_{i=1}^{n} \lambda_i v_i$. Dann $\langle v_j, v \rangle = \sum_{i=1}^{n} \lambda_i \langle v_j, v_i \rangle = \lambda_j$ für alle j. □

Definition 6.3.2 (Adjungiertes Endomorphismus)**.** Sei $(V, \langle\rangle)$ ein endlich-dimensionaler Vektorraum mit Skalarprodukt und sei $f\colon V \to V$ ein Endomorphismus. Ein adjungiertes Endomorphismus $f^*\colon V \to V$ ist ein Endomorphismus so dass

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle$$

Definition 6.3.3 (Adjungierte Matrix)**.** Sei $A \in \mathbb{K}^{n\times n}$ eine Matrix. Die adjungierte Matrix ist $\overline{A}^t$. Insbesondere $\overline{A}^t = A^t$ wenn $\mathbb{K} = \mathbb{R}$.

Proposition 6.3.4. *Sei $(V, \langle\, , \rangle)$ ein endlich-dimensionaler Vektorraum mit Skalaprodukt und sei $f\colon V \to V$ ein Endomorphismus.*

1. *Ein adjungiertes Endomorphismus $f^*\colon V \to V$ existiert und ist eindeutig.*

2. *Wenn $\mathcal{B} = (v_1, \dots, v_n)$ eine orthonormale Basis von V ist, dann*

$$f^*(v) = \sum_{i=1}^{n} \langle f(v_i), v \rangle \cdot v_i$$

3. *$(f^*)^* = f$.*

4. *Wenn $g\colon V \to V$ ein Endomorphismus ist, und $\lambda, \mu \in \mathbb{K}$, dann*

$$(\lambda f + \mu g)^* = \overline{\lambda} f^* + \overline{\mu} g^*$$

Sei $A \in \mathbb{K}^{n\times n}$ eine Matrix, und sei $\langle\, , \rangle_{I_n}$ das Standardskalarprodukt auf $\mathbb{K}^n$.

5. *$\langle Ax, y \rangle = \langle x, \overline{A}^t y \rangle$ für alle $x, y \in \mathbb{K}^n$.*

Sei $\mathcal{B}$ eine orthonormale Basis von V.

6. *Wenn $M^{\mathcal{B}}_{\mathcal{B}}(f) = A$, dann $M^{\mathcal{B}}_{\mathcal{B}}(f^*) = \overline{A}^t$.*

Beweis. 1. Wir zeigen zuerst dass f^* eindeutig ist: sei $f^*\colon V \to V$ ein adjungiertes Endomorphismus von f und sei $\mathcal{B} = (v_1, \dots, v_n)$ eine orthonormale Basis von V. Dann

$$f^*(v) = \sum_{i=1}^{n} \langle v_i, f^*(v) \rangle v_i = \sum_{i=1}^{n} \langle f(v_i), v \rangle v_i \qquad \text{für alle } v \in V$$

so dass f^* eindeutig ist. Um zu zeigen, dass f^* existiert, können wir einfach

$$f^*\colon V \to V, \qquad f^*(v) := \sum_{i=1}^{n} \langle f(v_i), v \rangle v_i$$

definieren und dann überprüfen dass f^* ein adjungiertes Endomorphismus ist. Wir können leicht überprüfen dass f^* linear ist: wenn $v, v' \in V$ und $\lambda, \lambda' \in \mathbb{K}$, dann

$$f^*(\lambda v+\lambda' v') = \sum_{i=1}^{n} \langle v_i, f(\lambda v+\lambda' v') \rangle v_i = \lambda \sum_{i=1}^{n} \langle v_i, f(v) \rangle v_i + \lambda' \sum_{i=1}^{n} \langle v_i, f(v') \rangle v_i = \lambda f^*(v) + \lambda' f^*(v').$$

Wir müssen denn überprüfen dass $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$ für alle $v, w \in V$. Sei $v = \sum_{i=1}^{n} x_i v_i$, $w = \sum_{i=1}^{n} y_i v_i$,dann

$$\langle f(v), w \rangle = \langle \sum_{i=1} x_i f(v_i), w \rangle = \sum_{i=1}^{n} \overline{x}_i \langle f(v_i), w \rangle$$

$$\langle v, f^*(w) \rangle = \langle v, \sum_{i=1}^{n} \langle f(v_i), w \rangle v_i \rangle = \sum_{i=1}^{n} \langle v, v_i \rangle \langle f(v_i), w \rangle = \sum_{i=1}^{n} \overline{x}_i \langle f(v_i), w \rangle$$

2. Das haben wir im Beweis des vorheriges Punktes gezeigt.

3. Wir müssen zeigen dass $\langle f^*(v), w \rangle = \langle v, f(w) \rangle$. Das ist äquivalent zu $\overline{\langle f^*(v), w \rangle} = \overline{\langle v, f(w) \rangle}$ und das ist äquivalent zu $\langle w, f^*(v) \rangle = \langle f(w), v \rangle$, was die bestimmende Eigenschaft von $f*$ ist.

4. Wir sehen dass

$$\langle (\lambda f+\mu g)(v), w \rangle = \overline{\lambda} \langle f(v), w \rangle + \overline{\mu} \langle g(v), w \rangle = \overline{\lambda} \langle v, f^*(w) \rangle + \overline{\mu} \langle v, g^*(w) \rangle = \langle v, (\overline{\lambda} f^* + \overline{\mu} g^*)(w) \rangle$$

5. Wir sehen dass

$$\langle Ax, y \rangle_{I_n} = \overline{(Ax)}^t y = \overline{x}^t \overline{A}^t y = \langle x, \overline{A}^t y \rangle_{I_n}$$

6. Wir wissen dass

$$f(v_j) = \sum_{i=1}^{n} \langle v_i, f(v_j) \rangle v_i, \qquad f^*(v_j) = \sum_{i=1}^{n} \langle f(v_i), v_j \rangle v_i$$

so dass $M_{\mathcal{B}}^{\mathcal{B}}(f^*)_{ij} = \langle f(v_i), v_j \rangle = \overline{\langle v_j, f(v_i) \rangle} = \overline{A}_{ji}$.

□

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 12.7.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Wir können eine weitere Interpretation orthogonaler und unitärer Endomorphismen mit dem adjungierten Endomorphismus geben

Lemma 6.3.5. *Sei $(V, \langle\, ,\, \rangle)$ ein Vektorraum mit einem Skalarprodukt. Ein Endomorphismus $f\colon V \to V$ ist orthogonal oder unitär, genau dann, wenn $f^* \circ f = \mathrm{id}_V$.*

Beweis. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine orthogonale Basis von V. Dann ist f orthogonal genau dann, wenn $\langle f(v_i), f(v_j)\rangle = \langle v_i, v_j\rangle$ für alle i, j. Das bedeutet

$$\langle v_i, v_j\rangle = \langle f(v_i), f(v_j)\rangle = \langle v_i, f^*(f(v_j))\rangle$$

und Lemma 6.3.1 zeigt dass dieses äquivalent zu $f^* f(v_j) = v_j$ ist, was äquivalent zu $f^* f = \mathrm{id}_V$ ist. □

Definition 6.3.6 (Normale Endomorphismen und Matrizen)**.** Sei $(V, \langle\, ,\, \rangle)$ ein Vektorraum mit einem Skalarprodukt. Ein Endomorphismus $f\colon V \to V$ ist normal, falls

$$f \circ f^* = f^* \circ f$$

Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt normal, falls

$$A\overline{A}^t = \overline{A}^t A$$

Beispiel 6.3.7. Ein orthogonales oder unitäres Endomorphismus $f\colon V \to V$ ist Normal:

$$f^* \circ f = \mathrm{id}_V = f \circ f^*$$

Eine orthogonale oder unitäre Matrix A ist auch normal:

$$\overline{A}^t A = I_n = A\overline{A}^t$$

Definition 6.3.8 (Selbstadjungierte Endomorphismen und Matrizen)**.** Sei $(V, \langle\, ,\, \rangle)$ ein Vektorraum mit einem Skalarprodukt. Ein Endomorphismus $f\colon V \to V$ ist selbstadjungiert falls

$$f = f^*$$

Beispiel 6.3.9. Die Identität id_V ist selbstadjungiert.

Bemerkung 6.3.10. Sei $\mathcal{B}$ eine orthonormale Basis von V. Dann zeigt die vorherige Proposition dass f selbstadjungiert ist genau dann, wenn $M^{\mathcal{B}}_{\mathcal{B}}(f)$ hermitesch ist.

Beispiel 6.3.11. Ein selbstadjungiertes Endomorphismus $f\colon V \to V$ ist normal:

$$f^* f = f^2 = f f^*$$

Symmetrische oder hermitesche Matrizen sind normal:

$$\overline{A}^t A = A^2 = A\overline{A}^t$$

6.4 Der Spektralsatz für normale Endomorphismen

Proposition 6.4.1. *Sei $(V, \langle\ ,\ \rangle)$ ein Vektorraum mit einem Skalarprodukt und sei $f\colon V \to V$ ein normales Endomorphismus.*

1. $\|f^*(v)\| = \|f(v)\|$ *für alle* $v \in V$.
2. $\operatorname{Ker} f = \operatorname{Ker} f^*$.
3. *Wenn $v \in V$ ein Eigenvektor von f mit Eigenwert λ ist, dann ist v auch ein Eigenvektor von f^* mit Eigenwert $\overline{\lambda}$.*
4. *Eigenvektoren von f mit verschiedene Eigenwerte sind orthogonal.*
5. *Sei $\lambda \in \mathbb{K}$. Dann ist $\operatorname{Ker}(f - \lambda \operatorname{id}_V)^\perp$ invariant für f und f^*:*

$$f(\operatorname{Ker}(f - \lambda \operatorname{id}_V)^\perp) \subseteq \operatorname{Ker}(f - \lambda \operatorname{id}_V)^\perp,\ f^*(\operatorname{Ker}(f - \lambda \operatorname{id}_V)^\perp) \subseteq \operatorname{Ker}(f - \lambda \operatorname{id}_V)^\perp$$

Beweis.

1. Wir sehen dass

$$\|f(v)\|^2 = \langle f(v), f(v)\rangle = \langle v, (f^* \circ f)(v)\rangle = \langle v, (f \circ f^*)(v)\rangle = \langle f^*(v), f^*(v)\rangle = \|f^*(v)\|^2$$

und das Ziehen der Quadratwurzeln aus beiden Seiten schließt den Beweis ab.

2. Wir sehen dass $\|f(v)\| = 0$ genau dann, wenn $\|f^*(v)\| = 0$ und dann wir nutzen dass eine Norm positiv definit ist.

3. Wir sehen dann $(f - \lambda \operatorname{id}_V)^* = f^* - \overline{\lambda} \operatorname{id}_V$. Insbesondere, kann mann leicht zeigen dass $(f - \lambda \operatorname{id}_V)$ normal ist. Dann zeigt Punkt 2 dass $\operatorname{Ker}(f - \lambda \operatorname{id}_V) = \operatorname{Ker}(f^* - \overline{\lambda} \operatorname{id}_V)$, und dass ist genau was wir zeigen wollten.

4. Seien $v, w \in V$ so dass $f(v) = \lambda v, f(w) = \mu w$, mit $\lambda \neq \mu$. Dann

$$\overline{\lambda}\langle v, w\rangle = \langle \lambda v, w\rangle = \langle f(v), w\rangle = \langle v, f^*(w)\rangle = \langle v, \overline{\mu} w\rangle = \overline{\mu}\langle v, w\rangle$$

und da $\overline{\lambda} \neq \overline{\mu}$, es müss sein dass $\langle v, w\rangle = 0$.

5. Sei $v \in \operatorname{Ker}(f - \lambda \operatorname{id}_V)^\perp$. Wir wollen zeigen dass $\langle f(v), w\rangle = \langle f^*(v), w\rangle$ für alle $w \in \operatorname{Ker}(f - \lambda \operatorname{id}_V)$. Wir sehen dass

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle = \langle v, \overline{\lambda} w\rangle = 0$$
$$\langle f^*(v), w\rangle = \langle v, f(w)\rangle = \langle v, \lambda w\rangle = 0$$

□

Satz 6.4.2 (Spektralsatz für Normale Endomorphismen)**.** *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler Vektorraum mit einem Skalarprodukt und sei $f\colon V \to V$ ein Endomorphismus so dass*

$$\chi_f(t) = a_n(t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_n} \qquad \text{in } \mathbb{K}[t]$$

in Linearfaktoren zerfallt, und $a_n \in \mathbb{K}, a_n \neq 0$.

Dann existiert eine orthonormale Basis von Eigenvektoren von V genau dann, wenn f normal ist.

Bevor wir diesen Satz beweisen, geben wir eine Neuinterpretation eines unserer Kriterien für die Diagonalisierbarkeit

Lemma 6.4.3. *Sei V ein endlicher dimensionaler $\mathbb{K}$-Vektorraum und sei $f\colon V \to V$ ein Endomorphismus so dass*

$$\chi_f(t) = a_n(t-\lambda_1)^{m_1} \cdot \dots \cdot (t-\lambda_r)^{m_n} \qquad \text{in } \mathbb{K}[t]$$

in Linearfaktoren zerfallt, mit $\lambda_1, \dots, \lambda_r$ paarweise verschiedene, und $a_n \in \mathbb{K}, a_n \neq 0$. Das Endomorphismus f ist diagonalisierbar, genau dann, wenn für jedes Eigenwert $\lambda_i \in \mathbb{K}$ ein Komplement W_i von $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ existiert, so dass

$$V = \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \oplus W_i$$

und $f(W_i) \subseteq W_i$.

Beweis. Wenn f diagonalisierbar ist, setzen wir W_i als die direkte Summe allen Eigenräume außer $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$:

$$W_i = \operatorname{Ker}(f - \lambda_1 \operatorname{id}_V) \oplus \dots \oplus \widehat{\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)} \oplus \dots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_V)$$

Dann $V = \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \oplus W_i$ und $f(W_i) \subseteq W_i$.

Andererseits, nehmen wir an dass wir eine solche direkte Summe haben, mit $f(W_i) \subseteq W_i$. Wir wollen zeigen dass die geometrische Vielfachheit $e_i = \dim \operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ und die algebraische Vielfachheit m_i gleich sind. Sei $\mathcal{B}' = (v_1, \dots, v_{e_i})$ eine Basis von $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ und sei $\mathcal{B}'' = (v_{e_i+1}, \dots, v_n)$ eine Basis von W_i. Dann $\mathcal{B} = (v_1, \dots, v_n)$ ist eine Basis von V, und

$$M^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} M^{\mathcal{B}'}_{\mathcal{B}'}(f_{|\operatorname{Ker}(f-\lambda_i \operatorname{id}_V)}) & 0 \\ 0 & M^{\mathcal{B}''}_{\mathcal{B}''}(f_{W_i}) \end{pmatrix} = \begin{pmatrix} \lambda_i I_{e_i} & 0 \\ 0 & M^{\mathcal{B}''}_{\mathcal{B}''}(f_{W_i}) \end{pmatrix}$$

so dass $\chi_f(t) = (-1)^{e_i}(t-\lambda_i)^{e_i} \cdot \chi_{f_{|W_i}}(t)$ Wir bemerken dass $\chi_{f_{|W_i}}(\lambda_i) \neq 0$, sonst existiert ein Eigenvektor $v \in \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \cap W_i$ und das ist unmöglich. Dann ist e_i genau die algebraische Vielfachheit: $e_i = m_i$. □

Wir beweisen nun den Spektralsatz:

Beweis des Spektralsatzes für normale Endomorphismen. Wir zeigen zuerst dass wenn f normal ist, dann existiert eine orthonormale Basis von Eigenvektoren. Wir zeigen zuerst dass f diagonalisierbar ist: sei λ_i ein Eigenwert von f. Dann

$$V = \operatorname{Ker}(f - \lambda_i \operatorname{id}_V) \oplus \operatorname{Ker}(f - \lambda_i \operatorname{id}_V)^{\perp}$$

und die vorherige Proposition zeigt dass $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)^{\perp}$ invariant für f ist. Dann zeigt das vorheriges Lemma dass f diagonalisierbar ist, so dass

$$V = \operatorname{Ker}(f - \lambda_1 \operatorname{id}_V) \oplus \dots \oplus \operatorname{Ker}(f - \lambda_r \operatorname{id}_V)$$

Wir können dann orthonormale Basen $\mathcal{B}_i$ von $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ nehmen, und dann ist $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_n$ eine orthonormale Basis von Eigenvektoren von f, weil Eigenvektoren mit verschiedene Eigenwerten orthogonal sind.

Andererseits, nehmen wir an, dass eine orthonormale Basis $\mathcal{B} = (v_1, \dots, v_n)$ von Eigenvektoren von V existiert: $f(v_i) = \lambda_i \cdot v_i$. Dann

$$(f \circ f^*)(v_i) = f(\overline{\lambda}_i \cdot v_i) = \overline{\lambda}_i f(v_i) = \overline{\lambda_i}\lambda_i v_i$$
$$(f^* \circ f)(v_i) = f^*(\lambda_i \cdot v_i) = \lambda_i f^*(v_i) = \lambda\overline{\lambda_i}\lambda_i v_i$$

so dass $(f \circ f^*)(v_i) = (f^* \circ f)(v_i)$, und da $\mathcal{B}$ eine Basis ist, bedeutet das, dass $f \circ f^* = f^* \circ f$. □

Bemerkung 6.4.4. Der Beweis zeigt auch, wie man eine orthonormale Basis von Eigenvektoren für f findet. Zuerst findet man eine Basis $\mathcal{B}_i$ des Eigenspace $\operatorname{Ker}(f - \lambda_i \operatorname{id}_V)$ für jeden Eigenwert λ_i für $i = 1, \dots, r$. Dann erhält man eine Orthonormalbasis $\mathcal{B}'_i$ über Gram-Schmidt. Schließlich wird die Vereinigung $\mathcal{B}' = \mathcal{B}'_1 \cup \ldots \cup \mathcal{B}'_r$ ist eine orthonormale Basis von V.

Korollar 6.4.5 (Spektralsatz für normale Matrizen)**.** *Sei $A \in \mathbb{K}^{n \times n}$ eine Matrix so dass*

$$\chi_f(t) = a_n(t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_n} \qquad \text{in } \mathbb{K}[t]$$

in Linearfaktoren zerfallt, mit und $a_n \in \mathbb{K}, a_n \neq 0$.

Dann existiert eine orthogonale (wenn $\mathbb{K} = \mathbb{R}$) oder unitäre (wenn $\mathbb{K} = \mathbb{C}$) Matrix $N \in \mathbb{K}^{n\times n}$ so dass $\overline{N}^t AN = D$ diagonal ist, genau dann, wenn A normal ist.

Beweis. Wir nehmen an dass A normal ist und wir betrachten die lineare Abbildung $L_A \colon \mathbb{K}^n \to \mathbb{K}^n$. Dann ist L_A auch normal, und der Spektralsatz zeigt dass eine orthonormale Basis von Eigenvektoren von A existiert: $\mathcal{B} = (v_1, \dots, v_n)$. Die Matrix $N = M^{\mathcal{B}}_{\mathcal{C}}(\operatorname{id}_{\mathbb{K}^n})$ ist orthogonal (wenn $\mathbb{K} = \mathbb{R}$) oder unitär (wenn $\mathbb{K} = \mathbb{C}$), so dass $N^{-1} = \overline{N}^t$, und

$$\overline{N}^t AN = N^{-1}AN = M^{\mathcal{C}}_{\mathcal{B}}(\operatorname{id}_{\mathbb{K}^n}) \cdot M^{\mathcal{C}}_{\mathcal{C}}(L_A) \cdot M^{\mathcal{B}}_{\mathcal{C}}(\operatorname{id}_{\mathbb{K}^n}) = M^{\mathcal{B}}_{\mathcal{B}}(L_A) = D$$

ist diagonal.

Andererseits, wenn $\overline{N}^t AN = D$, dann $A = ND\overline{N}^t$ und $\overline{A}^t = N\overline{D}\,\overline{N}^t$, so dass

$$A\overline{A}^t = ND\overline{N}^t \cdot N\overline{D}\overline{N}^t = ND\overline{D}\overline{N}^t = N\overline{D}D\overline{N}^t = N\overline{D}\overline{N}^t \cdot ND\overline{N}^t = \overline{A}^t A$$

□

Bemerkung 6.4.6. Wenn $(V, \langle\ ,\ \rangle)$ ein unitärer Vektorraum ist, dann zerfallt $\chi_f(t) \in \mathbb{C}[t]$ in Linearfaktoren, weil $\mathbb{C}$ algebraisch abgeschlossen ist. Dann ist jedes normales Endomorphismus $f \colon V \to V$ durch eine orthonormale Basis von Eigenvektoren diagonalisierbar.

Eine ähnliche Begrundung zeigt dass jede normale Matrix $A \in \mathbb{C}^{n\times n}$ diagonalisierbar durch eine orthonormale Basis von Eigenvektoren ist.

6.5 Der Spektralsatz für selbstadjungierte Endomorphismen

Lemma 6.5.1. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit einem Skalarprodukt und sei $f \colon V \to V$ ein selbstadjungiertes Endomorphismus. Dann zerfallt das charakteristisches Polynom in Linearfaktoren*

$$\chi_f(t) = a_n(t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_r)^{m_r} \ \text{in } \mathbb{K}[t]$$

mit $a_n \in \mathbb{K}, a_n \neq 0$, und außerdem sind alle Eigenwerte reell: $\lambda_i \in \mathbb{R}$ für alle i.

Beweis. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine orthonormale Matrix und sei $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$. Insbesondere ist A selbstadjungiert $A = \overline{A}^t$. Wir wissen dass $\chi_f(t) = \chi_A(t)$. Wenn wir A als komplexe Matrix $A \in \mathbb{C}^{n\times n}$ betrachten, wissen wir dass

$$\chi_A(t) = a_n(t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r} \text{ in } \mathbb{C}[t]$$

da $\mathbb{C}$ algebraisch abgeschlossen ist. Wir zeigen nun dass alle die Eigenwerte reell sind: sei $\lambda \in \mathbb{C}$ ein Eigenwert und sei $v \in \mathbb{C}^n$ ein Eigenvektor so dass $Av = \lambda v$. Wir betrachten das Standardskalarprodukt $\langle\,,\,\rangle_{I_n}$ auf $\mathbb{C}^n$. Dann

$$\overline{\lambda}\langle v,v\rangle_{I_n} = \langle \lambda v, v\rangle_{I_n} = \langle Av, v\rangle_{I_n} = \langle v, \overline{A}^t v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda\langle v,v\rangle$$

Da $\langle v,v\rangle \neq 0$, es muss sein dass $\overline{\lambda} = \lambda$, d.h. $\lambda \in \mathbb{R}$. Man kann außerdem auch leicht beweisen dass $a_n = (-1)^n$ (das gilt für alle $n \times n$ Matrizen), so dass

$$\chi_f(t) = \chi_A(t) = a_n(t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r} \text{ in } \mathbb{R}[t]$$

Das beweist die Außage gleichzeitig für $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$. □

Satz 6.5.2 (Spektralsatz für selbstadjungierte Endomorphismen)**.** *Sei $(V, \langle\,,\,\rangle)$ ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit einem Skalarprodukt und sei $f\colon V \to V$ ein Endomorphismus. Dann existiert eine orthonormale Basis von Eigenvektoren von f mit reelle Eigenwerte, genau dann, wenn f selbstadjungiert ist.*

Beweis. Wenn f selbstadjungiert ist, dann wissen wir dass das charakteristisches Polynom in Linearfaktoren in $\mathbb{K}[t]$ zerfallt. Wir wissen auch das f normal ist, und der Spektralsatz für normale Endomorphismen zeigt dass V eine orthonormale Basis von Eigenvektoren von f. Wir wissen auch dass alle Eigenwerte reell sind.

Andererseits, nehmen wir an dass V eine orthonormale Basis $\mathcal{B} = (v_1, \dots, v_n)$ von Eigenvektoren von f mit reelle Eigenwerte hat: $f(v_i) = \lambda_i v_i, \lambda_i \in \mathbb{R}$. Dann

$$f(v_i) = \lambda_i v_i = \overline{\lambda_i} v_i = f^*(v_i) \quad \text{für alle } i$$

so dass $f = f^*$. □

Korollar 6.5.3 (Spektralsatz für selbstadjungierte Matrizen)**.** *Sei $A \in \mathbb{K}^{n\times n}$ eine Matrix. Dann existiert eine orthogonale (wenn $\mathbb{K} = \mathbb{R}$) oder unitäre (wenn $\mathbb{K} = \mathbb{C}$) Matrix $N \in \mathbb{K}^{n\times n}$ so dass $\overline{N}^t AN = D$ diagonal und reell ist, genau dann, wenn A selbstadjungiert ist.*

Beweis. Wir nehmen an dass A selbstadjungiert ist ist und wir betrachten die lineare Abbildung $L_A\colon \mathbb{K}^n \to \mathbb{K}^n$. Dann ist L_A auch selbstadjungiert, und der Spektralsatz zeigt dass eine orthonormale Basis von Eigenvektoren von A mit reelle Eigenwerte existiert: $\mathcal{B} = (v_1, \dots, v_n)$. Die Matrix $N = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_{\mathbb{K}^n})$ ist orthogonal (wenn $\mathbb{K} = \mathbb{R}$) oder unitär (wenn $\mathbb{K} = \mathbb{C}$), so dass $N^{-1} = \overline{N}^t$ auch orthogonal/unitär ist, und

$$\overline{N}^t AN = N^{-1}AN = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_{\mathbb{K}^n}) \cdot M_{\mathcal{C}}^{\mathcal{C}}(L_A) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{id}_{\mathbb{K}^n}) = M_{\mathcal{B}}^{\mathcal{B}}(L_A) = D$$

ist diagonal mit reelle Eigenwerte.

Andererseits, wenn $\overline{N}^t AN = D$ mit D diagonal und reell, dann $A = ND\overline{N}^t$ und $\overline{A}^t = N\overline{D}\,\overline{N}^t = ND\,\overline{N}^t = A$. □

Insbesondere, wenn $\mathbb{K} = \mathbb{R}$ bekommen wir

Korollar 6.5.4 (Spektralsatz für symmetrische Matrizen)**.** *Sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann existiert eine orthogonale Matrix $N \in O_n(\mathbb{R})$ so dass*

$$N^t AN = D$$

diagonal ist.

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 16.7.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Beispiel 6.5.5. Wir betrachten die symmetrische Matrix

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

und wir suchen eine orthonormale Basis von Eigenvektoren für $\mathbb{R}^3$. Das charakteristisches Polynom ist

$$\begin{aligned} \chi_A(t) = \det \begin{pmatrix} -t & -1 & 1 \\ -1 & -t & 1 \\ 1 & 1 & -t \end{pmatrix} &= -t \det \begin{pmatrix} -t & 1 \\ 1 & -t \end{pmatrix} + 1 \cdot \det \begin{pmatrix} -1 & 1 \\ 1 & -t \end{pmatrix} - 1 \cdot \det \begin{pmatrix} -1 & -t \\ 1 & 1 \end{pmatrix} \\ &= -t(t^2-1) + (t-1) + (-1+t) = -t(t^2-1) + 2(t-1) = -t(t-1)(t+1) + 2(t-1) \\ &= (t-1)(-t(t+1)+2) = -(t-1)(t^2+t-2) = -(t-1)^2(t+2) \end{aligned}$$

und die Eigenwerte sind $1, -2$.

Wir können eine orthonormale Basis von $\mathrm{Ker}(A - I_3)$ finden:

$$\mathrm{Ker}(A - I_3) = \mathrm{Ker} \begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \langle v_1, v_2 \rangle$$

Wir sehen dass (v_1, v_2) eine Basis von $\mathrm{Ker}(A - I_3)$ ist, aber nicht orthonormal. Um eine orthogonale Basis zu finden, nutzen wir das Gram-Schmidt-Verfahren:

$$v_1' = v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$v_2' = v_2 - \frac{\langle v_1', v_2 \rangle}{\langle v_1', v_1' \rangle} v_1' = v_2 - \frac{1}{2} v_1' = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

und am Ende finden wir eine orthonormale Basis:

$$v_1'' = \frac{1}{\|v_1'\|} v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

$$v_2'' = \frac{1}{\|v_2'\|} v_2 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}$$

Wir können auch eine orthonormale Basis von $\operatorname{Ker}(A+2I_3)$ finden:

$$\operatorname{Ker}(A+2I_3) = \operatorname{Ker}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \left\langle \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle = \langle v_3 \rangle$$

und eine orthonormale Basis ist dann

$$v_3'' = \frac{1}{\|v_3\|} v_3 = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

Endlich, eine orthonormale Basis von Eigenvektoren von A ist:

$$\mathcal{B} = (v_1'', v_2'', v_3'') = \left(\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right)$$

Sei $N = (v_1''|v_2''|v_3'')$ die Matrix, die diese Vektoren als Spaltenvektoren hat. Dann kann man überprüfen dass

$$N^t A N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Beispiel 6.5.6. Wir betrachten die hermitesche Matrix

$$A = \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

und wir suchen eine orthonormale Basis von Eigenvektoren für $\mathbb{C}^3$. Das charakteristisches Polynom ist

$$\begin{aligned} \chi_A(t) = \det\begin{pmatrix} -t & -1 & i \\ -1 & -t & -i \\ -i & i & -t \end{pmatrix} &= -t\det\begin{pmatrix} -t & -i \\ i & -t \end{pmatrix} + 1\cdot\det\begin{pmatrix} -1 & -i \\ -i & -t \end{pmatrix} - i\cdot\det\begin{pmatrix} -1 & -t \\ -i & i \end{pmatrix} \\ &= -t(t^2+i^2) + (t-i^2) - i(-i+it) = -t(t^2-1) + (t+1) + (1+t) = \\ &= -t(t-1)(t+1) + 2(t+1) = (t+1)(-t(t-1)+2) \\ &= -(t+1)(t^2-t-2) = -(t+1)^2(t-2) \end{aligned}$$

und die Eigenwerte sind $-1, 2$.

Wir können eine orthonormale Basis von $\operatorname{Ker}(A+I_3)$ finden:

$$\operatorname{Ker}(A+I_3) = \operatorname{Ker}\begin{pmatrix} 1 & -1 & i \\ -1 & 1 & -i \\ -i & i & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -i \\ 0 \\ 1 \end{pmatrix} \right\rangle = \langle v_1, v_2 \rangle$$

Wir sehen dass (v_1, v_2) eine Basis von $\operatorname{Ker}(A - I_3)$ ist, aber nicht orthonormal. Um eine orthogonale Basis zu finden, nutzen wir das Gram-Schmidt-Verfahren:

$$v_1' = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$v_2' = v_2 - \frac{\langle v_1', v_2 \rangle}{\langle v_1', v_1' \rangle} v_1' = v_2 + \frac{i}{2} v_1' = \begin{pmatrix} -\frac{i}{2} \\ \frac{i}{2} \\ 1 \end{pmatrix}$$

und am Ende finden wir eine orthonormale Basis:

$$v_1'' = \frac{1}{\|v_1'\|} v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

$$v_2'' = \frac{1}{\|v_2'\|} v_2 = \begin{pmatrix} \frac{i}{\sqrt{6}} \\ -\frac{i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}$$

Wir können auch eine orthonormale Basis von $\operatorname{Ker}(A - 2I_3)$ finden:

$$\operatorname{Ker}(A + 2I_3) = \operatorname{Ker} \begin{pmatrix} -2 & -1 & i \\ -1 & -2 & -i \\ -i & i & -2 \end{pmatrix} = \left\langle \begin{pmatrix} i \\ -i \\ 1 \end{pmatrix} \right\rangle = \langle v_3 \rangle$$

und eine orthonormale Basis ist dann

$$v_3'' = \frac{1}{\|v_3\|} v_3 = \begin{pmatrix} \frac{i}{\sqrt{3}} \\ -\frac{i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

Endlich, eine orthonormale Basis von Eigenvektoren von A ist:

$$\mathcal{B} = (v_1'', v_2'', v_3'') = \left(\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{i}{\sqrt{6}} \\ -\frac{i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} \frac{i}{\sqrt{3}} \\ -\frac{i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right)$$

Sei $N = (v_1'' | v_2'' | v_3'')$ die Matrix, die diese Vektoren als Spaltenvektoren hat. Dann kann man überprüfen dass

$$\overline{N}^t A N = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

6.6 Bilinearformen und der Trägheitssatz von Sylvester

Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und sei

$$b \colon V \times V \to \mathbb{R}$$

eine Bilinearform: das bedeutet dass b bilinear und symmetrisch ist, aber nicht unbedingt positiv definit. Wenn $A = M_{\mathcal{B}}(b)$ ist die Gramsche Matrix von b bezüglich eine Basis $\mathcal{B} = (v_1, \dots, v_n)$, dann ist $A \in \mathbb{R}^{n\times n}$ symmetrisch und wir können die Bilinearform b durch die Bilinearform

$$\langle\, , \rangle_A \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (x, y) \mapsto x^t A y$$

interpretieren. Skalarprodukte sind wichtige Beispiele von Bilinearformen, aber es gibt andere:

Beispiel 6.6.1 (Der Minkowski-Raum)**.** Sei $c \in \mathbb{R}, c > 0$ konstant. Wir betrachten die symmetrische Matrix

$$M = \begin{pmatrix} -c & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

und die induzierte Bilinearform

$$\langle\, , \rangle_M \colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^4, \qquad (x, y) \mapsto x^t M y$$

Der Vektorraum $\mathbb{R}$ zusammen mit dem Bilinearform $\langle\, , \rangle_M$ heißt der Minkowski-Raum. Er war von Hermann Minkowski eingeführt als Modell für die Spezielle Relativität: die Vektoren in diesem Raum haben vier Koordinaten

$$x = \begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

wobei t eine Zeit-Koordinate ist und x_1, x_2, x_3 Raum-Koordinaten sind.

Wir wollen Bilinearformen mit Hilfe des Spektralsatzes klassifizieren:

Satz 6.6.2 (Trägheitssatz von Sylvester)**.** *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und sei $b \colon V \times V \to \mathbb{R}$ eine Bilinearform auf V.*

1. *Es existiert eine Basis $\mathcal{B}$ von V so dass die Gramsche Matrix diagonal mit Einträge in $\{1, -1, 0\}$ ist:*
$$M_{\mathcal{B}}(b) = \begin{pmatrix} I_{r_+} & 0 & 0 \\ 0 & -I_{r_-} & 0 \\ 0 & 0 & 0_{r_0} \end{pmatrix}$$
mit $0_{r_0} \in \mathbb{R}^{r_0 \times r_0}$ die Nullmatrix.

2. *Wenn die Gramsche Matrix diagonal ist, sind die Anzahl r_+ von positive Einträge in der Hauptdiagonal, die Anzahl r_- von negative Einträge und die Anzahl r_0 von nulle Einträge unabhängig von der Basis $\mathcal{B}$.*

Beweis. 1. Sei $\mathcal{B}''$ eine beliebige Basis von V und sei $A = M_{\mathcal{B}''}(b)$ die Gramsche Matrix von b. Da die Matrix A symmetrisch ist, existiert eine orthogonale Matrix $N \in O_n(\mathbb{R})$ so dass $N^t \cdot A \cdot N = D$ diagonal ist. Insbesondere, existiert eine Basis $\mathcal{B}' = (v'_1, \dots, v'_n)$ von V, so dass $b(v'_i, v'_j) = 0$ falls $i \neq j$. Wir setzen dann

$$v_i = \begin{cases} \frac{1}{\sqrt{|b(v'_i, v'_i)|}} v'_i, & \text{falls } b(v'_i, v'_i) \neq 0 \\ v'_i, & \text{falls } b(v'_i, v'_i) = 0 \end{cases}$$

und wir sehen dass $b(v_i, v_j) = 0$, falls $i \neq j$ und dass $b(v_i, v_j) \in \{-1, 0, 1\}$. Das bedeutet dass die Gramsche Matrix $M_{\mathcal{B}}(b)$ diagonal mit Einträge in $\{1, 0, -1\}$ ist.

2. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis so dass die Gramsche Matrix $M_{\mathcal{B}}(b)$ diagonal ist , und sei $\lambda_i = b(v_i, v_i)$. Wir nehmen an dass

$$b(v_i, v_i) = \begin{cases} \lambda_i > 0 & \text{falls } 1 \leq i \leq r_+ \\ \lambda_i < 0 & \text{falls } r_+ + 1 \leq i \leq r_+ + r_- \\ \lambda_i = 0 & \text{falls } r_+ + r_- + 1 \leq i \leq r_+ + r_- + r_0 \end{cases}$$

Wir sehen dass $r_+ + r_- + r_0 = \dim V = n$ unabhängig von der Basis $\mathcal{B}$ ist, und auch dass $r_+ + r_- = \operatorname{rk} M_{\mathcal{B}}(b)$ unabhängig von der Basis ist (Warum?). Wir müssen nun zeigen dass r_+ unabhängig von der Basis ist. Wir zeigen dass

$$r_+ = \max\{\dim W \mid W \subseteq V \text{ Untervektorraum }, \ b_{|W} \text{ positiv definit }\}$$

Wir betrachten den Untervektorraum $V_+ = \langle v_1, \dots, v_{r_+} \rangle$: sei $v = \sum_{i=1}^{r_+} x_i v_i \in V$ mit $x_i \in \mathbb{R}$ nicht alle Null. Dann

$$b(v, v) = \sum_{i=1}^{r_+} x_i^2 \lambda_i > 0$$

so dass $b_{|V_+}$ positiv definit ist. Sei nun $W \subseteq V$ ein Untervektorraum so dass $b_{|W}$ positiv definit ist, und nehmen wir an, dass $\dim W > r_+$. Wir betrachten den Untervektorraum $V'_+ = \langle v_{r_+ + 1}, \dots, v_n \rangle$: sei $v = \sum_{i=r_+ + 1}^{n} x_i v_i \in V'_+$, mit $x_i \in \mathbb{R}$ dann

$$b(v, v) = \sum_{i=r_+ + 1}^{n} x_i^2 \cdot \lambda_i \leq 0$$

Wir sehen dass

$$\begin{aligned} \dim(W \cap V'_+) &= \dim W + \dim V'_+ - \dim(W + V'_+) > r_+ + r_- + r_0 - \dim(W + V'_+) \\ &= \dim V - \dim(W + V'_+) \geq 0 \end{aligned}$$

Das bedeutet dass ein Vektor $w \in W \cap V'_+, w \neq 0$ existiert: da $w \in W$, wissen wir dass $b(w, w) > 0$ aber da $w \in V'_+$ haben wir gezeigt dass $b(w, w) \leq 0$. Widerspruch.

□

Korollar 6.6.3. *Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Es gibt eine Matrix $N \in GL_n(\mathbb{R})$ so dass $N^t A N$ diagonal mit Einträge in $\{-1, 0, 1\}$. Außerdem, wenn $N^t A N$ diagonal ist, sind die Anzahl r_+ von positive Einträge, die Anzahl r_- von negative Einträge und die Anzahl r_0 von nulle Einträge unabhängig von N.*

Beweis. Wir können den Satz von Sylvester auf die Bilinearform $\langle \, , \, \rangle_A$ anwenden. □

Definition 6.6.4 (Trägheitssindex oder Signatur)**.** Mit der vorherige Notation, heißt Das Tripel (r_+, r_-, r_0) das Trägheitsindex oder Signatur von der Bilinearform b oder der symmetrische Matrix A.

Bemerkung 6.6.5. Der Beweis des Tragheitssatzes von Sylvester zeigt dass

$$r_+ = \max\{\dim W \mid W \subseteq V \text{ Untervektorraum }, \ b_{|W} \text{ positiv definit }\}$$

Korollar 6.6.6. *Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist positiv definit genau dann, wenn alle Eigenwerte von A positiv sind.*

Beweis. Die Matrix A ist positiv definit genau dann, wenn die Signatur $(n, 0, 0)$ ist. Dank dem Spektralsatz existiert $N \in O_n(\mathbb{R})$ so dass $N^tAN = D$ diagonal ist, mit den Eigenwerten von A in der Hauptdiagonal. Dann hat A Signatur $(n, 0, 0)$ genau dann, wenn alle Eigenwerte positiv sind. □

6.6.1 Das Sylvester-Kriterium

Lemma 6.6.7. *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und positiv-definit. Dann* $\det(A) > 0$.

Beweis. Da A positiv definit ist, existiert $N \in GL_n(\mathbb{R})$ so dass $N^tAN = I_n$. Dann

$$1 = \det(N^tAN) = \det(A) \cdot \det(N)^2$$

und dann $\det(A) > 0$. □

Definition 6.6.8 (Führende Hauptminoren)**.** Sei $A \in \mathbb{R}^{n\times n}$ eine Matrix. Die führenden Hauptminoren von A sind die Determinanten der Untermatrizen von A, die man durch die ersten k Zeilen und die ersten k Spalten für jedes $1 \leq k \leq n$ erhält. Es gibt n führende Hauptminoren.

Satz 6.6.9 (Sylvester-Kriterium)**.** *Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist positiv definit genau dann, wenn alle führende Hauptminoren positiv sind.*

♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ **Ende Vorlesung 19.7.2024** ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣ ♣

Beweis. Sei A eine symmetrische Matrix und sei $b = \langle\,\rangle_A$ die entsprechende symmetrische Bilinearform. Für jede $1 \leq k \leq n$, sei A_k die Matrix, die wir erhalten, wenn wir die ersten k Zeilen und Spalten von A nehmen. Sei auch $V_k = \langle e_1, \dots, e_k\rangle$, wobei $\mathcal{C} = (e_1, \dots, e_n)$ die kanonische Basis von $\mathbb{R}^n$ ist. Wir sehen dass $A_k = M_{(e_1,\dots,e_k)}(b_{|V_k})$.

Wenn A positiv-definit ist, dann ist auch $b_{|V_k}$ positiv definit für alle k. Das bedeutet dass $\det(A_k) > 0$. Andererseits, nehmen wir an dass $\det(A_k) > 0$ für alle k. Wir zeigen durch Induktion auf n dass A positiv definit ist. Wenn $n = 1$ ist die Aussage klar. Nehmen wir an dass $n > 1$ und dass die Aussage für $n - 1$ wahr ist. Insbesondere sehen wir dass die symmetrische Bilinearform $b_{|V_{n-1}}$ positiv definit ist. Wenn (r_+, r_-, r_0) die Signatur von A ist, bedeutet dass, das $r_+ \geq n - 1$. Dann existiert eine $N \in GL_n(\mathbb{R})$ so dass

$$N^tAN = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Mit $\lambda_n \in \{-1, 0, 1\}$. Da $\det(A) > 0$, wissen wir dass $\lambda_n = \det(N^tAN) = \det(N)^2 \det(A) > 0$. Das zeigt dass $\lambda_n = 1$ so dass $r_+ = n$ und A ist positiv definit. □

Bemerkung 6.6.10. Alle Aussagen in diesem Abschnitt gelten auch für hermitesche Formen auf komplexen Vektorräumen und für hermitesche Matrizen mit komplexen Koeffizienten. Insbesondere gelten für sie der Tragheitssatz von Sylvester und das Sylvesterkriterium.