4 Lineare Algebra

1 Vektorräume und Untervektorräume

Wir erinnern uns an die Gruppe $(\mathbb{R}^2, +)$ mit komponentenweiser Addition und die Schreibweise der $k-$ fachen Summe $(a,b) + \ldots + (a,b)$ durch $k \cdot (a,b) = (ka, kb)$. Es stellt sich die Frage, ob auch Ausdrücke wie $(-3) \cdot (a,b)$, $\frac{3}{2} \cdot (a,b)$ oder $\pi \cdot (a,b)$ sinnvoll gedeutet werden können.

Mathematisch betrachtet handelt es sich um eine Abbildung $\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$: Jeder reellen Zahl α und jedem Zahlenpaar $x = (x_1, x_2)$ wird ein neues Zahlenpaar $\alpha \cdot x = (\alpha x_1, \alpha x_2)$ zugeordnet.

Diese Abbildung nennt man *skalare Multiplikation* oder *Skalarmultiplikation*:

$$\cdot \; : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (\alpha, (x_1, x_2)) & \mapsto & (\alpha x_1, \alpha x_2) \end{array} \right.$$

Für alle $\alpha, \beta \in \mathbb{R}$ und für alle $x = (x_1, x_2),\ y = (y_1, y_2) \in \mathbb{R}^2$ gelten leicht zu beweisende Regeln wie

$$\begin{array}{lrcl} (1) & \alpha \cdot (x + y) & = & \alpha \cdot x + \alpha \cdot y \\ (2) & (\alpha + \beta) \cdot x & = & \alpha \cdot x + \beta \cdot x \\ (3) & (\alpha \cdot \beta) \cdot x & = & \alpha \cdot (\beta \cdot x) \\ (4) & 1 \cdot x & = & x \end{array}$$

Man beachte, dass $\cdot$ und $+$ in unterschiedlichen Bedeutungen vorkommen! (Welchen?)

Frage: Sind (1) – (4) auch für $(\mathbb{R}^3, +)$ gültig?

Interessante Frage: Kann man beide Fälle $\mathbb{R}^2$ und $\mathbb{R}^3$ sinnvoll zusammenfassen?

Def 1.1 Sei V eine Menge mit einer binären Verknüpfung $+$ (genannt Addition) und einer Abbildung $\cdot : \mathbb{R} \times V \to V$ (genannt skalare Multiplikation). Das Tripel $(V, +, \cdot)$ heißt *reeller Vektorraum* $:\Longleftrightarrow$

(V1) $(V, +)$ ist abelsche Gruppe

(V2) Für die skalare Multiplikation gelten sinngemäß die Regeln (1) bis (4).

Für jede natürliche Zahl n ist $(\mathbb{R}^n, +, \cdot)$ ein reeller Vektorraum. Ängstliche Zeitgenossen dürfen sich zunächst auf die anschaulichen Fälle $n = 2$ oder $n = 3$ beschränken. Sie sollten aber nie vergessen, dass die Vektorraumdefinition viel allgemeinere Strukturen zulässt. Deshalb wollen wir uns an dieser Stelle einmal kurz mit einem ganz anderen Vektorraum beschäftigen.

Beispiel: $V = \operatorname{Abb}([0,1], \mathbb{R}) = \{f : [0,1] \to \mathbb{R}\}$ mit $f + g : x \mapsto f(x) + g(x),\ \alpha \cdot f : x \mapsto \alpha \cdot f(x)$ ist ein reeller Vektorraum, dessen Elemente Funktionen sind.[1] Wir wollen den Nachweis der Eigenschaften (V1) und (V2) teilweise hier und teilweise in den Übungen durchführen:

(V1): $(V, +)$ ist abelsche Gruppe:

- $V \neq \emptyset$; beispielsweise wird durch $o(x) := 0$ für alle $x \in [0,1]$ eine Abbildung $o \in V$ definiert.
- $+$ ist eine binäre Verknüpfung (warum?).
- $(V, +)$ ist assoziativ, da die Addition in $\mathbb{R}$ assoziativ ist: $f + (g + h) = (f + g) + h$; denn für alle $x \in [0,1]$ ist $(f + (g + h))(x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) = (f(x) + g(x)) + h(x) = (f + g)(x) + h(x) = ((f + g) + h)(x)$.

[1] Es sei noch einmal darauf hingewiesen, dass $+$ und $\cdot$ jeweils zwei verschiedene Bedeutungen haben!

- $o \in V$ ist neutrales Element (zur Erinnerung: $o(x) := 0 \quad \forall x \in [0,1]$)
- Zu f ist $-f : x \mapsto -f(x)$ invers.
- $(V,+)$ ist abelsch: $(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x) \implies f+g = g+f$.

(V2): Es gelten die Regeln der Skalarmultiplikation:

- Für alle $\alpha \in \mathbb{R}$ und für alle $f, g \in V$ ist $\alpha \cdot (f+g) = \alpha \cdot f + \alpha \cdot g$; denn für alle $x \in [0,1]$ ist $(\alpha \cdot (f+g))(x) = \alpha \cdot ((f+g)(x)) = \alpha \cdot (f(x)+g(x)) = \alpha \cdot f(x) + \alpha \cdot g(x) = (\alpha \cdot f)(x) + (\alpha \cdot g)(x) = (\alpha \cdot f + \alpha \cdot g)(x)$.
- Für alle $\alpha, \beta \in \mathbb{R}$ und für alle $f \in V$ ist $(\alpha+\beta)\cdot f = \alpha \cdot f + \beta \cdot f$
- Für alle $\alpha, \beta \in \mathbb{R}$ und für alle $f \in V$ ist $(\alpha\beta)\cdot f = \alpha \cdot (\beta \cdot f)$
- Für alle $f \in V$ ist $1 \cdot f = f$

Wenn in einem Vektorraum $(V,+,\cdot)$ Klarheit über Addition und Skalarmultiplikation besteht, spricht man auch vom Vektorraum V. (*Frage*: Was ist in diesem Sinn der Vektorraum $\mathbb{R}^4$?)

Die abkürzende Schreibweise $g-h$ für $g+(-h)$, die wir aus Gruppen kennen, ist auch in Vektorräumen üblich, also beispielweise $x-y = x+(-y)$, ebenfalls gilt $\alpha(-v) = -\alpha v$ usw.

Reelle Vektorräume sind spezielle Vektorräume. Ersetzt man in der Definition $\mathbb{R}$ durch einen anderen Körper $\mathbb{K}$, spricht man von einem Vektorraum *über dem Körper* $\mathbb{K}$.

Um Verwechslungen auszuschließen, werden wir das neutrale Element der Gruppe $(V,+)$ mit o und das neutrale Element der Gruppe $(\mathbb{K},+)$ mit 0 bezeichnen.

Während die Notwendigkeit der Forderungen (1) – (3) der skalaren Multiplikation (mehr oder weniger) einsichtig ist, liegt die Bedeutung des vierten Axioms $1 \cdot x = x$ zunächst nicht auf der Hand: Es wird lediglich benötigt, um Sonderfälle auszuschließen. Sonst hätte beispielsweise jede abelsche Gruppe $(G,\circ)$ mit neutralem Element e zusammen mit der langweiligen Abbildung $\left\{ \begin{array}{ccc} \mathbb{R}\times G & \to & G \\ (r,g) & \mapsto & e \end{array} \right.$ das Recht auf die Bezeichnung reeller Vektorraum!

Frage: Warum fordert man nicht auch $0 \cdot x = o \ \forall x \in V$?

Antwort: Weil es aus den anderen Axiomen folgt, es gilt

Satz 1.1 Sei $(V,+,\cdot)$ ein reeller Vektorraum mit neutralem Element o. Dann gelten

(1) $\quad 0 \cdot x = o \qquad \forall x \in V$

(2) $\quad r \cdot o = o \qquad \forall r \in \mathbb{R}$

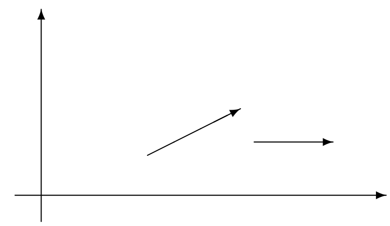
Beweis: Wir benutzen in beiden Teilen die Kürzungsregel aus der Gruppentheorie:

zu (1): $\quad 0\cdot x + o = 0 \cdot x = (0+0)\cdot x = 0\cdot x + 0 \cdot x \implies o = 0 \cdot x$

zu (2): $\quad r \cdot o + o = r \cdot o = r \cdot (o+o) = r\cdot o + r \cdot o \implies o = r \cdot o$

Bezeichnet man sinnvollerweise die Elemente eines Vektorraumes als Vektoren, folgt für den $\mathbb{R}^2$, dass jedes Zahlenpaar ein Vektor ist. Eventuell ist man es aus der Schule gewohnt, „Pfeile“ im $\mathbb{R}^2$ als Vektoren aufzufassen.

Für uns sind diese Pfeile *freie Vektoren*, festgelegt durch Anfangs– und Endpunkt (Länge) und Richtung. Jeder freie Vektor lässt sich als Differenz zweier Vektoren (Endpunkt minus Anfangspunkt oder Spitze minus Schaft) schreiben. Jeder Vektor ist ein freier Vektor mit Anfangspunkt $(0,0)$, manchmal spricht man dann auch von einem *Ortsvektor*.



Wir interessieren uns für spezielle Teilmengen von Vektorräumen. Als *Beispiel* untersuchen wir im reellen Vektorraum $\mathbb{R}^2$ die folgende Mengen

$$T_1 := \{(x,y) \in \mathbb{R}^2 \mid x,y \in \mathbb{Z}\} \qquad T_2 := \{(x,y) \in \mathbb{R}^2 \mid y = x\}$$
$$T_3 := \{(x,y) \in \mathbb{R}^2 \mid y = 1\} \qquad T_4 := \{(x,y) \in \mathbb{R}^2 \mid x,y \geq 0\}$$
$$T_5 := \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

Frage: Für welche $i = 1,2,3,4,5$ ist $(T_i, +, \cdot)$ selbst ein reeller Vektorraum?

Diese Frage werden wir in der Vorlesung beantworten und erkennen, dass nur wenige Teilmengen eines Vektorraumes selbst Vektorräume sind. Ist dies der Fall, spricht man von einem *Untervektorraum* oder kürzer *Unterraum*.

Beispiele: 1) T_2 ist Untervektorraum von $\mathbb{R}^2$, anschaulich handelt es sich um eine Gerade durch $(0,0)$.

2) $\{(x_1,x_2,x_3,0) \mid x_i \in \mathbb{R}\}$ ist Untervektorraum von $\mathbb{R}^4$.

3) $\{o\}$, wobei mit o der Nullvektor gemeint ist, und V sind stets Untervektorräume von V.

Will man überprüfen, ob eine Teilmenge U eines Vektorraums V ein Untervektorraum ist, muss man ähnlich wie bei Untergruppen zum Glück nicht alle Vektorraumaxiome überprüfen. Dafür sorgt

Satz 1.2 Sei V ein Vektorraum. $U \subset V$ ist (Unter)vektorraum $\iff$ U erfüllt

(U1) $U \neq \emptyset$

(U2) $x + y \in U \quad \forall x,y \in U$

(U3) $\alpha \cdot x \in U \quad \forall \alpha \in \mathbb{K},\ \forall x \in U$

Vor dem Beweis wollen wir kurz etwas zur Bedeutung dieses Satzes sagen. Nicht beide Beweisrichtungen sind mathematisch gleich wichtig. Während die Aussage „$\Rightarrow$“ ziemlich uninteressant ist, bedeutet die Kenntnis von „$\Leftarrow$“ eine große Arbeitserleichterung: Um eine nicht leere Teilmenge eines Vektorraumes als Untervektorraum zu erkennen, genügt es, in dieser Teilmenge die Abgeschlossenheit der Addition und der Skalarmultiplikation nachzuweisen.

Beweis: „$\Rightarrow$“: Aus den Voraussetzungen U Vektorraum und $U \subset V$ sind die Eigenschaften (U1) – (U3) nachzuweisen. Dies ist einfach, denn (U1) und (U2) sind direkte Konsequenzen aus der Vorgabe $(U,+)$ abelsche Gruppe, während (U3) gilt, weil $\cdot$ eine Skalarmultiplikation auf U ist.

„$\Leftarrow$“: Hier ist aus den Voraussetzungen V Vektorraum, $U \subset V$ und (U1) – (U3) der Nachweis zu führen, dass U Vektorraum ist.

Zur Skalarmultiplikation: Wegen (U3) ist die auf V gegebene Skalarmultiplikation auch auf U abgeschlossen. Die Bedingungen (1) – (4) sind für $\cdot : \mathbb{K} \times U \to U$ erfüllt, da sie wegen der Voraussetzung V Vektorraum sogar für jeden Vektor aus V gelten.

Zur abelschen Gruppe $(U,+)$: Wegen (U3) ist mit u auch $(-1)\cdot u \in U$. Aus den Übungen wissen wir $(-1)\cdot u = -u$. Zusammen mit (U2) folgt dann $x-y \in U \ \forall x,y \in U$. Weil $U \neq \emptyset$ wegen (U1), folgt die Behauptung aus Satz III.3.1 (Nachweis Untergruppe).

Da die Bedingung (U3) auch für $\alpha = 0$ erfüllt sein muss, und weil für jeden Vektor x gilt $0 \cdot x = o$, muss der Nullvektor in jedem Unterraum enthalten sein. Gehört umgekehrt der Nullvektor *nicht* zu einer Teilmenge eines Vektorraumes, so liegt garantiert *kein* Untervektorraum vor!

Wir interessieren uns für *alle* Untervektorräume des $\mathbb{R}^2$. Wie bereits bekannt, gibt es die beiden trivialen Unterräume $\{(0,0)\}$ und $\mathbb{R}^2$. Was ist sonst noch möglich?

Sei $(0,0) \neq u \in U \subset \mathbb{R}^2$. Damit U Unterraum sein kann, muss U außer u auch alle Vielfachen von u enthalten (warum?):

$$u \in U \text{ und } U \text{ Unterraum} \implies \alpha u \in U \ \ \forall \alpha \in \mathbb{R}, \text{ d.h., } \{\alpha u \mid \alpha \in \mathbb{R}\} \subset U\ .$$

Anschaulich handelt es sich bei $\{\alpha u \mid \alpha \in \mathbb{R}\}$ um eine Gerade durch $(0,0)$. Jede solche Gerade ist ein Unterraum:

Beh. 1: Sei $u \in \mathbb{R}^2 \backslash \{(0,0)\} \implies \mathbb{R}u := \{\alpha u \mid \alpha \in \mathbb{R}\}$ ist Unterraum.

Beweisskizze: Die Bedingungen (U1) – (U3) aus Satz 1.2 sind erfüllt; denn

(U1): $\mathbb{R}u \neq \emptyset$, beispielsweise ist $u = 1 \cdot u \in \mathbb{R}u$.

(U2): $x \in U$ und $y \in U \implies x = \alpha u,\ y = \beta u \implies x + y = \alpha u + \beta u = (\alpha+\beta)u \in \mathbb{R}u$

(U3): $\alpha \in \mathbb{R}$ und $x \in U \implies \alpha \cdot x = \alpha \cdot (\beta u) = (\alpha\beta)u \in \mathbb{R}u$.

Damit haben wir bereits alle Unterräume gefunden!

Beh. 2: Sei $U \subset \mathbb{R}^2$ Unterraum. Dann ist $U = \{(0,0)\}$ oder $U = \mathbb{R}u$ (Gerade durch den Ursprung) oder $U = \mathbb{R}^2$.

Beweis: Wir müssen zeigen: Besteht $U \neq \{(0,0)\}$ nicht nur aus einer Geraden der Gestalt $\mathbb{R}u$, so ist U bereits der gesamte Vektorraum $\mathbb{R}^2$.

Sei also $v = (v_1, v_2) \in U,\ u = (u_1, u_2) \in U$, beide von $(0,0)$ verschieden, und $v \notin \mathbb{R}u$.

Wir zeigen $u_1v_2 - u_2v_1 \neq 0$. Hierzu leiten wir, ausgehend vom Gegenteil $u_1v_2 - u_2v_1 = 0 \iff u_1v_2 = u_2v_1$, einen Widerspruch zur Voraussetzung $v \notin \mathbb{R}u$ her. Angenommen, $u_1v_2 = u_2v_1$:

1. Fall $u_1 = 0 \implies u_2 \neq 0 = v_1 \implies v = (0, v_2) = \frac{v_2}{u_2}(0, u_2) \in \mathbb{R}u$
2. Fall $u_1 \neq 0 \implies v_2 = \frac{u_2}{u_1}v_1 \implies v = (v_1, v_2) = \frac{v_1}{u_1}(u_1, u_2) \in \mathbb{R}u$

Da U nach Voraussetzung ein Unterraum ist, der u und v enthält, gilt $\alpha u + \beta v \in U$ für alle $\alpha, \beta \in \mathbb{R}$. Dies und $u_1v_2 - u_2v_1 \neq 0$ nutzen wir aus, um $U = \mathbb{R}^2$ zu zeigen:

Sei $(c_1, c_2) \in \mathbb{R}^2$ beliebig. Durch einfaches Nachrechnen erhält man für $\alpha = \frac{c_1v_2 - c_2v_1}{u_1v_2 - u_2v_1}$ und $\beta = \frac{u_1c_2 - u_2c_1}{u_1v_2 - u_2v_1}$

$$(c_1, c_2) \ = \ \frac{c_1v_2 - c_2v_1}{u_1v_2 - u_2v_1} \cdot (u_1, u_2) + \frac{u_1c_2 - u_2c_1}{u_1v_2 - u_2v_1} \cdot (v_1, v_2) \in U.$$

Außer den sogenannten trivialen Unterräumen $\{(0,0)\}$ und $\mathbb{R}^2$ ist (geometrisch gesprochen) genau jede Gerade durch den Nullpunkt $(0,0)$ ein Untervektorraum von $\mathbb{R}^2$.

Frage: Wie sehen alle Unterräume von $\mathbb{R}^3$ aus?

Kehren wir zu beliebigen Vektorräumen zurück!

Satz 1.3 Sei V ein beliebiger Vektorraum mit Untervektorräumen U_1, U_2. Dann gelten

(a) $U_1 \cap U_2$ ist Untervektorraum von V

(b) $U_1 \cup U_2$ ist Untervektorraum von $V \quad \Longleftrightarrow \quad U_1 \subset U_2$ oder $U_2 \subset U_1$

Beweis: Zu (a): Wir haben (U1) – (U3) für $U_1 \cap U_2$ zu zeigen; wir wissen, dass diese Regeln in U_1 und U_2 gelten.

(U1): Da $o \in U_1$ und $o \in U_2 \implies o \in U_1 \cap U_2$.

(U2): Seien $x, y \in U_1 \cap U_2 \implies x, y \in U_1 \wedge x, y \in U_2 \implies x + y \in U_1, U_2 \implies x + y \in U_1 \cap U_2$.

(U3): Seien $\alpha \in \mathbb{R},\ x \in U_1 \cap U_2 \implies \alpha x \in U_1 \wedge \alpha x \in U_2 \implies \alpha x \in U_1 \cap U_2$.

Zu (b): Da „$\Leftarrow$“ trivial ist (warum?), ist nur „$\Rightarrow$“ zu zeigen. Wir gehen indirekt vor:

Angenommen, $U_1 \cup U_2$ ist ein Untervektorraum und es gelte $U_1 \not\subset U_2$ und $U_2 \not\subset U_1$. Damit gibt es Elemente $u_1 \in U_1 \backslash U_2$ und $u_2 \in U_2 \backslash U_1$. Nach (U2) ist $u_1 + u_2 \in U_1 \cup U_2$, aber

$u_1 + u_2 \notin U_1$, denn sonst wäre $u_2 = u_2 + o = u_2 + (u_1 - u_1) = (u_1 + u_2) + (-1)u_1 \in U_1$, Widerspruch!

$u_1 + u_2 \notin U_2$, denn sonst wäre $u_1 = u_1 + o = u_1 + (u_2 - u_2) = (u_1 + u_2) + (-1)u_2 \in U_2$, Widerspruch!

D.h., $u_1 + u_2$ kann nicht in $U_1 \cup U_2$ liegen. Wenn $U_1 \cup U_2$ Untervektorraum ist, muss also eine der beiden Teilmengenbeziehungen gelten.

2 Lineare Unabhängigkeit

In diesem Abschnitt werden mit *linear (un)abhängig* und *Basis* grundlegende Begriffe der linearen Algebra eingeführt. V sei stets ein beliebiger reeller Vektorraum.

Def 2.1 Seien $v_1, \ldots, v_r \in V$. Dann heißt

$$L(v_1, \ldots, v_r) := \{\alpha_1 v_1 + \ldots + \alpha_r v_r \mid \alpha_i \in \mathbb{R} \text{ für alle } i = 1, \ldots, r\} \subset V$$

lineare Hülle von $v_1, \ldots, v_r$.

Man nennt $\alpha_1 v_1 + \ldots + \alpha_r v_r = \sum\limits_{i=1}^{r} \alpha_i v_i$ auch *Linearkombination* der Vektoren $v_1, \ldots, v_r$. Wenn die Summationsgrenzen klar sind, werden wir sie manchmal der Übersichtlichkeit wegen weglassen. Die lineare Hülle besteht aus allen Vektoren, die durch skalare Multiplikation und Vektoraddition aus den gegebenen Vektoren $v_1, \ldots, v_r$ erzeugt werden.

Beispiele: 1) In jedem reellen Vektorraum ist $L(o) = \{\alpha o \mid \alpha \in \mathbb{R}\} = \{o\}$.

2) $V = \mathbb{R}^2$: $L((4,7)) = \{\alpha(4,7) \mid \alpha \in \mathbb{R}\} = \mathbb{R}((4,7))$

3) $V = \mathbb{R}^2$: $L((1,0),(0,1)) = \{\alpha_1(1,0) + \alpha_2(0,1) \mid \alpha_i \in \mathbb{R}\} = \{(\alpha_1, \alpha_2) \mid \ldots\} = \mathbb{R}^2$.

4) $V = \mathbb{R}^2$: $L((1,0),(2,0)) = \{\alpha_1(1,0) + \alpha_2(2,0) \mid \alpha_i \in \mathbb{R}\} = \{(\alpha_1 + 2\alpha_2)(1,0) \mid \ldots\} = \mathbb{R}((1,0))$.

5) $V = \mathbb{R}^2$: $L((1,1),(1,-1)) = \{\alpha_1(1,1) + \alpha_2(1,-1) \mid \alpha_i \in \mathbb{R}\} = ?$

6) $V = \mathbb{R}^3$: Gesucht sind $L((1,0,0),(0,1,0))$; $L((1,0,0),(1,2,0))$; $L((1,2,0),(2,4,0))$.

7) Nur der Vollständigkeit halber sei $L(\emptyset) := \{o\}$ erwähnt.

Wie wir in den Beispielen 3) und 5) gesehen haben, können verschiedene Vektoren durchaus die gleiche lineare Hülle haben. Allen linearen Hüllen ist eine andere wichtige Eigenschaft gemeinsam:

Satz 2.1 Jede lineare Hülle L ist ein Untervektorraum.

Beweis: Einmal mehr haben wir die Gültigkeit von (U1) – (U3) zu zeigen.

(U1): $L \neq \emptyset$; denn wegen $o = 0 \cdot v_1 + \ldots + 0 \cdot v_r$ liegt o in jeder linearen Hülle.

(U2): Seien $x, y \in L \implies \exists \alpha_1, \ldots, \alpha_r,\ \beta_1, \ldots, \beta_r \in \mathbb{R}:\ x = \sum \alpha_i v_i,\ \ y = \sum \beta_i v_i$
$\implies x + y = \sum \alpha_i v_i + \sum \beta_i v_i = \sum (\alpha_i + \beta_i) v_i \in L.$

(U3): Sei $x \in L,\ \alpha \in \mathbb{R}$. Analog folgt $\alpha \cdot x = \alpha \cdot \sum \alpha_i v_i \ = \ \sum (\alpha \alpha_i) v_i \in L.$

Beispiele: $L((1,0),(1,1),(1,2)) = \mathbb{R}^2,\ \ L((1,0,1),(0,0,1))$ ist eine Ebene im $\mathbb{R}^3$. (Welche?)

Wie gerade im Beweis ausgenutzt, liegt o in jeder linearen Hülle. Wir interessieren uns jetzt für die Linearkombinationen $\sum \alpha_i v_i$ der Vektoren $v_1, \ldots, v_r$, die den Nullvektor ergeben. Natürlich gibt es immer die (triviale) Linearkombination mit $\alpha_i = 0$ für alle $i = 1, \ldots, r$:

$$0 \cdot v_1 + \ldots + 0 \cdot v_r = o$$

Frage: Gibt es in $L(v_1, \ldots, v_r)$ weitere Linearkombinationen $\alpha_1 v_1 + \ldots + \alpha_r v_r = o$?

Beispiele: 1) $V = \mathbb{R}^2,\ L((1,0),(1,1),(0,1))$. Hier gilt beispielsweise $(0,0) = 1 \cdot (1,0) - 1 \cdot (1,1) + 1 \cdot (0,1)$.
2) $V = \mathbb{R}^2,\ L((1,0),(0,1))$. Die Suche nach $\alpha, \beta \in \mathbb{R}$ mit $(0,0) = \alpha \cdot (1,0) + \beta \cdot (0,1) = (\alpha, \beta)$ ergibt zwangsweise $\alpha = \beta = 0$.

Wichtige **Feststellung:** r–Tupel von Vektoren $(v_1, \ldots, v_r)$ können unterschiedliche Eigenschaften bzgl. der Darstellung des Nullvektors als Linearkombination haben: Manchmal geht es wie im letzten Beispiel ausschließlich auf triviale Weise, manchmal sind wie im Beispiel davor auch andere Linearkombinationen möglich.

Def 2.2 Sei V ein beliebiger Vektorraum, $v_i \in V$.

(1) $(v_1, \ldots, v_r)$ heißt *linear unabhängig* $:\iff$ Für alle $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$ gilt

$$\alpha_1 v_1 + \ldots + \alpha_r v_r \ = \ o \quad \implies \quad \alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$$

Man sagt auch: Der Nullvektor lässt sich nur trivial aus $v_1, \ldots, v_r$ kombinieren.

(2) $(v_1, \ldots, v_r)$ heißt *linear abhängig* $:\iff$ $(v_1, \ldots, v_r)$ ist nicht linear unabhängig.

Beispiel: Im Vektorraum $\mathbb{R}^2$ ist $((1,0),(0,1))$ linear unabhängig, während $((1,0),(1,1),(0,1))$ linear abhängig ist.

Auch folgende Ausdrucksweisen für lineare (Un)abhängigkeit sind üblich: Die Vektoren $v_1, \ldots, v_r$ sind linear (un)abhängig, $\{v_1, \ldots, v_r\}$ ist linear (un)abhängig, usw.

Frage: Wie steht es um lineare Abhängigkeit bei $((1,0),(2,0)),\quad ((1,1,1),(0,1,2)),\quad ((0,0),(1,2))$?

Wir können lineare Unabhängigkeit auch folgendermaßen beschreiben:

Satz 2.2 $(v_1, \ldots, v_r)$ linear unabhängig $\iff$ Keiner der Vektoren v_i ist eine Linearkombination der übrigen Vektoren $v_1, \ldots, v_r$ (ohne v_i).

Beweis: Wir zeigen beide Richtungen indirekt.

„$\Rightarrow$“: Sei oBdA v_r *doch* als Linearkombination der übrigen Vektoren darstellbar:

$$v_r = \alpha_1 v_1 + \ldots + \alpha_{r-1} v_{r-1} \iff \alpha_1 v_1 + \ldots + \alpha_{r-1} v_{r-1} + (-1) v_r = o$$

$\implies$ $(v_1, \ldots, v_r)$ linear abhängig, ein Widerspruch zur Voraussetzung der linearen Unabhängigkeit dieser Vektoren.

„$\Leftarrow$“: Angenommen $(v_1, \ldots, v_r)$ linear abhängig, d.h., es gibt eine nicht–triviale Linearkombination des Nullvektors o:

$$\alpha_1 v_1 + \ldots + \alpha_r v_r = o \qquad \text{mit oBdA } \alpha_1 \neq 0$$

$\implies v_1 = \left(-\frac{\alpha_2}{\alpha_1}\right) v_2 + \ldots + \left(-\frac{\alpha_r}{\alpha_1}\right) v_r$ ist *doch* als Linearkombination der übrigen Vektoren darstellbar.

Will man Vektoren auf lineare (Un)abhängigkeit untersuchen, muss man stets *alle* beteiligten Vektoren im Auge behalten.

Beispiel: $((1,0),(1,1),(0,1))$ ist linear abhängig, obwohl je zwei dieser Vektoren linear unabhängig sind.

Wir fassen zusammen: Sind Vektoren $v_1, \ldots, v_r$ linear abhängig, kann man mindestens einen von ihnen als Linearkombination der übrigen aufschreiben (OBdA $v_r = \alpha_1 v_1 + \ldots + \alpha_{r-1} v_{r-1}$). Dies bedeutet in unserem Fall $L(v_1, \ldots, v_{r-1}) = L(v_1, \ldots, v_r)$. Um unnötigen Schreib– und Rechenaufwand zu vermeiden, ist man an linear unabhängigen Vektoren interessiert.

Def 2.3 Sei V ein beliebiger Vektorraum, $v_i \in V$. $(v_1, \ldots, v_r)$ heißt *Basis* von V $:\iff$

(B1): $(v_1, \ldots, v_r)$ ist linear unabhängig

(B2): $L(v_1, \ldots, v_r) = V$

Eine Basis ist ein *minimales Erzeugendensystem* eines Vektorraumes. Jeder Vektor aus V liegt in der linearen Hülle einer Basis. Entfernt man einen Vektor aus einer Basis, umfasst die lineare Hülle der restlichen Vektoren nicht mehr *alle* Vektoren aus V.

Anstelle von einer Basis $(v_1, \ldots, v_r)$ spricht man auch von einer Basis $\{v_1, \ldots, v_r\}$ oder man sagt, die Vektoren $v_1, \ldots, v_r$ bilden eine Basis.[2]

Ist eine Basis $(b_1, \ldots, b_n)$ gegeben, kann jeder Vektor v wegen (B2) als Linearkombination dieser Basisvektoren geschrieben werden. Weil die Basisvektoren wegen (B1) linear unabhängig sind, ist diese Darstellung eindeutig:

$$\forall\, v \in V \quad \exists_1\, (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \quad v = \alpha_1 b_1 + \ldots + \alpha_n b_n$$

[2] In dieser Vorlesung werden wir nicht auf Unterschiede eingehen, die sich aus der Schreibweise einer Basis als geordnetes Tupel oder ungeordnete Menge ergeben und je nach Gutdünken eine der beiden Darstellungen wählen.

Beispiele: 1) $V = \mathbb{R}^2$: $((1,0),(0,1))$ und $((1,1),(1,2))$ sind Basen, für $(2,5) \in \mathbb{R}^2$ gilt
$$(2,5) = 2 \cdot (1,0) + 5 \cdot (0,1) \quad \text{bzw.} \quad (2,5) = (-1) \cdot (1,1) + 3 \cdot (1,2)\ .$$

2) $((1,2),(1,3),(1,4))$ ist keine Basis, denn (B1) ist verletzt.

3) $V = \mathbb{R}^3$: $((1,0,0),(0,1,0),(0,0,1))$ ist eine Basis. $((1,0,0),(0,0,1))$ ist keine Basis, denn (B2) ist verletzt.

3) $V = \mathbb{R}^n$ mit $n \in \mathbb{N}$: Sei $e_i := (0,\dots,0,1,0,\dots,0)$ mit der 1 an i–ter Stelle. Diese besonders einfache Basis $(e_1,\dots,e_n)$ wird *kanonische Basis* genannt.

4) In der Schule haben Sie bereits mehr oder weniger bewusst die kanonische Basis benutzt, um Punkte der Anschauungsebene $\mathbb{R}^2$ eindeutig anzugeben: $P = (x,y)$ bedeutet nichts anderes als $P = (x,y) = x \cdot (1,0) + y \cdot (0,1) = x \cdot e_1 + y \cdot e_2$.

In der Schule wird Ihnen wahrscheinlich nie eine andere als die kanonische Basis begegnen, trotzdem ist es manchmal von Vorteil, mit anderen Basen zu arbeiten.

3 Dimension

Wir haben gesehen, dass ein Vektorraum viele verschiedene Basen haben kann. Es wird uns allerdings nicht gelingen, eine Basis des $\mathbb{R}^2$ mit nur einem oder eine mit mehr als zwei Vektoren zu finden. Jede Basis des $\mathbb{R}^n$ besteht aus genau n Vektoren. Niemand von Ihnen wird bezweifeln, dass jeder Vektorraum eine Basis besitzt. Der Beweis dieser simplen Aussage ist allerdings so schwierig, dass er in dieser Vorlesung nicht durchgeführt werden kann.

Satz 3.1 $(v_1,\dots,v_n)$ und $(w_1,\dots,w_m)$ seien Basen eines Vektorraumes. Dann gilt $n = m$.

Bevor wir Angaben zum Beweis machen, halten wir als wichtige Folgerung fest:

Def 3.1 a) Sei $(v_1,\dots,v_n)$ eine Basis eines Vektorraumes V. Dann heißt n die *Dimension* von V, geschrieben $\dim V = n$ oder $\dim_{\mathbb{K}} V = n$.

b) Sei V ein Vektorraum, der für kein $n \in \mathbb{N}$ eine Basis $(v_1,\dots,v_n)$ besitzt. Dann heißt V *unendlich dimensional*, geschrieben $\dim V = \infty$.

Diese letzte Definition steht nicht im Widerspruch zur Behauptung, dass es immer eine Basis gibt! In b) wird lediglich festgestellt, dass nicht jeder Vektorraum eine Basis aus endlich vielen Vektoren besitzt.

Beispiel: Für die uns vertrauten reellen Vektorräume $\mathbb{R}^n$ gilt $\dim \mathbb{R}^n = n$ (auch für $n = 1$).

Um Satz 3.1 beweisen zu können, benötigen wir zwei Hilfsmittel.

Satz 3.2 (*Basisergänzungssatz*, BErg)

Sei V ein beliebiger Vektorraum, $\{v_1,\dots,v_r\},\{w_1,\dots,w_s\}$ seien Teilmengen von V und es gelte $L(v_1,\dots,v_r,w_1,\dots,w_s) = V$. Wenn $(v_1,\dots,v_r)$ linear unabhängig ist, kann $(v_1,\dots,v_r)$ durch Hinzufügen von Vektoren aus $\{w_1,\dots,w_s\}$ zu einer Basis von V ergänzt werden.

Beweisskizze: 1. Wenn $L(v_1,\dots,v_r) = V$ gilt, sind keine Vektoren hinzuzufügen – auch dieser Fall wird vom BErg eingeschlossen.

2. Sei $L(v_1,\dots,v_r) \neq V = L(v_1,\dots,v_r,w_1,\dots,w_s)$. Man kann zeigen, dass es $w_i \in \{w_1,\dots,w_s\}$ gibt mit $w_i \notin L(v_1,\dots,v_r) \implies (v_1,\dots,v_r,w_i)$ linear unabhängig.

3. Man überprüft, ob $L(v_1,\dots,v_r,w_i) = V$ gilt. Wenn ja, ist man fertig, sonst führt man 2. analog für $L(v_1,\dots,v_r,w_i) \neq V$ durch: $\exists w_j \in \{w_1,\dots,w_s\}\backslash\{w_i\}$ mit $w_j \notin L(v_1,\dots,v_r,w_i) \implies (v_1,\dots,v_r,w_i,w_j)$ linear unabhängig, usw.

4. Wegen $s < \infty$ ist man nach endlich vielen Schritten fertig.

Beispiel: Im $\mathbb{R}^3$ ist $(v_1,v_2) = ((1,0,0),(0,1,0))$ linear unabhängig, für $(w_1,w_2,w_3) = ((1,1,0),(0,1,1),(1,1,1))$ gilt (ohne Beweis) $L(v_1,v_2,w_1,w_2,w_3) = \mathbb{R}^3$.

Wegen $L(v_1,v_2) = \{(\alpha_1,\alpha_2,0) \mid \alpha_i \in \mathbb{R}\} \neq \mathbb{R}^3$ gibt es $w_i \in \{w_1,w_2,w_3\}$ mit $w_i \notin L(v_1,v_2)$.

Wegen $w_1 = v_1 + v_2 \in L(v_1,v_2)$ kommt w_1 nicht in Frage, aber es ist $w_2 \notin L(v_1,v_2)$.

(v_1,v_2,w_2) ist linear unabhängig, man kann zeigen, dass dies eine Basis von $\mathbb{R}^3$ ist.

Frage: Was ist mit (v_1,v_2,w_3)?

Satz 3.3 (*Austauschlemma*, ATL)

Seien $\{v_1,\dots,v_n\},\{w_1,\dots,w_m\}$ Basen eines beliebigen Vektorraumes V. Dann gilt:

$$\forall v_i \in \{v_1,\dots,v_n\}\ \exists w_j \in \{w_1,\dots,w_m\}\ \text{ mit }\ \{v_1,\dots,v_n\}\backslash\{v_i\} \cup \{w_j\} \text{ ist Basis von } V.$$

Beweisskizze: OBdA sei $i = 1$. Wegen $L(w_1,\dots,w_m) = V \neq L(v_2,\dots,v_n)$ gibt es $w_j \in \{w_1,\dots,w_m\}$ mit $(w_j,v_2,\dots,v_n)$ linear unabhängig. BErg sagt uns, dass wir $(w_j,v_2,\dots,v_n)$ durch Hinzufügen von Vektoren aus $\{v_1\}$ zu einer Basis von V ergänzen können. Die einzige Möglichkeit, dies zu tun, ist aber nichts hinzuzufügen! $\{w_j,v_2,\dots,v_n\}$ ist bereits eine Basis, wir haben v_1 durch w_j ausgetauscht.

Beispiel: (e_1,e_2,e_3) und $((1,1,0),(1,0,1),(0,1,1))$ sind zwei Basen des $\mathbb{R}^3$. Wir wenden zweimal ATL an (die Lücken werden in der Vorlesung ausgefüllt):

$((1,0,0),(0,1,0),\qquad\qquad)$ und $((1,0,0),\qquad\qquad,\qquad\qquad)$ sind Basen des $\mathbb{R}^3$.

Der Beweis von Satz 3.1 folgt jetzt direkt aus ATL; denn nur Basen gleicher Länge können vollständig ausgetauscht werden.

Beispiele: 1) Der Körper $\mathbb{C}$ kann als ein Vektorraum über $\mathbb{R}$ aufgefasst werden. Da jede komplexe Zahl in der Form $z = a + b \cdot i$ mit $a,b \in \mathbb{R}$ geschrieben werden kann, ist $(1,i)$ eine Basis, damit ist $\dim_{\mathbb{R}} \mathbb{C} = 2$.

2) Was sind $\dim_{\mathbb{R}} \mathbb{R}$, $\dim_{\mathbb{Q}} \mathbb{R}$?

Der Nachweis der linearen (Un)abhängigkeit bei konkret gegebenen Vektoren ist oft nicht einfach. In einem Fall können Sie sich allerdings jede Rechenarbeit ersparen, nämlich wenn Sie in einem Vektorraum der Dimension n mehr als n Vektoren auf lineare Unabhängigkeit untersuchen sollen.

In Satz 1.3 wurde bewiesen, dass der Durchschnitt von Untervektorräumen ebenfalls ein Vektorraum ist. Was wissen wir über $\dim(U_1 \cap U_2)$? Wir wollen diese Problematik zuerst an einem *Beispiel* erörtern.

Sei $V = \mathbb{R}^3$, $U_i = L(e_i)$ für $i = 1,2,3$ und $U_4 = L(e_1,e_2)$.
Was ist $\dim U_i$, $\dim(U_1 \cap U_2)$, $\dim(U_1 \cap U_4)$, $\dim(U_3 \cap U_4)$?

Wie wir (hoffentlich) gesehen haben, hängt die Dimension des Durchschnitts nicht nur von den Dimensionen der beteiligten Räume ab, sondern auch von ihrer Lage zueinander. Ab jetzt kürzen wir *U ist Untervektorraum von V* durch $U \leq V$ ab.[3]

Def 3.2 Seien $U_1, U_2 \leq V$. Dann heißt

$$U_1 + U_2 := \{x + y \mid x \in U_1,\ y \in U_2\} \qquad \textit{Summe} \text{ von } U_1 \text{ und } U_2$$

Frage: Was unterscheidet $U_1 + U_2$ von $U_1 \cup U_2$?

Wir greifen das letzte *Beispiel* auf und erkennen

$U_1 + U_2 = \{(\alpha, 0, 0) + (0, \beta, 0) \mid \alpha, \beta \in \mathbb{R}\} = L(e_1, e_2)$

$U_1 + U_4 = \{(\alpha, 0, 0) + (\beta_1, \beta_2, 0) \mid \alpha, \beta_i \in \mathbb{R}\} = \{(\alpha + \beta_1, \beta_2, 0) \mid \alpha, \beta_i \in \mathbb{R}\}$
$= \{(\gamma, 0, 0) + (0, \beta_2, 0) \mid \gamma, \beta_2 \in \mathbb{R}\} = L(e_1, e_2)$

$U_3 + U_4 = \{(0, 0, \gamma) + (\beta_1, \beta_2, 0) \mid \gamma, \beta_i \in \mathbb{R}\} = \{(\beta_1, 0, 0) + (0, \beta_2, 0) + (0, 0, \gamma) \mid \beta_i, \gamma \in \mathbb{R}\}$
$= L(e_1, e_2, e_3)$

In allen Beispielen ist die Summe von Untervektorräumen selbst ein Vektorraum. Dies ist kein Zufall:

Satz 3.4 $U_1, U_2 \leq V \implies U_1 + U_2 \leq V$

Beweis: Wir überprüfen (U1) – (U3):

Zu (U1): Es ist $o = o + o \in U_1 + U_2$

Zu (U2): $x, y \in U_1 + U_2 \implies \exists x_i, y_i \in U_i$ mit $x = x_1 + x_2,\ \ y = y_1 + y_2$
$\implies x + y = (x_1 + x_2) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) \in U_1 + U_2$

Zu (U3): $x \in U_1 + U_2 \implies \exists x_i \in U_i$ mit $x = x_1 + x_2$
$\implies \alpha x = \alpha(x_1 + x_2) = \alpha x_1 + \alpha x_2 \in U_1 + U_2$ für jedes $\alpha \in \mathbb{K}$.

Für die untersuchten Mengen gilt $U_1 \cap U_2 \subset U_i \subset U_1 \cup U_2 \subset U_1 + U_2 \quad (i = 1, 2)$. Wir wollen das Verhältnis der Mengen $U_1 \cup U_2$ und $U_1 + U_2$ zueinander genauer untersuchen.

Satz 3.5 Für Vektorräume U_1, U_2 gilt

$$U_1 \cup U_2 \ = \ U_1 + U_2 \quad \iff \quad U_1 \subset U_2 \text{ oder } U_2 \subset U_1$$

Der **Beweis** ist eine Übungsaufgabe. (Hinweis: Man erinnere sich an den Beweis von Satz 1.3.)

Wir merken uns: Während die *Summe* von Untervektorräumen stets selbst wieder einen Vektorraum ergibt, trifft dies für die *Vereinigung* nur in uninteressanten Spezialfällen zu.

Satz 3.6 *Dimensionsformel für Untervektorräume*

Seien U_1, U_2 endlichdimensionale Untervektorräume. Dann gilt:

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

[3]Erinnern Sie sich an die analoge Bezeichnung für Untergruppen!

Beweisskizze: Sei $(v_1,\ldots,v_r)$ eine Basis von $U_1 \cap U_2$, des kleinsten der beteiligten Untervektorräume. Mit Hilfe vom BErg können wir diese Basis einerseits zu einer Basis $(v_1,\ldots,v_r,w_1,\ldots,w_s)$ von U_1 mit $\dim U_1 = r+s$ und andererseits zu einer Basis $(v_1,\ldots,v_r,z_1,\ldots,z_t)$ von U_2 mit $\dim U_2 = r+t$ ergänzen.

Man zeigt dann, dass $(v_1,\ldots,v_r,w_1,\ldots,w_s,z_1,\ldots,z_t)$ eine Basis von U_1+U_2 ist. Hieraus folgt

$$\dim(U_1+U_2) \;=\; r+s+t \;=\; (r+s)+(r+t)-r \;=\; \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

Die Hauptarbeit dieses Beweises steckt im Nachweis der Basiseigenschaft (B1) von $(v_1,\ldots,v_r,w_1,\ldots,w_s,z_1,\ldots,z_t)$; für Interessierte folgt der vollständige Beweis (kein Vorlesungsstoff):

Sei $\sum\limits_{\nu=1}^{r} \alpha_\nu v_\nu + \sum\limits_{\nu=1}^{s} \beta_\nu w_\nu + \sum\limits_{\nu=1}^{t} \gamma_\nu z_\nu \;=\; o \qquad (*)$

$$\Longrightarrow \sum_{\nu=1}^{t} \gamma_\nu z_\nu \;=\; -\sum_{\nu=1}^{r} \alpha_\nu v_\nu - \sum_{\nu=1}^{s} \beta_\nu w_\nu \;\in\; L(v_1,\ldots,v_r,w_1,\ldots,w_s) \;=\; U_1$$

Da $\sum\limits_{\nu=1}^{t} \gamma_\nu z_\nu \in U_2 \Longrightarrow \sum\limits_{\nu=1}^{t} \gamma_\nu z_\nu \in U_1 \cap U_2 \;=\; L(v_1,\ldots,v_r)$

$$\Longrightarrow \sum_{\nu=1}^{t} \gamma_\nu z_\nu \;=\; \sum_{\nu=1}^{r} \mu_\nu v_\nu \quad \Longleftrightarrow \quad \sum_{\nu=1}^{t} \gamma_\nu z_\nu + \sum_{\nu=1}^{r} (-\mu_\nu) v_\nu \;=\; o$$

Da $(v_1,\ldots,v_r,z_1,\ldots,z_t)$ als Basis von U_2 linear unabhängig ist, erkennen wir $\gamma_1 = \ldots = \gamma_t = 0$.

Wenn wir dies in $(*)$ einsetzen, folgt $\sum\limits_{\nu=1}^{r} \alpha_\nu v_\nu + \sum\limits_{\nu=1}^{s} \beta_\nu w_\nu \;=\; o$

Da $(v_1,\ldots,v_r,w_1,\ldots,w_s)$ eine Basis von U_1 ist, gilt $\alpha_i = \beta_j = 0$ für alle $i = 1,\ldots,r$ und $j = 1,\ldots,s$, damit ist (B1) bewiesen.

Für endlichdimensionale Vektorräume V kann man mit Hilfe dieses Satzes zeigen, dass *echte* Unterräume von V stets eine kleinere Dimension als V haben müssen.

Beispiel: Im $\mathbb{R}^3$ ist für $U = L((1,2,4),(1,0,2))$ und $V = L((2,0,4),(3,1,2))$ die Dimension von $U \cap V$ gesucht. (Geometrisch handelt es sich um den Schnitt zweier Ebenen.)

Lösung: $((1,2,4),(1,0,2))$ und $((2,0,4),(3,1,2))$ sind jeweils linear unabhängig: $\dim U = \dim V = 2$.

$\Longrightarrow \dim(U \cap V) = \dim U + \dim V - \dim(U+V) = 2+2-\dim(U+V)$ mit $\dim(U+V) \in \{2,3\}$.

$\dim(U+V) = 2$ ist nicht möglich: $\dim(U+V) = 2 \Longrightarrow U = U+V = V$, aber $(1,2,4) \in U \backslash V$. Daher muss $\dim(U+V) = 3$ sein. Der gesuchte Schnitt hat die Dimension 1, es handelt sich um eine Gerade durch den Ursprung.

4 Lineare Abbildungen

Grob gesprochen sind lineare Abbildungen bei Vektorräumen dasselbe wie Homomorphismen bei Gruppen, nämlich strukturerhaltende Abbildungen. Auch in diesem Kapitel seien V, W endlichdimensionale reelle Vektorräume, die durchaus unterschiedliche Dimensionen haben dürfen. Wenn nicht ausdrücklich etwas anderes erwähnt wird, geht es in den Beispielen um Vektorräume $\mathbb{R}^n$ mit $n \in \mathbb{N}$ und kanonischen Basen $(e_1,\ldots,e_n)$.

Def 4.1 $f : V \to W$ heißt *lineare Abbildung* oder *Homomorphismus* $:\iff$

(L1) $f(x+y) = f(x) + f(y) \quad \forall x, y \in V$

(L2) $f(\alpha x) = \alpha f(x) \quad \forall \alpha \in \mathbb{R},\ \forall x \in V$

Die Menge aller Homomorphismen von V nach W schreibt man $\mathrm{Hom}\,(V, W) := \{f : V \to W \mid f \text{ linear }\}$.

Man beachte, dass in (L1) und (L2) jeweils für verschiedene Verknüpfungen das gleiche Symbol benutzt wurde. Um die Nullvektoren der möglicherweise verschiedenen Vektorräume V und W unterscheiden zu können, bezeichnen wir sie mit o_V und o_W.

Beispiele: 1) $V = W = \mathbb{R}$: $f(x) := x + 1$ ist nicht linear, $g(x) := 5x$ ist linear.

2) $V = \mathbb{R}^2$, $W = \mathbb{R}$: $f((x_1, x_2)) := x_1 - x_2$ ist linear, $g((x_1, x_2)) := x_1 x_2$ ist nicht linear.

3) $V = \mathbb{R}^3$, $W = \mathbb{R}^2$: $f((x_1, x_2, x_3)) := (x_1 + x_2, x_3)$ ist linear.

Spezielle lineare Abbildungen haben besondere Namen. Eine lineare Abbildung $f : V \to W$ heißt *Endomorphismus*, falls $V = W$ gilt, und *Isomorphismus*, falls f bijektiv ist. Treffen beide Bedingungen zu, liegt ein *Automorphismus* vor.[4]

Man wird bei beliebigen linearen Abbildungen $f : V \to W$ nicht erwarten können, dass jeder Vektor aus W ein Urbild besitzt. Stets gilt aber $f(o_V) = o_W$:

$$f(o_V) = f(0 \cdot o_V) = 0 \cdot f(o_V) = o_W$$

Def 4.2 Sei $f \in \mathrm{Hom}\,(V, W)$. Dann heißt

$\mathrm{Kern}\, f := \{v \in V \mid f(v) = o_W\}$ *Kern* von f

$\mathrm{Bild}\, f := \{w \in W \mid \exists v \in V \text{ mit } f(v) = w\} = \{f(v) \mid v \in V\}$ *Bild* von f

Der Kern einer linearen Abbildung besteht aus allen Vektoren, deren Bild der Nullvektor ist. $\mathrm{Bild}\, f$ wird in der Literatur auch $\mathrm{im}\, V$ („image") oder $f(V)$ genannt.

Beispiel: Was sind $\mathrm{Bild}\, f$ und $\mathrm{Kern}\, f$ von $f : \begin{cases} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x, y) & \mapsto & (x, x) \end{cases}$?

Für jede lineare Abbildung $f : V \to W$ ist $\mathrm{Kern}\, f \subseteq V$ und $\mathrm{Bild}\, f \subseteq W$. Es gilt sogar

Satz 4.1 Sei $f \in \mathrm{Hom}\,(V, W)$. Dann gelten

(a) $\mathrm{Kern}\, f \leq V$

(b) $\mathrm{Bild}\, f \leq W$

(c) f injektiv $\iff$ $\mathrm{Kern}\, f = \{o_V\}$

Beweis: Um (a) und (b) zu beweisen, müssen wir jeweils (U1) – (U3) überprüfen.

zu (a), (U1): Wir wissen bereits $o_V \in \mathrm{Kern}\, f$, also $\mathrm{Kern}\, f \neq \emptyset$.

(U2): $v, w \in \mathrm{Kern}\, f \implies f(v + w) = f(v) + f(w) = o_W + o_W = o_W \implies v + w \in \mathrm{Kern}\, f$.

(U3): $\alpha \in \mathbb{R},\ v \in \mathrm{Kern}\, f \implies f(\alpha v) = \alpha f(v) = \alpha o_W = o_W \implies \alpha v \in \mathrm{Kern}\, f$.

[4] Man vergleiche die analogen Bezeichnungen bei Gruppen!

zu (b), (U1): Da $V \neq \emptyset \implies \exists v \in V$ mit $f(v) \in W \implies$ Bild $f \neq \emptyset$

(U2), (U3): $\alpha \in \mathbb{R},\ x, y \in \text{Bild}\, f \implies \exists v, w \in V$ mit $f(v) = x, f(w) = y \implies$
$f(v+w) = f(v) + f(w) = x + y \implies x + y \in \text{Bild}\, f$,
$f(\alpha v) = \alpha f(v) = \alpha x \implies \alpha x \in \text{Bild}\, f$.

zu (c): „$\Rightarrow$": Ist klar, da wir den wesentlichen Schritt bereits bewiesen haben (wo?).

„$\Leftarrow$": Seien $v, w \in V$ mit $f(v) = f(w) \iff o_W = f(v) - f(w) = f(v - w)$
$\implies v - w \in \text{Kern}\, f \implies v - w = o_V \implies v = w \implies f$ injektiv.

Wir stellen einen Zusammenhang zwischen den Dimensionen der Untervektorräume Kern f und Bild f her. Aus Kern $f \leq V$ folgt $0 \leq \dim \text{Kern}\, f \leq \dim V$ und aus Bild $f \leq W$ folgt $0 \leq \dim \text{Bild}\, f \leq \dim W$.

Beispiele: 1) $f : \begin{cases} V & \to & W \\ v & \mapsto & o_W \end{cases} \implies \dim \text{Kern}\, f = \dim V, \quad \dim \text{Bild}\, f = 0.$

2) $f : \begin{cases} V & \to & V \\ v & \mapsto & v \end{cases} \implies \dim \text{Kern}\, f = 0, \quad \dim \text{Bild}\, f = \dim V.$

3) $f : \begin{cases} \mathbb{R}^3 & \to & \mathbb{R}^2 \\ (x, y, z) & \mapsto & (x + y, z) \end{cases}$: Wegen $f((1,0,0)) = (1,0)$ und $f((0,0,1)) = (0,1)$ enthält das Bild zwei linear unabhängige Vektoren, damit ist $\dim \text{Bild}\, f = 2$.

$(x, y, z) \in \text{Kern}\, f \iff f(x, y, z) = (x + y, z) = (0, 0) \implies x = -y,\ z = 0$. Damit ist Kern $f = \{(x, -x, 0) \mid x \in \mathbb{R}\} = L((1, -1, 0)) \implies \dim \text{Kern}\, f = 1$.

Satz 4.2 *Dimensionsformel für lineare Abbildungen*

Sei $f : V \to W$ linear. Dann gilt

$$\dim V = \dim\ \text{Kern}\ f + \dim\ \text{Bild}\ f$$

Beweisskizze: Jede Basis $(v_1, \ldots, v_r)$ des Kerns können wir dank BErg zu einer Basis von V ergänzen, sei $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ die ergänzte Basis. Für ein beliebiges $v \in V$ gilt dann

$$f(v) = f\left(\sum_{i=1}^{n} \alpha_i v_i\right) = f\left(\sum_{i=1}^{r} \alpha_i v_i\right) + f\left(\sum_{i=r+1}^{n} \alpha_i v_i\right) = o_W + \sum_{i=r+1}^{n} \alpha_i f(v_i) = \sum_{i=r+1}^{n} \alpha_i w_i$$

Damit ist Bild $f = L(w_{r+1}, \ldots, w_n)$. Aus der linearen Unabhängigkeit von $(w_{r+1}, \ldots, w_n)$ (Beweis für Interessierte anschließend) folgt $\dim \text{Bild} f = n - r$, was zu zeigen war.

Beh: $(w_{r+1}, \ldots, w_n)$ linear unabhängig *Bew*: Sei $\beta_{r+1} w_{r+1} + \ldots + \beta_n w_n = o_W$. Wegen $w_i = f(v_i)$ folgt

$$o_W = \sum_{i=r+1}^{n} \beta_i f(v_i) = f\left(\sum_{i=r+1}^{n} \beta_i v_i\right)$$

$\sum_{i=r+1}^{n} \beta_i v_i$ gehört damit zu Kern $f = L(v_1, \ldots, v_r)$. Es folgt $\sum_{i=r+1}^{n} \beta_i v_i = \sum_{i=1}^{r} \gamma_i v_i$, also

$$\gamma_1 v_1 + \ldots + \gamma_r v_r - \beta_{r+1} v_{r+1} - \ldots - \beta_n v_n = o_V \implies \gamma_1 = \ldots = \gamma_r = \beta_{r+1} = \ldots = \beta_n = 0 \quad .$$

In der Dimensionsformel taucht $\dim W$ aus leicht erklärlichen Gründen nicht explizit auf: Bei jeder linearen Abbildung $f : V \to W$ kann W durch einen „größeren“ Vektorraum W' ersetzt werden, also $W \leq W'$, ohne dass sich Kern f und Bild f und damit deren Dimensionen ändern. Im Fall $\dim W < \dim V$ kann f nicht injektiv sein, denn $\dim \operatorname{Kern} f = \dim V - \dim \operatorname{Bild} f \geq \dim V - \dim W > 0$.

Weil die Dimension des Bildes auch *Rang* von f genannt wird, ist $\dim V = \dim \operatorname{Kern} f + \operatorname{Rang} f$ eine andere Fassung dieser Formel.

Frage/Beispiel: Wieviele lineare Abbildungen $f : \mathbb{R}^2 \to \mathbb{R}^3$ gibt es mit
$f(e_1) = (2, 4, 0)$ und $f(e_2) = (1, 0, 2)$?

Antwort: Sei $(x, y) \in \mathbb{R}^2$ beliebig. Wegen $(x, y) = xe_1 + ye_2$ folgt aus den Eigenschaften einer linearen Abbildung

$$f((x, y)) = f(xe_1 + ye_2) = xf(e_1) + yf(e_2) = x(2, 4, 0) + y(1, 0, 2) = (2x + y, 4x, 2y).$$

Damit hat genau eine lineare Abbildung die verlangten Eigenschaften.

Das ist kein Zufall: Ist $(b_1, \ldots, b_n)$ eine Basis von V und sind $w_1, \ldots, w_n$ beliebige, nicht notwendig verschiedene Vektoren aus W, so gibt es immer genau eine lineare Abbildung $f : V \to W$ mit $f(b_i) = w_i$ für $i = 1, \ldots, n$. Damit wir diesen Sachverhalt nicht so schnell vergessen, sei er hier noch einmal als Satz (ohne formalen Beweis) formuliert:

Satz 4.3 Sei $(b_1, \ldots, b_n)$ eine Basis von V, sei $(w_1, \ldots, w_n)$ ein beliebiges n–Tupel von Vektoren aus W. Dann gilt

$$\exists_1\, f \in \operatorname{Hom}(V, W) \quad f(b_i) = w_i \text{ für } i = 1, \ldots, n\,,$$

nämlich $f(v) = f\left(\sum\limits_{i=1}^{n} \alpha_i b_i\right) = \sum\limits_{i=1}^{n} \alpha_i f(b_i) = \sum\limits_{i=1}^{n} \alpha_i w_i \quad \forall\, v \in V.$

Wichtige **Feststellung**: Die gesamte Information über eine lineare Abbildung ist bereits in den Bildern der Basisvektoren enthalten!

Gemessen an der Gesamtzahl aller Abbildungen zwischen zwei Vektorräumen gibt es relativ wenige lineare Abbildungen. Wir wollen diese Behauptung ausnahmsweise am Beispiel der endlichen Vektorräume $\mathbb{K}^3$ und $\mathbb{K}^2$ für $\mathbb{K} = \mathbb{Z}_2$ belegen:

Es gibt (Kombinatorik!) insgesamt $4^8 > 65\,000$ Abbildungen von $\mathbb{Z}_2^3$ nach $\mathbb{Z}_2^2$. Da man $4 \cdot 4 \cdot 4 = 64$ Tripel (w_1, w_2, w_3) mit $w_i \in \mathbb{Z}_2^2$ bilden kann, sind nur 64 dieser Abbildungen linear.

Die Verkettung linearer Abbildungen ergibt allerdings stets wieder eine lineare Abbildung. (Versuchen Sie den einfachen Nachweis!)

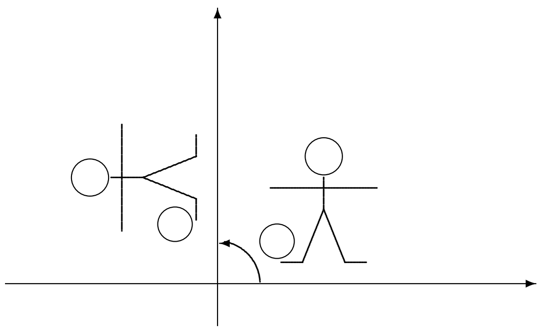
5 Beispiele von linearen Abbildungen in $\mathbb{R}^2$ und $\mathbb{R}^3$

In diesem Abschnitt gibt es keine neue Theorie und keine Sätze. Es werden nur einige hoffentlich aus der Schulzeit bekannte Abbildungen behandelt. Wir legen jeweils die kanonische Basis zu Grunde, alle Überlegungen beziehen sich auf Satz 4.3, wonach jede lineare Abbildung eindeutig durch die Bilder der Basiselemente festliegt.

A. Wir untersuchen einige lineare Abbildungen der Anschauungsebene $\mathbb{R}^2$, also $f : \mathbb{R}^2 \to \mathbb{R}^2$:

1. Sei $f(e_1) = e_2$ und $f(e_2) = -e_1$. Für alle $(x, y) \in \mathbb{R}^2$ ist dann

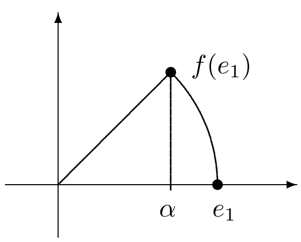
$$f((x, y)) = f(xe_1 + ye_2) = xf(e_1) + yf(e_2) = xe_2 - ye_1 = (-y, x)$$



Es handelt sich um eine Drehung mit Drehzentrum $((0, 0))$, der Drehwinkel beträgt $\frac{\pi}{2} = 90°$.

Alle Drehungen um $(0, 0)$ sind lineare Abbildungen, bezüglich der Hintereinanderausführung bilden sie eine abelsche Gruppe. Drehungen $\alpha \neq id$ mit einem anderem Drehzentrum sind wegen $\alpha((0, 0)) \neq (0, 0)$ nicht linear. Ab jetzt soll bei Drehungen stets $(0, 0)$ das Drehzentrum sein.

2. Sei f eine Drehung um $45°$, gesucht ist $f((x, y))$. Wir benötigen die Bilder der Basisvektoren.



$f(e_1) = \alpha e_1 + \beta e_2$ mit $\alpha = \beta$ und $\alpha^2 + \beta^2 = 1$
$\Longrightarrow\ f(e_1) = (\alpha, \alpha)$ mit $\alpha = \frac{\sqrt{2}}{2}$.
$f(e_2)$ und damit f werden in den Übungen bestimmt.

3. Sei g die Spiegelung an der Geraden (Achse) $y = -x$, gesucht ist $g((x, y))$. Da auch Spiegelungen mit einer Achse durch den Ursprung lineare Abbildungen sind, ist

$$g((x, y)) = xg(e_1) + yg(e_2) = x(-e_2) + y(-e_1) = (-y, -x)\ .$$

4. Gesucht ist $f \circ g$ mit f von 1. und g von 3. Für die lineare Abbildung $f \circ g$ gilt

$$(f \circ g)((x, y)) = f(g((x, y))) = f((-y, -x)) = (x, -y)\ .$$

Es handelt sich um eine Spiegelung an der x–Achse.

5. Translationen (Verschiebungen) bilden zwar mit der Hintereinanderausführung eine abelsche Gruppe, sind aber mit Ausnahme der Identität nicht linear (warum?).

6. *Fragen*: Wie lautet $h((x, y))$, wenn h eine zentrische Streckung um $(0, 0)$ mit Streckungsfaktor 2 ist? Was passiert mit dem Fußballspieler von 1. unter der linearen Abbildung $e_1 \mapsto e_1,\ e_2 \mapsto e_1 + e_2$?

B. Jetzt untersuchen wir lineare Abbildungen des Anschauungsraumes $\mathbb{R}^3$, hierbei beschränken wir uns auf spezielle Drehungen und Spiegelungen. Die Drehungen haben als Drehachse eine Gerade durch den Koordinatenursprung. Der Drehwinkel wird in einer Ebene senkrecht zur Drehachse gemessen, wobei auf die eindeutige Angabe des Drehwinkels (Drehrichtung) geachtet werden muss. Gespiegelt wird an einer Ebene, die ebenfalls durch $(0,0,0)$ verläuft.

1. Drehung um die z–Achse mit Drehwinkel von $90°$. (Wir blicken in z–Richtung und drehen entgegen dem Uhrzeigersinn.)

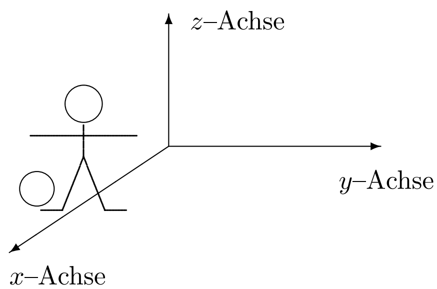
$$e_1 \mapsto -e_2, \qquad e_2 \mapsto e_1, \qquad e_3 \mapsto e_3$$

Für ein beliebiges $(x,y,z) \in \mathbb{R}^3$ erhalten wir $f((x,y,z)) = -xe_2 + ye_1 + ze_3 = (y,-x,z)$.

2. Spiegelung an der yz–Ebene: $e_1 \mapsto -e_1, \qquad e_2 \mapsto e_2, \qquad e_3 \mapsto e_3, \qquad g((x,y,z)) = (-x,y,z)$.

3. Wir untersuchen $g \circ f$: $(g \circ f)((x,y,z)) = g(f((x,y,z))) = g((y,-x,z)) = (-y,-x,z)$. Es handelt sich um eine Spiegelung, die zugehörige „Achse“, an der gespiegelt wird, ist die Ebene $\{(x,-x,z) \mid x,z \in \mathbb{R}\}$.

4. *Frage*: Welche Abbildung ist $f \circ g$?



Frage: Wo befindet sich der Fußballspieler nach Durchführung der Abbildungen $f \circ g$ bzw. $g \circ f$?

5. Für die Abbildung h mit $e_1 \mapsto e_2, \quad e_2 \mapsto e_3, \quad e_3 \mapsto e_1$ gilt $h((x,y,z)) = (z,x,y)$. Wir untersuchen, welche Punkte unter h festbleiben: $(x,y,z) = h((x,y,z)) = (z,x,y) \iff x = y = z$. Genau die Punkte auf der Geraden $\mathbb{R}((1,1,1)) = L(e_1 + e_2 + e_3)$ sind Fixpunkte, h ist eine Drehung. Wie groß ist der Drehwinkel?

Alle Drehungen mit einer Drehachse durch $(0,0,0)$, alle Spiegelungen an einer Ebene durch (0,0,0) und alle zentrischen Streckungen mit Zentrum $(0,0,0)$ sind lineare Abbildungen.

6 Lineare Abbildungen und Matrizen

Noch einmal zur Aussage von Satz 4.3: Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ ist eindeutig bestimmt durch n Vektoren $f(e_1) = (a_{11},\ldots,a_{m1}),\ldots,f(e_n) = (a_{1n},\ldots,a_{mn}) \in \mathbb{R}^m$, also durch $n \cdot m$ viele reelle Zahlen a_{ij}, die nicht verschieden sein müssen. Umgekehrt legt jedes $n \cdot m$–Tupel reeller Zahlen genau eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ fest:

$$f(x) = f\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} x_i (a_{1i},\ldots,a_{mi})$$

Wir wollen hierfür eine übersichtlichere Schreibweise kennenlernen.

Def 6.1 Seien a_{ij} reelle Zahlen. Das Schema

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \text{heißt } (m \times n)\text{–}Matrix\,.$$

Die Menge aller $(m \times n)$ – Matrizen schreiben wir $M(m \times n)$.

Beispiel: $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 5 \end{pmatrix}$ ist eine (2×4) – Matrix, $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ eine (3×1)–Matrix.

Eine $(m \times n)$ – Matrix hat m *Zeilen* und n *Spalten*, geschrieben $A = (a_{ij}) \in M(m \times n)$, und besteht aus $m \cdot n$ vielen Zahlen.

Was haben Matrizen mit linearen Abbildungen zu tun? Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ ist ebenfalls durch $m \cdot n$ viele Zahlen eindeutig festgelegt. Schreiben wir jeden der n die lineare Abbildung bestimmenden Bildvektoren $f(e_i) = (a_{1i}, \ldots, a_{mi})$ als *Spaltenvektor* [5]

$$f(e_i) = (a_{1i}, \ldots, a_{mi}) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} ,$$

erhalten wir als zugehörige Matrix

$$A = (f(e_1)\ \ldots\ f(e_n)) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M\,(m \times n)$$

Wichtige **Feststellung:** Zu jeder linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ gehört eine Matrix $A \in M(m \times n)$.

Beispiel: Zur Drehung in der Anschauungsebene (Anfang Kapitel 5) gehört die Matrix

$$(f(e_1)\ \ f(e_2)) = (e_2 \quad -e_1) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M\,(2 \times 2)$$

Wir sind es gewohnt, Abbildungen f in der Form $f(x) = \ldots$ anzugeben, *beispielsweise* die eben erwähnte Drehung durch $f((x, y)) = (-y, x)$. Auch diese Darstellung kann man bei linearen Abbildungen leicht aus der zugehörigen Matrix gewinnen; denn es gibt folgende Verbindung zwischen Matrix A und Urbild x einerseits und Bild $f(x)$ andererseits:

Def 6.2 Sei $A \in M(m \times n),\ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n = M(n \times 1)$. Dann ist

$$A \cdot x = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \in M(m \times 1) = \mathbb{R}^m$$

[5] Wir erlauben uns die Freiheit, Vektoren wahlweise in einer Zeile $(x_1, \ldots, x_n)$ (beachte die Kommata) oder als $(n \times 1)$ – Matrix (Spaltenvektor) zu schreiben.

Matrizen und Vektoren kann man genau dann miteinander multiplizieren, wenn die Matrix so viele Spalten wie der Vektor Zeilen hat, die Rechenvorschrift kann man sich gut durch die Regel *Zeile mal Spalte* merken.

Beispiele: $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \qquad\qquad , \quad \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \end{pmatrix} =$

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \qquad\qquad , \quad \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} =$$

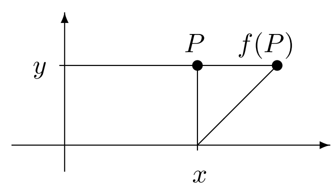
Kehren wir zu der linearen Abbildung f zurück! Wir stellen fest: Das Ergebnis der soeben definierten Multiplikation $A \cdot x$ ist nichts anderes als $f(x)$, geschrieben als Spaltenvektor:

Satz 6.1 a) Sei $A \in M(m \times n)$. Dann ist $f : \left\{ \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^m \\ x & \mapsto & Ax \end{array} \right.$ eine lineare Abbildung.

b) Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ linear. Dann ist $f(x) = Ax$ mit $A = (f(e_1) \dots f(e_n)) \in M(m \times n)$.

Beispiele: 1) Zur Drehung aus Kapitel 5: Es ist $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} = (-y, x)$.

2) $f((x,y)) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \end{pmatrix} = (x+y, y)$.



Bei dieser Abbildung handelt es sich um eine sogenannte *Scherung*.

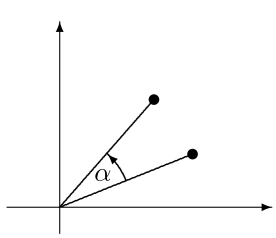
3) $f : \mathbb{R}^3 \to \mathbb{R}^3$, Projektion auf die xy–Ebene: $e_1 \mapsto e_1,\ e_2 \mapsto e_2,\ e_3 \mapsto 0$.

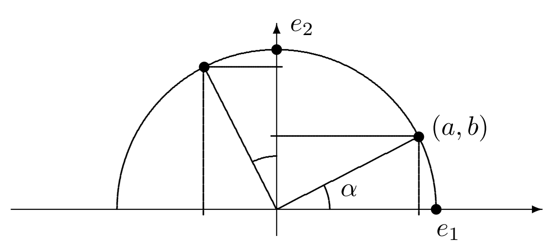
Die zugehörige Matrix ist $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

4) $f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^4 \\ (x,y,z) & \mapsto & (2x-z, 3y, x+y-z, y-3z) \end{array} \right.$ Wie lautet die zugehörige Matrix?

5) Welche Abbildung gehört zu $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$?

6) Gesucht ist im $\mathbb{R}^2$ eine Drehung um den Ursprung mit dem Drehwinkel $\alpha = 27°$.





Es ist $b = \sin\alpha$ und $a = \cos\alpha$

Wegen $f(e_1) = ae_1 + be_2$ und $f(e_2) = -be_1 + ae_2$ ist die zugehörige Matrix

$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, it ist $f((x,y)) = (ax - by, bx + ay)$.

Wir erinnern uns an die Verkettung von Abbildungen. $f : \mathbb{R}^m \to \mathbb{R}^l$ und $g : \mathbb{R}^n \to \mathbb{R}^m$ können hintereinander ausgeführt werden, und zwar erst g und dann f. Sind beide Abbildungen linear, ist auch die Verkettung $f \circ g : \mathbb{R}^n \to \mathbb{R}^l$ linear:

$$\begin{aligned} (f\circ g)(x+y) &= f(g(x+y)) = f(g(x)+g(y)) = f(g(x)) + f(g(y)) = (f\circ g)(x) + (f\circ g)(y) \\ (f\circ g)(\alpha x) &= f(g(\alpha x)) = f(\alpha g(x)) = \alpha f(g(x)) = \alpha (f\circ g)(x) \end{aligned}$$

Nach Satz 6.1 gibt es Matrizen A, B, C mit $f(x) = Ax,\ g(x) = Bx,\ (f\circ g)(x) = Cx$. Wir untersuchen den Zusammenhang dieser Matrizen:

$$Cx = (f\circ g)(x) = f(g(x)) = f(Bx) = A(Bx)$$

Beispiel: $f : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \mapsto (x+y, 2x-y)$ mit $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$,

$g : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \mapsto (x-y, x)$ mit $B = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$.

Es ist $(f\circ g)((x,y)) = f(g((x,y))) = f((x-y,x)) = (x-y+x, 2(x-y)-x) = (2x-y, x-2y)$
und $(g\circ f)((x,y)) = g(f((x,y))) = g((x+y, 2x-y)) = (x+y-(2x-y), x+y) = (-x+2y, x+y)$.
Die Matrix $C \in M(2\times 2)$ mit $(f\circ g)((x,y)) = C(x,y) = C\begin{pmatrix} x \\ y \end{pmatrix}$ ist daher $C = \begin{pmatrix} 2 & -1 \\ 1 & -2 \end{pmatrix}$, zu $g\circ f$ gehört $\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}$.

Wir untersuchen allgemeiner $f, g \in \mathrm{Hom}(\mathbb{R}^2, \mathbb{R}^2)$ mit zugehörigen Matrizen $(a_{ij}), (b_{ij})$:

$$\begin{aligned} (f\circ g)(x) = A(Bx) &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \left(\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \\ &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11}x_1 + b_{12}x_2 \\ b_{21}x_1 + b_{22}x_2 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}(b_{11}x_1 + b_{12}x_2) + a_{12}(b_{21}x_1 + b_{22}x_2) \\ a_{21}(b_{11}x_1 + b_{12}x_2) + a_{22}(b_{21}x_1 + b_{22}x_2) \end{pmatrix} \\ &= \begin{pmatrix} (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2 \\ (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Cx \end{aligned}$$

Analog ergibt $(g\circ f)(x) = B(Ax) = \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Def 6.3 Sei $A \in M(m \times r),\ B \in M(r \times n)$. Dann ist $A \circ B = (a_{ij}) \circ (b_{ij}) = (c_{ij}) \in M(m \times n)$ definiert durch $c_{ij} := \sum\limits_{\nu=1}^{r} a_{i\nu} b_{\nu j}$ für $i = 1, \ldots, m$ und $j = 1, \ldots, n$.

Diese *Multiplikation von Matrizen* ist eine Verallgemeinerung der Definition 6.2 auf mehrere Spalten, auch hier gilt die Merkregel *Zeile mal Spalte*.

$A \circ B$ ist nur möglich, falls die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt. Statt $A \circ B$ schreiben wir kürzer AB.

Beispiele: 1) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},\ B = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \Longrightarrow AB = \begin{pmatrix} 4 & 7 \\ 8 & 15 \end{pmatrix},\ BA = \begin{pmatrix} 3 & 4 \\ 11 & 16 \end{pmatrix}$.

2) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \Longrightarrow AB = \begin{pmatrix} \quad & \quad \\ \quad & \quad \end{pmatrix},\ BA = \begin{pmatrix} 1 & 2 & 3 \\ & & \\ & & \end{pmatrix}$.

3) $A = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix},\ B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \Longrightarrow AB = \begin{pmatrix} \quad & \quad \end{pmatrix}$, BA ist nicht definiert.

Wie wir an den Beispielen gesehen haben, ist die Multiplikation von Matrizen nicht kommutativ. Dies ist aber kein Wunder:

Wichtige **Feststellung:** Die Matrizenmultiplikation entspricht genau der Verkettung von linearen Abbildungen: Wenn Matrix A zur Abbildung f und Matrix B zu g gehört, liefert $f \circ g$ die Matrix $A \circ B$.

Satz 6.2 (Assoziativgesetz für Matrizenmultiplikation)

Seien $A \in M(m \times r), B \in M(r \times s), C \in M(s \times n)$. Dann ist $A(BC) = (AB)C$.

Wir geben zwei ganz unterschiedliche Beweise an:

1. Beweis: (Elegant) Nach Satz 6.1 gibt es zu den Matrizen A, B und C eindeutig bestimmte lineare Abbildungen f, g und h. Auch die verketteten Abbildungen $f \circ (g \circ h)$ und $(f \circ g) \circ h$ sind linear, hierzu gehören die Matrizen $A(BC)$ bzw.$(AB)C$. Da die Hintereinanderausführung von Funktionen grundsätzlich assoziativ ist, folgt $f \circ (g \circ h) = (f \circ g) \circ h$ und damit die Behauptung $A(BC) = (AB)C$.

2. Beweis: (Brutal)

$$\begin{aligned}
A(BC) &= (a_{ij})\,((b_{ij})(c_{ij})) = (a_{ij}) \left(\sum_{\nu=1}^{s} b_{i\nu} c_{\nu j} \right) = \left(\sum_{\mu=1}^{r} a_{i\mu} \left(\sum_{\nu=1}^{s} b_{\mu\nu} c_{\nu j} \right) \right) \\
&= \left(\sum_{\mu=1}^{r} \sum_{\nu=1}^{s} a_{i\mu} b_{\mu\nu} c_{\nu j} \right) = \left(\sum_{\nu=1}^{s} \left(\sum_{\mu=1}^{r} a_{i\mu} b_{\mu\nu} \right) c_{\nu j} \right) = \left(\sum_{\mu=1}^{r} a_{i\mu} b_{\mu j} \right) (c_{ij}) \\
&= ((a_{ij})(b_{ij}))\,(c_{ij}) = (AB)C
\end{aligned}$$

Beispiel: Im $\mathbb{R}^2$ seien f die Spiegelung an der y–Achse und g die Spiegelung an der Geraden $y = x$, gesucht sind die Matrizen für die verketteten Abbildungen $f \circ g \circ f$ und $g \circ f \circ g$:

Wegen $f(e_1) = -e_1,\ f(e_2) = e_2,\ g(e_1) = e_2,\ g(e_2) = e_1$ folgt für die zugehörigen Matrizen

$$F = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$\Longrightarrow$ Zu $f \circ g \circ f$ gehört $FGF = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$, zu $g \circ f \circ g$ die Matrix $GFG = \begin{pmatrix} & \\ & \end{pmatrix}$.

Die bekannte und in jedem Körper gültige Regel $xy = 0 \implies x = 0$ oder $y = 0$ gilt *nicht* für die Matrizenmultiplikation:

Beispiel: Was ist $\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} =$??

Man sagt hierzu auch: Matrizen sind *nicht nullteilerfrei.*

Wir schließen diesen inhaltsreichen Abschnitt mit einem interessanten Nebenprodukt unserer teilweise abstrakten Überlegungen.

Satz 6.3 Für sin und cos gelten die Additionsregeln

$$\begin{aligned} \sin(\alpha + \beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta \\ \cos(\alpha + \beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{aligned}$$

Beweis: Wie wir gesehen haben, gehört in der Anschauungsebene zu einer Drehung um den Nullpunkt mit einem beliebigen Winkel γ die Matrix $\begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$. Für eine Drehung um den Winkel $\alpha + \beta$ lautet die entsprechende Matrix

$$M = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}$$

Die gleiche lineare Abbildung und damit die gleiche Matrix M erhalten wir, wenn wir nacheinander um die Winkel β und α drehen. Für die zugehörigen Matrizen folgt dann

$$\begin{aligned} &\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \\ = &\begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -(\sin\alpha\cos\beta + \cos\alpha\sin\beta) \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{pmatrix} \end{aligned}$$

Aus der Gleichheit von M mit der letzten Matrix folgt die Behauptung.

Wir erhalten übrigens die gleiche Matrix, wenn die beiden Drehungen in umgekehrter Reihenfolge vorgenommen werden.

7 Reguläre Matrizen

Lineare Abbildungen und Matrizen gehören eindeutig zusammen[6], wobei der Verkettung der Abbildungen genau die Matrizenmultiplikation entspricht. Da die bijektiven linearen Abbildungen (Automorphismen) $f : \mathbb{R}^n \to \mathbb{R}^n$ für jedes $n \in \mathbb{N}$ mit der Verkettung eine Gruppe bilden, gilt dies auch für die zugehörigen Matrizen mit der Matrizenmultiplikation als Verknüpfung. Wir halten fest:

Def 7.1 und **Satz 7.1** Die zu einem Automorphismus f gehörende Matrix heißt *invertierbar* oder *regulär*. Für jedes $n \in \mathbb{N}$ bilden alle regulären $(n \times n)$–Matrizen bzgl. $\circ$ eine Gruppe.

Neutrales Element in der Matrizengruppe ist die zur Identität gehörende *Einheitsmatrix* E_n, bei der genau die Einträge auf der *Hauptdiagonalen* den Wert 1 und alle anderen den Wert 0 haben.

$$E_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in M(n \times n)$$

Wie üblich bezeichnen wir die zu A *inverse Matrix* mit A^{-1}.

Fragen: Kann man einer Matrix „ansehen“, ob sie regulär ist? Wie sieht die zugehörige inverse Matrix aus? Wir wollen diese Fragen nur für den Fall $n = 2$ beantworten:

Satz 7.2 Sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Dann gelten

a) A regulär $\implies a_{11}a_{22} - a_{12}a_{21} \neq 0$

b) $a_{11}a_{22} - a_{12}a_{21} \neq 0 \implies A$ regulär mit inverser Matrix $A^{-1} = \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

Hierbei bedeutet αA, dass α mit jedem Element der Matrix A multipliziert wird.

Beweis: a) A ist als regulär vorausgesetzt, also ist die zugehörige lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ bijektiv und deshalb gilt nach Satz 4.1 (c) $\text{Kern}\, f = \{(0,0)\}$ $(*)$.

Wir zeigen, dass nicht nur die kanonische Basis (e_1, e_2), sondern auch deren Bilder $(f(e_1), f(e_2))$ linear unabhängig sind.

Seien $\alpha, \beta \in \mathbb{R}$ mit $(0,0) = \alpha f(e_1) + \beta f(e_2)$. Wegen der Linearität von f folgt

$$(0,0) = f(\alpha e_1 + \beta e_2) = f((\alpha, \beta)) , \qquad \text{also} \quad (\alpha, \beta) \in \text{Kern}\, f .$$

Aus $(*)$ folgt dann $\alpha = \beta = 0$, was zu zeigen war.

Damit ist $f(e_1) = (a_{11}, a_{21})$ kein Vielfaches von $f(e_2) = (a_{12}, a_{22})$, insbesondere sind a_{11} und a_{12} garantiert nicht beide 0. Wir können oBdA $a_{12} \neq 0$ annehmen und erhalten

$$(a_{11}, a_{21}) \neq \frac{a_{11}}{a_{12}}(a_{12}, a_{22}) = \left(a_{11}, \frac{a_{11}a_{22}}{a_{12}}\right) \implies a_{21} \neq \frac{a_{11}a_{22}}{a_{12}} \implies a_{12}a_{21} \neq a_{11}a_{22} \quad .$$

b) Sei erneut f die lineare Abbildung zu A. Auch zur Matrix $B = \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$ gehört eine lineare Abbildung, nennen wir sie g. Wir müssen zeigen, dass f bijektiv ist und dass g die zu f inverse Abbildung ist.

[6] Wir beziehen uns weiterhin auf die kanonische Basis. Bei anderen Basen ist der Zusammenhang komplizierter.

Beh. 1: $f \circ g = id$

Beweis: Wir führen die Matrizenmultiplikation $A \circ B$ durch und erhalten durch simples Rechnen

$$A \circ B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \left(\frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \right) = \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

$$= \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & -a_{21}a_{12} + a_{22}a_{11} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2.$$

Weil die Einheitsmatrix zur identischen Abbildung gehört, folgt $f \circ g = id$.[7]

Wir wissen jetzt zwei Dinge: $(f \circ g)((x,y)) = (x,y)$ für jedes Paar (x,y) und $\text{Kern}\,(f \circ g) = \{(0,0)\}$.

Beh. 2: $\text{Kern}\, g = \{(0,0)\}$

Beweis: Für $(x,y) \in \text{Kern}\, g$, d.h., $g((x,y)) = (0,0)$, untersuchen wir $(f \circ g)((x,y))$.

$$(f \circ g)((x,y)) = f(g((x,y))) = f((0,0)) = (0,0).$$

Es folgt $(x,y) \in \text{Kern}\,(f \circ g) = \{(0,0)\}$, also $(x,y) = (0,0)$.

Damit ist (vgl. Satz 4.1 (c)) g injektiv mit $\dim \text{Kern}\, g = 0$.
Aus dem Dimensionssatz folgt $2 = \dim \text{Kern}\, g + \dim \text{Bild}\, g = \dim \text{Bild}\, g$, also ist g auch surjektiv und somit bijektiv mit inverser bijektiver Abbildung g^{-1}.

Beh. 3: f ist bijektiv.

Beweis: $g^{-1} = id \circ g^{-1} = (f \circ g) \circ g^{-1} = f \circ (g \circ g^{-1}) = f$.

Der Rest ist einfach: Weil f bijektiv ist, ist A regulär. Die inverse Matrix A^{-1} gehört zur Abbildung $f^{-1} = g$ und deshalb ist wie gefordert $A^{-1} = B$.

Beispiele: 1) Wir untersuchen im $\mathbb{R}^2$ die Drehung um (0,0) mit $\alpha = 90°$. Es handelt sich um eine bijektive lineare Abbildung (Automorphismus), die zugehörige Matrix
ist $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Wegen $a_{11}a_{22} - a_{12}a_{21} = 1$ ist $A^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (Drehung um (0,0) mit 270°).

2) Wenn wir im $\mathbb{R}^2$ an der Geraden $y = x$ spiegeln, ist $A = A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

$\Big($ Beachte $a_{11}a_{22} - a_{12}a_{21} = -1$, also $A^{-1} = -1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.\Big)$

3) Gesucht ist die inverse Matrix zu $A = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$: Mit $a_{11}a_{22} - a_{12}a_{21} = -3$ ist $A^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$.

Die Kenntnis von inversen Matrizen kann auch bei der Lösung von Gleichungssystemen nützlich sein.

Beispiel: Gesucht sind alle reelle Zahlen x_1 und x_2, die folgende Gleichungen erfüllen:

$$\begin{aligned} 3x_1 + 2x_2 &= 3 \\ 7x_1 + 5x_2 &= 2 \end{aligned} \iff \begin{pmatrix} 3 & 2 \\ 7 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \iff Ax = b\,.$$

[7] Dies heißt leider noch nicht, dass f bijektiv ist, wie man am *Beispiel* $f : \mathbb{R}^3 \to \mathbb{R}^2, (x,y,z) \mapsto (x,y)$ und $g : \mathbb{R}^2 \to \mathbb{R}^3$, $(x,y) \mapsto (x,y,y)$ sieht: Auch hier ist $f \circ g = id$.

A ist regulär, daher existiert A^{-1} und es gilt $Ax = b \iff A^{-1}(Ax) = A^{-1}b \iff E_n x = x = A^{-1}b$.
Im obigen Beispiel erhalten wir die Lösung $x_1 = 11$ und $x_2 = -15$ aus

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 7 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -7 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ -15 \end{pmatrix}.$$

Für Mathematiker sind im Zusammenhang mit dem Lösen von Gleichungen folgende Fragen wichtig:

- Wie erkennt man die Regularität von $A \in M(n \times n)$ für ein beliebiges $n \in \mathbb{N}$ und wie berechnet man gegebenenfalls A^{-1}?
- Was macht man andernfalls (A nicht regulär oder A nicht quadratisch, d.h., die Anzahl der Unbekannten ist verschieden von der Anzahl der Gleichungen)?

Wir merken uns: Lineare Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$ und Matrizen gehören eindeutig zusammen, zu bijektiven linearen Abbildungen passen reguläre Matrizen. Im nächsten Abschnitt werden wir untersuchen, wie der Begriff des Rangs einer linearen Abbildung (zur Erinnerung: Dies ist die Dimension des Bildraumes) auf die zugehörige Matrix übertragen werden kann.

8 Rang einer Matrix

Wir erinnern an $\operatorname{rang} f = \dim \operatorname{Bild} f$ für lineare Abbildungen. Inzwischen wissen wir, dass zu jeder linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ eine Matrix $A = (f(e_1) \ \dots \ f(e_n))$ gehört mit $\operatorname{Bild} f = L(f(e_1), \dots, f(e_n))$. Der Rang von f ist daher nichts anderes als die maximale Anzahl linear unabhängiger Spalten von A, an Stelle von $\operatorname{rang} f$ können wir gleichwertig vom *Spaltenrang* $\operatorname{SR} A$ von A sprechen: $\operatorname{rang} f = \operatorname{SR} A$.

Beispiel: Zu $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}$ gehört $f : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^2 \\ (x, y, z) & \mapsto & (x + 2z, x + y) \end{array} \right.$, es ist $\operatorname{SR} A = 2$.

Was Spalten recht ist, sollte Zeilen billig sein!

Frage: Was ist der *Zeilenrang* $\operatorname{ZR} A$ einer Matrix A? *Antwort*: Die maximale Anzahl linear unabhängiger Zeilen von A.

Beispiel: Was ist der Zeilenrang von $A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}$?

Obwohl die Matrix im Beispiel mehr Spalten als Zeilen hat, stimmen Spalten– und Zeilenrang überein, weiter Beispiele verhalten sich analog.

Satz 8.1 Für jede Matrix $A \in M(m \times n)$ gilt $\operatorname{SR} A = \operatorname{ZR} A$.

Einen **Beweis** findet man beispielsweise in den Lehrbüchern von *Beutelspacher* und *Jänich*, jeweils mit dem Titel *Lineare Algebra.* [8]

Insgesamt gilt für jede lineare Abbildung f mit zugehöriger Matrix A: $\operatorname{rang} f = \operatorname{SR} A = \operatorname{ZR} A$.

[8] Diese Bücher sind insgesamt gut als ergänzende Lektüre für den Stoff dieses Semesters geeignet.

Def 8.1 Sei $A \in M(m \times n)$. Dann heißt $\operatorname{rg} A := \operatorname{rang} f = \operatorname{SR} A = \operatorname{ZR} A$ *Rang* von A.

Wir wollen in einigen Beispielen den Rang bestimmen.

Beispiele: 1) $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$: Wir sehen uns die Zeilen von A an. Da sie linear unabhängig sind, ist $\operatorname{rg} A = 2$. Wir können auch die Spalten betrachten: Da die ersten beiden linear unabhängig sind, ist $\operatorname{rg} A = 2$. *Frage*: Warum ist $\operatorname{rg} A = 3$ unmöglich?

2) $B = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 5 & 7 \\ 2 & 6 & 8 \end{pmatrix}$ Zeilen 1 und 3 linear abhängig, Zeilen 1 und 2 unabhängig, daher $\operatorname{rg} B = 2$.

3) $C = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \end{pmatrix}$, $\operatorname{rg} C = ??$ (einfach)

4) $D = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 1 & 1 & 0 & 2 \\ -1 & 0 & 1 & -1 \end{pmatrix}$, $\operatorname{rg} D = ??$ (nicht so einfach)

Schön wäre es, wenn wir die Matrix D auf eine Gestalt ähnlich zur Matrix C bringen könnten, ohne ihren Rang zu verändern. Welche Manipulationen von Zeilen oder Spalten ändern den Rang nicht?

1. Überlegung: Vertauschung zweier Zeilen [Spalten] ändert den Rang nicht; denn lineare (Un)abhängigkeit hängt nicht von der Reihenfolge der Vektoren ab.

2. Überlegung: Multiplikation einer Zeile [Spalte] mit $\alpha \in \mathbb{R}^* = \mathbb{R}\backslash\{0\}$ ändert den Rang nicht; denn lineare (Un)abhängigkeit hängt nicht von der Vielfachheit einzelner Vektoren ab.

3. Überlegung: Addition des Vielfachen einer Zeile [Spalte] zu einer anderen Zeile [Spalte] ändert den Rang nicht; denn lineare (Un)abhängigkeit hängt nicht von Linearkombinationen der beteiligten Vektoren ab.

Wir wollen von *elementaren Umformungen* sprechen, genauer von *elementaren Zeilenumformungen* EZU bzw. *elementaren Spaltenumformungen* ESU, noch genauer von EZU 1 – EZU 3 bzw. ESU 1 – ESU 3.

Feststellung: Elementare Umformungen ändern den Rang einer Matrix nicht.

Nehmen wir uns noch einmal die Matrix D aus Beispiel 4 vor und führen elementare Umformungen durch:

Beispiel: $\begin{pmatrix} 1 & 2 & 1 & 3 \\ 1 & 1 & 0 & 2 \\ -1 & 0 & 1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Longrightarrow \operatorname{rg} D = 2.$

Frage: Welche Umformungen wurden benutzt?

Wir geben ein „Kochrezept“ an, wie wir elementare Umformungen benutzen können, um den Rang einer beliebigen Matrix zu bestimmen:

1. Sorge für $a_{11} \neq 0$ (Hilfsmittel: EZU 1, ESU 1). Ist dies nicht möglich, so liegt die Nullmatrix mit $\operatorname{rg} A = 0$ vor.
2. Überführe die erste Spalte von A in $\begin{pmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ (alle Umformungen erlaubt, günstig sind EZU 3).

3. Wenn A in $\begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & & & \\ 0 & & A' & \\ 0 & & & \end{pmatrix}$ überführt ist, fahre mit der Teilmatrix A' mit dem ersten Schritt fort, bis eine Matrix erreicht ist, der man den Rang „ansieht“.

Für quadratische Matrizen A kann man auf diese Weise zusätzlich feststellen, ob A regulär ist und ob A^{-1} existiert. Es gilt sogar (ohne Beweis)

Satz 8.2 Sei $A \in M(n \times n)$ regulär. Dann gilt: Die gleichen EZU, die A in E überführen, verwandeln E in A^{-1}.

Beispiele: 1) Die Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix}$ ist wegen $1 \cdot 7 - 2 \cdot 3 = 1 \neq 0$ regulär, nach Satz 7.2 ist $A^{-1} = \begin{pmatrix} 7 & -2 \\ -3 & 1 \end{pmatrix}$. Wir wollen das Ergebnis mit dem neuen Verfahren verifizieren:

$$(A|E) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 7 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & -3 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 7 & -2 \\ 0 & 1 & -3 & 1 \end{array}\right) = (E|A^{-1})$$

2) $(A|E) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 6 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -3 & 1 \end{array}\right) \to \quad ???$

Falls A nicht regulär ist, bricht das Verfahren ohne Ergebnis ab. An Stelle von EZU kann man auch ESU benutzen. Satz 8.2 gilt allerdings nicht, falls EZU und ESU gemischt angewandt werden.

9 Determinanten

Zum Verständnis der folgenden Überlegungen stellen wir uns Vektoren $(a_{11}, a_{21}), (a_{12}, a_{22}) \in \mathbb{R}^2$ als Spalten einer 2×2 – Matrix vor. Wir wissen aus Satz 7.2

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ regulär } \iff a_{11}a_{22} - a_{12}a_{21} \neq 0 \quad .$$

Wir haben gelernt, dass zu jeder Matrix eine lineare Abbildung f gehört, wobei in der i-ten Spalte von A das Bild $f(e_i)$ des i-ten Einheitsvektors e_i steht. A ist regulär, wenn diese Abbildung bijektiv ist. Dies bedeutet, dass die Spaltenvektoren von A linear unabhängig sind. Weil es nach Satz 4.3 auch zu den Vektoren (a_{11}, a_{21}) und (a_{12}, a_{22}) eine lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ gibt mit $f(e_1) = (a_{11}, a_{21})$ und $f(e_2) = (a_{12}, a_{22})$, können wir mit der Abbildung

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto |A| := a_{11}a_{22} - a_{12}a_{21} \ ,$$

der jeder 2×2 – Matrix eine reelle Zahl zuordnet, feststellen, ob die Vektoren (a_{11}, a_{21}) und (a_{12}, a_{22}) linear (un)abhängig sind.

Beispiel: Die Vektoren $(1, 3)$ und $(2, 5)$ sind linear unabhängig, da $\left|\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}\right| = 1 \cdot 5 - 2 \cdot 3 \neq 0$.

Die Vektoren $(1,3)$ und $(2,6)$ sind linear abhängig (selbst nachrechnen!).

Wir werden $|A|$ später als *Determinante* von A kennenlernen. Vorher wollen wir drei interessante Eigenschaften dieser Zuordnung (Abbildung) erwähnen:

1) Genau dann, wenn eine Matrix A aus linear abhängigen Zeilen oder Spalten besteht, ist $|A| = 0$ (siehe oben).

2) Für alle reelle Zahlen α, a_{ij}, a'_{ij} ist wegen $\alpha(a_{11}a_{22} - a_{12}a_{21}) = \alpha a_{11}a_{22} - \alpha a_{12}a_{21}$

$$\alpha \left|\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right| = \left|\begin{pmatrix} \alpha a_{11} & a_{12} \\ \alpha a_{21} & a_{22} \end{pmatrix}\right| = \left|\begin{pmatrix} a_{11} & \alpha a_{12} \\ a_{21} & \alpha a_{22} \end{pmatrix}\right| = \left|\begin{pmatrix} \alpha a_{11} & \alpha a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right| = \left|\begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{21} & \alpha a_{22} \end{pmatrix}\right|$$

und

$$\left|\begin{pmatrix} a_{11} + a'_{11} & a_{12} \\ a_{21} + a'_{21} & a_{22} \end{pmatrix}\right| = \left|\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right| + \left|\begin{pmatrix} a'_{11} & a_{12} \\ a'_{21} & a_{22} \end{pmatrix}\right| ,$$

gleiches gilt für die zweite Spalte und ebenso für beide Zeilen.

3) $\left|\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| = |E_2| = 1.$

Frage: Gibt es auch zu anderen $n \in \mathbb{N}$ eine Abbildung $M(n \times n) \to \mathbb{R}$ mit analogen Eigenschaften?

Die Antwort erfolgt im nächsten Satz, den wir nicht beweisen werden. Gleichzeitig wird die sogenannte *Determinantenfunktion* allgemein vorgestellt und für die Fälle $n = 2$ und $n = 3$ explizit angegeben:

Def 9.1 und **Satz 9.1** Zu jedem $n \in \mathbb{N}$ gibt es genau eine *Determinantenfunktion* $M(n \times n) \to \mathbb{R}$, $A \mapsto |A|$, die die genannten Eigenschaften 1) – 3) erfüllt. Speziell gilt für

$$n = 2: \ |A| = a_{11}a_{22} - a_{12}a_{21}$$

$$n = 3: \ |A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Als Eselsbrücke merken wir uns für $n = 2$: *Hauptdiagonale – Nebendiagonale*. Der Fall $n = 3$ läuft auch unter den Begriffen *Jägerzaunregel* oder *Regel von Sarrus* (wird in der Vorlesung näher erklärt).

Beispiel: Determinante von $\begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix}$ mit der Eselsbrücke $\left(\begin{array}{ccc|cc} 1 & 3 & 0 & 1 & 3 \\ 2 & 1 & 4 & 2 & 1 \\ 0 & 1 & 2 & 0 & 1 \end{array}\right)$:

$$\left|\begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix}\right| = 1 \cdot 1 \cdot 2 + 3 \cdot 4 \cdot 0 + 0 \cdot 2 \cdot 1 - 0 \cdot 1 \cdot 0 - 1 \cdot 4 \cdot 1 - 3 \cdot 2 \cdot 2 = -14 \ .$$

Eine der Behauptungen von Satz 9.1 ist die *Eindeutigkeit* der Determinantenfunktion. Wir wollen uns dies für $n = 1$ explizit vor Augen führen, zumal wir auf diese Weise einiges des bisher Gelernten anwenden können:

Für $n = 1$ ist abgesehen von den Klammern $M(1 \times 1) = \mathbb{R}$. Die Eigenschaft 2) beinhaltet für $n = 1$ genau die beiden Linearitätsbedingungen $\alpha f(r) = f(\alpha r)$ und $f(r_1 + r_2) = f(r_1) + f(r_2)$. Gesucht ist also eine *lineare* Abbildung $\mathbb{R} \to \mathbb{R}$. Wegen 3) soll $f(1) = |(1)| = 1$ erfüllt sein. Da jede lineare Abbildung durch die Bilder der Basisvektoren eindeutig festgelegt ist (siehe Satz 4.3), und weil 1 die kanonische Basis des reellen Vektorraumes $\mathbb{R}^1 = \mathbb{R}$ ist, kann es höchstens eine Abbildung der gesuchten Art geben, nämlich

$$|(r)| = |(r \cdot 1)| = r \cdot |(1)| = r \cdot 1 = r.$$

Es liegt die identische Abbildung vor. Da sie auch die bisher noch nicht beachtete erste Forderung erfüllt (warum?), gibt es für $n = 1$ genau eine Determinantenfunktion.

Beispiel: Die durch $f(x) = -2x$ auf $\mathbb{R}$ definierte Funktion ist bijektiv und linear. Die zugehörige Matrix ist $A = (-2)$ mit Determinante $|A| = |(-2)| = -2$; die Determinantenstriche dürfen nicht mit dem Betrag einer Zahl verwechselt werden!

Die Jägerzaunregel ist nur für $n = 3$ und kein anderes n anwendbar. *Frage*: Wie berechnet man die Determinante bei größerem n?

Def 9.2 Sei $A \in M(n \times n)$. Dann ist $A_{ij} \in M((n-1) \times (n-1))$ die Matrix, die durch Streichen der i–ten Zeile und j–ten Spalte aus A entsteht. (*Streichungsmatrix*)

Beispiel: $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 0 & a & b \\ c & d & e & f \end{pmatrix} \implies A_{23} = \qquad\qquad\qquad , \ (A_{23})_{23} =$

Ohne Beweis geben wir den *Entwicklungssatz* für Determinanten an:

Satz 9.2 Sei $A \in M(n \times n)$. Dann gilt

$$\begin{array}{llll} \text{a)} & |A| = & \sum_{i=1}^{n} (-1)^{i+j} a_{ij}\, |A_{ij}| & \text{für alle } j = 1, \ldots, n \\ \text{b)} & |A| = & \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\, |A_{ij}| & \text{für alle } i = 1, \ldots, n \end{array}$$

Man spricht bei a) von der Entwicklung nach der j–ten Spalte und bei b) von der Entwicklung nach der i–ten Zeile.

Beispiel:

Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ a & b & c \\ x & y & z \end{pmatrix}$, Entwicklung nach der ersten Zeile ($i = 1$ fest):

$$|A| = \sum_{j=1}^{3} (-1)^{1+j} a_{1j}\, |A_{1j}| = 1 \cdot \left| \begin{pmatrix} b & c \\ y & z \end{pmatrix} \right| - 2 \cdot \left| \begin{pmatrix} a & c \\ x & z \end{pmatrix} \right| + 3 \cdot \left| \begin{pmatrix} a & b \\ x & y \end{pmatrix} \right| = \cdots$$

In der Vorlesung soll auch noch nach der zweiten Spalte ($j = 2$ fest) entwickelt werden:

$$|A| = \sum_{i=1}^{3} (-1)^{i+2} a_{i2}\, |A_{i2}| =$$

Der Einsatz des Entwicklungssatzes lohnt sich besonders, wenn es Zeilen oder Spalten mit vielen Nullen gibt. Dieser angenehme Fall lässt sich zum Glück mit relativ einfachen Mitteln herbeiführen. Zum besseren Verständnis des Folgenden dient

Beh.: $\left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| = \left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + \alpha a_{11} & a_{22} + \alpha a_{12} \end{pmatrix} \right|$. *Beweis*: Nachrechnen!

Wir haben soeben das α – fache der ersten Zeile zur zweiten Zeile addiert. Allgemein gilt für jedes $n \in \mathbb{N}$:

- Der Wert einer Determinante ändert sich nicht, wenn zu einer Zeile [Spalte] einer Matrix das Vielfache einer anderen Zeile [Spalte] addiert oder subtrahiert wird.
- Multipliziert man eine Zeile [Spalte] mit einer reellen Zahl α, ändert sich die Determinante um den Faktor α.
- Vertauscht man zwei Zeilen [Spalten], ändert sich das Vorzeichen der Determinante.

Hier sind wieder uns die elementaren Umformungen aus Kapitel 4.8 begegnet!

Beispiel: $A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$. Da in der zweiten Spalte zwei Nullen stehen, entwickelt man nach dieser Spalte:

$$|A| = -0 \cdot \left|\cdots\right| + 1 \cdot \left|\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\right| - 0 \cdot \left|\cdots\right| + 1 \cdot \left|\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix}\right| = \cdots = 1.$$

Für das richtige Vorzeichen im Entwicklungssatz bietet das *Schachbrettmuster* eine Eselsbrücke:

$$\left|\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & \cdots & \\ + & - & \cdots & & \\ \vdots & & & & \end{pmatrix}\right| \qquad \text{oder} \qquad \text{Vorzeichen } a_{ij} = \begin{cases} + & \text{falls} \quad i+j \text{ gerade} \\ - & \text{falls} \quad i+j \text{ ungerade} \end{cases}$$

Beispiele: 1) Die Determinante von $A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ soll auf verschiedene Weisen berechnet werden.

a) *Regel von Sarrus*: $|A| = 1 \cdot 2 \cdot 3 + 0 + 2 \cdot 3 \cdot 2 - 2 \cdot 2 \cdot 1 - 1 \cdot 1 \cdot 2 - 0 = 12$

b) *Entwicklung* nach erster Zeile: $|A| = 1 \left|\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}\right| + 2 \left|\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}\right| = (6-2) + 2(6-2) = 12$

c) Anwendung von Zeilenoperationen, um durch Umwandlung eine *Dreiecksmatrix* zu erhalten, deren Determinante einfach berechnet werden kann:

$$|A| = \left|\begin{pmatrix} 1 & 0 & 2 \\ 3 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix}\right| = \left|\begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & -5 \\ 0 & 2 & 1 \end{pmatrix}\right| = \left|\begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & -5 \\ 0 & 0 & 6 \end{pmatrix}\right| = 1 \cdot 2 \cdot 6 = 12$$

2) Im folgenden *Beispiel* subtrahieren wir zuerst die zweite von der fünften Zeile, um auch an der Stelle a_{52} den Eintrag 0 zu haben. Dann entwickeln wir nach dieser Spalte:

$$|A| = \left|\begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 2 & 3 & 0 & 0 & 1 \\ 1 & 0 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ 1 & 3 & -2 & 2 & 1 \end{pmatrix}\right| = \left|\begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 2 & 3 & 0 & 0 & 1 \\ 1 & 0 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ -1 & 0 & -2 & 2 & 0 \end{pmatrix}\right| = 3 \cdot \left|\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 5 \\ 0 & 1 & 2 & 3 \\ -1 & -2 & 2 & 0 \end{pmatrix}\right|$$

Erneut zwei Zeilenoperationen und Vertauschen der neuen Zeilen 2,3,4 liefert eine Dreiecksmatrix, deren Determinante durch Entwicklung nach der jeweils ersten Zeile leicht bestimmt werden kann:

$$|A| = 3 \cdot \left| \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 5 & 4 \end{pmatrix} \right| = 3 \cdot (-1)^2 \cdot \left| \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right| = 3 \cdot 1 \cdot 1 \cdot 5 \cdot 1 = 15$$

Ohne Beweis schließen wir mit dem nützlichen *Determinantenmultiplikationssatz*:

Satz 9.3 $A, B \in M(n \times n) \implies |AB| = |A||B|$

Hieraus folgt

Korollar A regulär $\implies |A^{-1}| = |A|^{-1}$.

Beweis: $1 = |E| = |A^{-1}A| = |A^{-1}||A| \implies |A^{-1}| = |A|^{-1}$.

10 Lineare Gleichungssysteme

Nachdem wir schon mehrfach Lösungen für Gleichungen gesucht haben, soll diese Problematik jetzt systematisch untersucht werden. Um uns auf das Wesentliche konzentrieren zu können, bleiben wir bei dem Körper der reellen Zahlen, obwohl die Theorie für beliebige Körper gültig ist. Ferner identifizieren wir die Menge der entarteten Matrizen mit nur einer Zeile oder einer Spalte mit der entsprechenden Menge $\mathbb{R}^k$.

Auf das folgende *Beispiel* wird mehrfach zurückgegriffen: Gesucht sind alle reelle Zahlen x_1, x_2, x_3 mit

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 1 \\ x_1 - 2x_2 + 2x_3 &= 0 \end{aligned}$$

Unter Verwendung der Matrizenrechnung sieht unser Problem folgendermaßen aus:

$$\begin{pmatrix} 2 & 3 & 1 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{oder allgemein} \qquad Ax = b, \quad x = (x_1, x_2, x_3)$$

Def 10.1 Sei $A \in M(m \times n)$, $b \in M(m \times 1) = \mathbb{R}^m$. $Ax = b$ heißt *lineares Gleichungssystem* (LGS) für ein unbekanntes $x \in M(n \times 1) = \mathbb{R}^n$.

Ist b der Nullvektor, heißt das System *homogen*, andernfalls *inhomogen*.

Ausführlich aufgeschrieben sieht ein LGS folgendermaßen aus:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

Hierbei sind die Einträge der Matrix A und b bekannt, gesucht sind die Komponenten $x_1, \ldots, x_n$ von x.

In unserem *Beispiel* handelt es sich um ein inhomogenes LGS. Im Folgenden bezeichnen wir die gesuchte Menge aller Lösungen von $Ax = b$ mit $\text{Lös}\,(A, b)$ und nennen ein LGS *lösbar*, falls $\text{Lös}\,(A, b) \neq \emptyset$.

Beispiel: Das Gleichungssystem $\begin{array}{rcl} 2x_1 + 3x_2 + \ z_3 & = & 1 \\ x_1 - 2x_2 + 2x_3 & = & 0 \end{array}$ ist lösbar, es gilt $(-2, 1, 2) \in \text{Lös}\,(A, b)$.

Nicht jedes LGS ist lösbar, wie uns das *Beispiel* $\begin{pmatrix} 1 \\ 2 \end{pmatrix} x = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ zeigt.

Um ein Kriterium für die Lösbarkeit von LGS zu bekommen, beschäftigen wir uns noch einmal mit dem Begriff Untervektorraum. Für $A \in M(m \times n, \mathbb{R})$ ist $\text{Lös}\,(A, b) \subseteq \mathbb{R}^n$. Wann ist dies ein Untervektorraum, d.h., wann gilt $\text{Lös}\,(A, b) \leq \mathbb{R}^n$?

Im folgenden Satz sei mit o wie üblich der Nullvektor gemeint.

Satz 10.1 $\text{Lös}\,(A, b) \leq \mathbb{R}^n \iff b = o \quad (\iff \text{LGS ist homogen})$

Beweis: „$\Rightarrow$“: $\text{Lös}\,(A, b)$ Untervektorraum $\implies o \in \text{Lös}\,(A, b) \implies b = Ao = o$.

„$\Leftarrow$“: Für $b = o$ ist $\text{Lös}\,(A, b) = \text{Lös}\,(A, o) = \{x \in \mathbb{R}^n \mid Ax = o\} = \text{Kern}\,A \leq \mathbb{R}^n$ (Satz 4.1).

Homogene LGS besitzen stets die Lösung $x = o$, inhomogene LGS müssen nicht lösbar sein.

Def 10.2 Für das LGS $Ax = b$ heißt die Matrix $(A, b) := \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$ *erweiterte Matrix*, für $A \in M(m \times n)$ ist $(A, b) \in M(m \times (n + 1))$.

In unserem *Beispiel* ist die erweiterte Matrix $(A, b) = \begin{pmatrix} 2 & 3 & 1 & 1 \\ 1 & -2 & 2 & 0 \end{pmatrix}$.

Wie bereits festgestellt, ist $x = (-2, 1, 2)$ eine Lösung dieses LGS:

$$b = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot (-2) + 3 \cdot 1 + 1 \cdot 2 \\ 1 \cdot (-2) - 2 \cdot 1 + 2 \cdot 2 \end{pmatrix} = (-2) \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ -2 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Weil eine Lösung existiert, kann b mit Hilfe dieser Lösung als Linearkombination der Spaltenvektoren von A geschrieben werden. Mit anderen Worten: b ist linear abhängig von den Spalten der Matrix A oder noch anders ausgedrückt: b liegt in der linearen Hülle der Spaltenvektoren von A. Durch Hinzunahme von b zu den Spaltenvektoren von A (= Übergang zur erweiterten Matrix) wird der Rang von A (= maximale Anzahl der linear unabhängigen Spalten, Einzelheiten im Abschnitt 4.8) nicht verändert.

Wir fassen dies im folgenden Satz zusammen, der nicht formal bewiesen wird:

Satz 10.2 $Ax = b$ lösbar $\iff \text{rg}\,A = \text{rg}\,(A, b)$

Beispiele: 1) $\begin{array}{rcl} 2x + 3y + \ z & = & 1 \\ x - 2y + 2z & = & 0 \end{array}$: $\quad \text{rg}\,A = \text{rg} \begin{pmatrix} 2 & 3 & 1 \\ 1 & -2 & 2 \end{pmatrix} = 2 = \text{rg}\,(A, b)$.

2) $\begin{pmatrix} 1 \\ 2 \end{pmatrix} x = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$: $\quad$ $\text{rg } A = 1 \neq \text{ rg } (A, b) = \text{ rg } \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix} = 2$, daher ist dieses LGS nicht lösbar.

Mit $(-2, 1, 2)$ kennen wir eine Lösung des ersten Beispiels, aber auch $(6, -2, -5)$ oder $(-802, 301, 702)$ sind Lösungen.

Wir sind an *allen* Lösungen interessiert.

Satz 10.3 Sei x_0 eine beliebige Lösung von $Ax = b$. Dann gilt

$$\text{Lös}\,(A, b) = x_0 + \text{ Kern } A = \{x_0 + x \mid Ax = 0\}$$

Beweis: „$\subseteq$": Für jede Lösung v gilt $A(v - x_0) = Av - Ax_0 = b - b = o$. Also ist $v - x_0$ aus dem Kern von A bzw. $v \in x_0 +$ Kern A.

„$\supseteq$": Sei $x \in \text{Kern}\, A \implies A(x_0 + x) = Ax_0 + Ax = b + o = b$, damit ist $x_0 + x \in \text{Lös}\,(A, b)$.

Feststellung: Wir erhalten *alle* Lösungen eines LGS $Ax = b$, wenn wir zu *einer* speziellen Lösung *alle* Lösungen des zugehörigen homogenen Systems (= Kern von A) addieren.

Beispiele: 1) Gesucht sind alle Lösungen (x, y) von $x - 2y = 5 \iff (1 \ \ -2) \begin{pmatrix} x \\ y \end{pmatrix} = (5)$: Wegen $\text{rg}\, A = \text{rg}\,(1 \ \ -2) = 1 = \text{rg}\,(A, b) = (1 \ \ -2 \ \ 5)$ existieren Lösungen. Eine spezielle Lösung ist $(5, 0)$ (in die umgeformte Gleichung $x = 5 + 2y$ wurde $y = 0$ eingesetzt). Wir bestimmen den Kern: $\text{Kern}\, A = \{(x, y) \mid x - 2y = 0\} = \{(2y, y) \mid y \in \mathbb{R}\} = L((2, 1))$. Die gesuchte Lösungsmenge ist $\text{Lös}\,(A, b) = (5, 0) + L((2, 1))$.

2) In unserem Standardbeispiel ist $(-2, 1, 2) \in \text{Lös}\,(A, b)$, gesucht ist $\text{Kern}\, A = \{x \in \mathbb{R}^3 \mid Ax = 0\}$. Durch Rechnung (hierzu mehr im nächsten Kapitel) erhält man $\text{Kern}\, A = L((-8, 3, 7))$, damit ist $\text{Lös}\,(A, b) = (-2, 1, 2) + L((-8, 3, 7))$.

Frage: Was ist die geometrische Gestalt dieser Lösungsmenge?

Mit Satz 10.2 haben wir ein notwendiges und hinreichendes Kriterium für die Existenz von Lösungen eines LGS kennengelernt. Gibt es auch Kriterien für *eindeutige* Lösbarkeit?

Satz 10.4 Sei $A \in M(m \times n)$. Dann gilt

$$Ax = b \text{ eindeutig lösbar} \quad \iff \quad \text{rg } A = \text{ rg } (A, b) = n$$

Beweis: Wir benutzen die Dimensionsformel für lineare Abbildungen (A kann als lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ aufgefasst werden): $n = \dim \text{ Kern } A + \dim \text{ Bild } A = \dim \text{ Kern } A + \text{ rg } A$.

„$\Rightarrow$": Nach Satz 10.3 bedeutet die eindeutige Lösbarkeit, dass der Kern von A nur aus dem Nullvektor besteht, also die Dimension 0 hat; aus der Formel folgt dann $n = \text{ rg } A$. Da wenigstens eine Lösung existiert, ist nach Satz 10.2 $\text{rg}\, A = \text{rg}\,(A, b)$.

„$\Leftarrow$": Wegen $\text{rg}\, A = \text{rg}\,(A, b)$ ist das LGS lösbar, sei $x_0 \in \text{Lös}\,(A, b)$. Wegen $\text{rg}\, A = n$ ist die Dimension des Kerns 0. Nach Satz 10.3 ist damit $\text{Lös}\,(A, b) = \{x_0\}$.

In unserem *Standardbeispiel* gilt $\text{rg}\, A = 2 = \text{rg}\,(A, b) \neq 3 = n$, es ist nicht eindeutig lösbar.

Generell gilt : Enthält ein LGS $Ax = b$ mehr Unbekannte als Gleichungen, kann es nie eindeutig lösbar sein, denn $\text{rg}\, A$ kann nicht größer sein als die Anzahl m der Gleichungen, und es ist $m < n$ (Anzahl der Unbekannten).

Wenn die Anzahl der Gleichungen mit der Anzahl der Unbekannten übereinstimmt, ist die zugehörige Matrix quadratisch und besitzt deshalb eine Determinante. In diesem Spezialfall erhalten wir wegen Satz 9.1 ein weiteres Kriterium für eindeutige Lösbarkeit:

Satz 10.5 Sei $A \in M(n \times n)$. Dann gilt $Ax = b$ eindeutig lösbar $\iff$ $|A| \neq 0$.

Beispiele: 1) $\begin{array}{rcl} 3x+2y &=& \alpha \\ 4x+4y &=& \beta \end{array}$: Es ist $m = n = 2$, $|A| = 3\cdot 4 - 4\cdot 2 \neq 0 \implies$ Das LGS ist unabhängig von speziellen Werten α und β stets eindeutig lösbar.

2) $\begin{array}{rcl} 3x+2y &=& \alpha \\ 6x+4y &=& \beta \end{array}$: $|A| = 3\cdot 4 - 6\cdot 2 = 0 \implies$ Das LGS ist auf keinen Fall eindeutig lösbar. Wegen $\operatorname{rg} A = 1$ und $\operatorname{rg}(A, b) = \operatorname{rg}\begin{pmatrix} 3 & 2 & \alpha \\ 6 & 4 & \beta \end{pmatrix}$ ist es für $\beta \neq 2\alpha$ nicht lösbar und für $\beta = 2\alpha$ mehrdeutig lösbar.

11 Über das Lösen von linearen Gleichungssystemen

Während es im vorherigen Kapitel im Wesentlichen um die Theorie ging, sollen jetzt LGS konkret gelöst werden.

Satz 11.1 *Cramersche Regel*

Sei das LGS $Ax = b$ mit $A \in M(n\times n)$ und $|A| \neq 0$ gegeben. Dann gilt für die Unbekannte $x = (x_1, \ldots, x_n)$

$$x_i = \frac{|A_i|}{|A|} \quad \text{mit} \quad A_i := \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix},$$

wobei A_i aus A entsteht, indem man die i–te Spalte von A durch b ersetzt.

Beweis: Wir wissen aus Satz 10.5 des letzten Kapitels, dass $Ax = b$ eindeutig lösbar ist:

$$\exists x_1, \ldots, x_n \in \mathbb{R}: \; x_1 s_1 + \ldots + x_n s_n = b$$

Hierbei sei s_j die j–te Spalte der Matrix A. Eine einfache Umformung und ein kleiner Trick in der Schreibweise ergeben

$$x_1 s_1 + \ldots + 1 \cdot (x_j s_j - b) + \ldots + x_n s_n = o \qquad \text{(Nullvektor)}$$

Die Vektoren $s_1, \ldots, s_{j-1}, x_j s_j - b, s_{j+1}, \ldots, s_n$ sind also linear abhängig, damit gilt für die zugehörige Determinante

$$\begin{vmatrix} s_1 & \cdots & x_j s_j - b & \cdots & s_n \end{vmatrix} = 0$$

$\Bigg($ Man beachte, dass für die j–te Spalte gilt $\quad x_j s_j - b = x_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} - \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} x_j a_{1j} - b_1 \\ \vdots \\ x_j a_{nj} - b_n \end{pmatrix}. \Bigg)$

Wir entwickeln die Determinante nach dieser j–ten Spalte:

$$\begin{aligned} 0 &= \sum_{i=1}^{n}(-1)^{i+j}(x_j a_{ij} - b_i)|A_{ij}| \\ &= x_j \sum_{i=1}^{n}(-1)^{i+j} a_{ij}|A_{ij}| - \sum_{i=1}^{n}(-1)^{i+j} b_i|A_{ij}| \\ &= x_j|A| - |A_j| \qquad \Rightarrow \qquad \text{Behauptung} \end{aligned}$$

Beispiel: $\begin{aligned} 3x + 2y &= 1 \\ 4x + 3y &= 2 \end{aligned} \qquad \Longrightarrow \qquad x = \dfrac{\left|\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\right|}{\left|\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}\right|} = \quad , \quad y = \dfrac{\left|\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}\right|}{\left|\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}\right|} =$

Die Cramersche Regel hat zwei große Nachteile: Sie ist nur für reguläre Matrizen anwendbar und der Rechenaufwand nimmt bei größeren n in einem unerträglichen Maß zu. Um ein allgemein anwendbares Verfahren zu finden, erinnern wir uns an das Kochrezept zur Rangbestimmung einer Matrix, denn LGS mit zugehörigen Matrizen in Dreiecksgestalt lassen sich leicht lösen:

Beispiel: $\begin{aligned} x - y + 2z &= 5 \\ 2y - z &= 4 \\ 3z &= 6 \end{aligned} \qquad \Longrightarrow \qquad \begin{aligned} z &= 2 \\ 2y &= 4 + z \\ x &= 5 - 2z + y \end{aligned} \qquad \Longrightarrow \qquad \begin{aligned} z &= 2 \\ y &= 3 \\ x &= 4 \end{aligned}$

Satz 11.2 Sei $Ax = b$ ein LGS, durch EZU sei aus der erweiterten Matrix (A, b) die erweiterte Matrix (A', b') entstanden. Dann gilt $\text{Lös}\,(A, b) = \text{Lös}\,(A', b')$.

Der **Beweis** dieses Satzes ist anschaulich klar (auf beiden Seiten einer Gleichung werden dieselben Operationen durchgeführt).

Die sinnvolle Anwendung des Satzes ist der *Gaußsche Algorithmus*, den wir zuerst für den Spezialfall $A \in M(n \times n)$ und $|A| \neq 0$ kennenlernen wollen.

Problem: Gesucht ist die eindeutige Lösung x von $Ax = b$ mit $A \in M(n \times n)$ regulär.

Lösung: 1) Mit Hilfe des „Kochrezepts“ (Seite 93) überführen wir (A, b) durch EZU in (A', b').

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \end{pmatrix} \begin{pmatrix} a'_{11} & * & \cdots & \cdots & * & b'_1 \\ 0 & a'_{22} & * & \cdots & * & b'_2 \\ \vdots & 0 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & * & \vdots \\ 0 & 0 & \cdots & 0 & a'_{nn} & b'_n \end{pmatrix}$$

Da A als regulär vorausgesetzt war, ist $a'_{ii} \neq 0$ für alle i.

2) Wir bestimmen $(x_1, \ldots, x_n)$ „rückwärts“:

$$x_n = \frac{b'_n}{a'_{nn}}, \quad x_{n-1} = \frac{1}{a'_{n-1,n-1}}\left(b'_{n-1} - a'_{n-1,n}x_n\right), \qquad \text{usw.}$$

Allgemein gilt für $k = n-1, n-2, \ldots, 1:$ $\quad x_k = \frac{1}{a'_{kk}}\left(b'_k - \sum\limits_{\nu=k+1}^{n} a'_{k\nu}x_\nu\right).$

Durch weitere Umformungen kann man (A, b) auf die Gestalt (E, b'') mit Einheitsmatrix E bringen. Die Lösung erhält man dann recht einfach durch $x_i = b''_i$. Welche Methode günstigster ist, hängt häufig von weiteren (Neben)bedingungen ab. Bei umfangreichen Systemen ist es enorm wichtig, den Rechenaufwand zu optimieren. Solche Fragestellungen werden neben vielen anderen in der *Numerischen Mathematik* oder auch in der *Komplexitätstheorie* untersucht.

Beispiel:
$$\begin{array}{rcr} 2x_2 - 2x_3 &=& 3 \\ x_1 + \ x_2 + \ x_3 &=& 1 \\ x_1 + 3x_2 + 7x_3 &=& -1 \end{array} \iff (A,b) = \begin{pmatrix} 0 & 2 & -2 & 3 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 7 & -1 \end{pmatrix}$$

Durch EZU erhalten wir aus (A, b) die Matrix $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 3 \\ 0 & 0 & 8 & -5 \end{pmatrix}$. Die gesuchte Lösung ist

$x_3 = -\frac{5}{8}, \quad x_2 = \frac{1}{2}\left(3 - (-2)\frac{-5}{8}\right) = \frac{7}{8}, \quad x_1 = 1 - x_2 - x_3 = \frac{3}{4}.$

Wir kommen zum allgemeinen Fall $A \in M(m \times n)$ und versuchen die gleiche Vorgehensweise wie im Spezialfall an einem *Beispiel*:

$\begin{pmatrix} 1 & 2 & -1 & -1 \\ 3 & 6 & 1 & -1 \\ 2 & 4 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ ? \end{pmatrix}$, wobei wir für das Fragezeichen erst später eine konkrete Zahl einsetzen werden. Durch EZU erhalten wir die Matrix (A', b')

$$\begin{pmatrix} 1 & 2 & -1 & -1 & 2 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & 0 & 2 & 1 & ?? \end{pmatrix}$$

Der nächste Schritt (Ziel: $a'_{22} \neq 0$) ist durch erlaubte EZU nicht möglich. Wir ignorieren diesen Mangel und gelangen mit weiteren EZU zur sogenannten *Zeilenstufenform*

$$\begin{pmatrix} 1 & 2 & -1 & -1 & 2 \\ 0 & 0 & 4 & 2 & -1 \\ 0 & 0 & 0 & 0 & ??? \end{pmatrix}$$

An den Stellen x_1 und x_3 liegen Stufen vor, durch die Stellen x_2 und x_4 werden *freie Parameter* ins Spiel gebracht, ihre Anzahl ist stets gleich der Dimension von Kern A.

Zwei verschiedene Fälle können eintreten:

1. Fall: $??? \neq 0$: Das gegebene LGS $Ax = b$ ist nicht lösbar.

2. Fall: $??? = 0$: Das LGS ist lösbar.

Wenn wir in unserem Beispiel $? = \frac{7}{2}$ setzen, erhalten wir $??? = 0$. Es ist

x_4:	keine Stufe	$\Rightarrow$	frei wählbar, $x_4 = \lambda_1$		
x_3:	Stufe	$\Rightarrow$	$4x_3 = -1 - 2x_4$	$\Rightarrow$	$x_3 = -\frac{1}{4} - \frac{1}{2}\lambda_1$
x_2:	keine Stufe	$\Rightarrow$	frei wählbar, $x_2 = \lambda_2$		
x_1:	Stufe	$\Rightarrow$	$x_1 = 2 - 2x_2 - (-1)x_3 - (-1)x_4$	$\Rightarrow$	$x_1 = \frac{7}{4} - 2\lambda_2 + \frac{1}{2}\lambda_1$

$$\Longrightarrow \text{Lös}\,(A,b) = \left\{\left(\tfrac{7}{4} - 2\lambda_2 + \tfrac{1}{2}\lambda_1, \lambda_2, -\tfrac{1}{4} - \tfrac{1}{2}\lambda_1, \lambda_1\right) \mid \lambda_1, \lambda_2 \in \mathbb{R}\right\}$$
$$= \left\{\left(\tfrac{7}{4}, 0, -\tfrac{1}{4}, 0\right) + \left(-2\lambda_2 + \tfrac{1}{2}\lambda_1, \lambda_2, -\tfrac{1}{2}\lambda_1, \lambda_1\right) \mid \lambda_1, \lambda_2 \in \mathbb{R}\right\}$$
$$= \left(\tfrac{7}{4}, 0, -\tfrac{1}{4}, 0\right) + L\left((1,0,-1,2), (-2,1,0,0)\right).$$

Beispiel:

$$\begin{array}{rcr} x_1 - 2x_2 + 3x_3 + 5x_4 - 4x_5 & = & 2 \\ 2x_1 - 4x_2 + 6x_3 + 5x_4 + 2x_5 & = & -6 \\ 2x_1 - 4x_2 + 7x_3 + 7x_4 + 3x_5 & = & -7 \\ -x_1 + 2x_2 - 2x_3 - 3x_4 + 5x_5 & = & -3 \end{array} \Longrightarrow (A,b) = \left(\begin{array}{rrrrr|r} 1 & -2 & 3 & 5 & -4 & 2 \\ 2 & -4 & 6 & 5 & 2 & -6 \\ 2 & -4 & 7 & 7 & 3 & -7 \\ -1 & 2 & -2 & -3 & 5 & -3 \end{array}\right)$$

Durch EZU erhält man irgendwann die Zeilenstufenform

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 3 & 5 & -4 & 2 \\ 0 & 0 & 1 & 0 & 5 & -5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Wir stellen fest, daß das LGS lösbar ist mit Stufen für x_1, x_3, x_4. Durch die anderen Stellen x_2, x_5 kommen zwei freie Parameter in unserer Lösungsmenge vor. Wir bestimmen eine beliebige Lösung $x = (x_1, \ldots, x_5) \in \text{Lös}\,(A,b)$:

x_5:	keine Stufe	$\Rightarrow$	frei wählbar, $x_5 = \lambda_1$
x_4:	Stufe	$\Rightarrow$	$x_4 = 2 - (-2)x_5 = 2 + 2\lambda_1$
x_3:	Stufe	$\Rightarrow$	$x_3 = -5 - 5x_5 = -5 - 5\lambda_1$
x_2:	keine Stufe	$\Rightarrow$	frei wählbar, $x_2 = \lambda_2$
x_1:	Stufe	$\Rightarrow$	$x_1 = 2 - (-2)x_2 - 3x_3 - 5x_4 - (-4)x_5 = 7 + 2\lambda_2 + 9\lambda_1$

$$\Longrightarrow \text{Lös}\,(A,b) = \{(7 + 2\lambda_2 + 9\lambda_1, \lambda_2, -5 - 5\lambda_1, 2 + 2\lambda_1, \lambda_1) \mid \lambda_1, \lambda_2 \in \mathbb{R}\}$$

oder $\text{Lös}\,(A,b) = (7,0,-5,2,0) + L((9,0,-5,2,1),(2,1,0,0,0))$

Die Darstellung der Lösungsmenge ist nicht eindeutig, da die spezielle Lösung und die Basis des Kernes anders gewählt werden können.

Beispiel: $\text{Lös}\,(A,b) = (0,1,0,0,-1) + L((7,-1,-5,2,1),(11,1,-5,2,1))$ ist ebenfalls Lösung des vorherigen Beispiels, wobei es sich natürlich um die gleiche Lösungsmenge handelt.

LGS, die „ähnlich“ aussehen, müssen nicht „ähnliche“ Lösungen besitzen:

Frage: Welche Lösungen besitzen die LGS

$$\begin{pmatrix} 1 & 1 \\ 1 & 1.001 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 1 \\ 1 & 1.001 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2.001 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 \\ 1 & 1.001 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2.01 \end{pmatrix} \qquad ?$$

Genaues Rechnen ist manchmal dringend erforderlich!

Wir beenden diesen Abschnitt mit einer Aufgabe, die aus einem Rechenbuch aus dem alten China stammt und hier verkürzt wiedergegeben werden soll:

Frage: Wieviele Hähne, Hennen und Küken kann man für (genau) 100 Münzen kaufen, wenn man 100 Vögel haben will und ein Hahn fünf Münzen, eine Henne vier Münzen und vier Küken eine Münze kosten?

Es handelt sich um ein LGS $Ax = b$ mit einer *Nebenbedingung*, denn man ist schließlich nur an ganzen Tieren interessiert – gesucht ist $\text{Lös}\,(A, b) \cap (\mathbb{N} \times \mathbb{N} \times \mathbb{N})$.

Lösung: Mit den Abkürzungen Ha, Hu und K erhalten wir das LGS

$$\begin{array}{rcrcrcl} Ha & + & Hu & + & K & = & 100 \\ 5Ha & + & 4Hu & + & \frac{1}{4}K & = & 100 \end{array}$$

Wir formen (A, b) in Zeilenstufenform um:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 100 \\ 5 & 4 & \frac{1}{4} & 100 \end{array}\right) \quad \rightarrow \quad \left(\begin{array}{ccc|c} 1 & 1 & 1 & 100 \\ 0 & 1 & \frac{19}{4} & 400 \end{array}\right)$$

Damit ist $K = \lambda$ (freier Parameter), $Hu = 400 - \frac{19}{4}\lambda$, $Ha = 100 - 400 + \frac{19}{4}\lambda - \lambda = -300 + \frac{15}{4}\lambda$.

Aus der Nebenbedingung folgt

$$\begin{array}{rccccc} & 4|\lambda & \wedge & 400 \geq \frac{19}{4}\lambda \quad \wedge \quad 300 \leq \frac{15}{4}\lambda \\ \Longleftrightarrow & \lambda = 4l,\ l \in \mathbb{N} & \wedge & \frac{300}{15} \leq l \leq \frac{400}{19} \\ \Longleftrightarrow & l \in \mathbb{N} & \wedge & 20 \leq l < 21{,}06 \end{array}$$

Damit kommt nur $l \in \{20, 21\}$ in Frage und wir erhalten genau zwei Lösungen $(0, 20, 80)$ und $(15, 1, 84)$.

12 Spezielle Eigenschaften der Anschauungsebene $\mathbb{R}^2$

Es geht ausschließlich um den reellen Vektorraum $\mathbb{R}^2$, dessen Elemente jetzt auch Punkte genannt und mit großen Buchstaben geschrieben werden, beispielsweise $A = (a_1, a_2)$, $B, P, X \in \mathbb{R}^2$. Mit $A + B$ ist die übliche Vektorraumaddition gemeint.

Im Kapitel 4.1 haben wir alle eindimensionalen Untervektorräume von $\mathbb{R}^2$ kennengelernt und unter Berücksichtigung unseres Schulwissens festgestellt, dass es sich genau um die Geraden durch $O := (0, 0)$ handelt. Jetzt wollen wir Geraden „offiziell“ einführen.

Def 12.1 Seien $A, B \in \mathbb{R}^2$ verschiedene Punkte. Dann heißt

$g_{A,B} := \{X = A + \lambda(B - A) \in \mathbb{R}^2 \mid \lambda \in \mathbb{R}\}$ *Gerade* des $\mathbb{R}^2$.

Man nennt A auch einen *Aufpunkt* und $B - A$ einen *Richtungsvektor* der Geraden. Weil $g_{A,B}$ von der Variablen λ abhängt, nennt man $g_{A,B}$ auch *Parameterform* einer Geraden.

Manchmal findet man statt $g_{A,B}$ auch die Bezeichnungen AB oder $< A, B >$. Letzteres ist nicht zu empfehlen, da $< , >$ üblicherweise (wie später auch in dieser Vorlesung) für einen anderen Sachverhalt benötigt wird.

Fragen: 1) Wo liegen die Punkte mit $\lambda = 0, 1, \frac{1}{2}$?

2) Warum heißt es *ein* und nicht *der* Aufpunkt bzw. Richtungsvektor?

Wir stellen einige einfache Tatsachen über Geraden zusammen. Bei den folgenden Sätzen sind A und B stets *verschiedene* Punkte.

Satz 12.1 Für alle Geraden gilt $A, B \in g_{A,B}$ und $g_{A,B} = g_{B,A}$

Beweis: Für $\lambda = 0$ ist $A = A + 0 \cdot (B - A) \in g_{A,B}$, für $\lambda = 1$ ist $B = A + 1 \cdot (B - A) \in g_{A,B}$.

Sei $X \in g_{A,B} \implies \exists \alpha \in \mathbb{R}: \ X = A + \alpha(B - A) = B + (1 - \alpha)(A - B) \in g_{B,A}$, damit ist $g_{A,B} \subseteq g_{B,A}$. Die umgekehrte Teilmengenbeziehung zeigt man analog, man vertausche nur A und B.

Satz 12.2 Für alle Geraden $g_{A,B}$ sind folgende Aussagen äquivalent:

(1) $O = (0, 0) \in g_{A,B}$

(2) A, B linear abhängig

(3) $B - A \in g_{A,B}$

(4) $g_{A,B} \leq \mathbb{R}^2$ ($\leq$ bedeutet Untervektorraum)

Beweis: Wir beweisen (4) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1) $\Rightarrow$ (4), durch diesen Ringschluss sind alle Äquivalenzen gezeigt.

(4) $\Rightarrow$ (3): Aus Satz 12.1 wissen wir $A, B \in g_{A,B}$. Auf Grund der Voraussetzung Untervektorraum gilt dann auch $B - A = (-1)(A - B) \in g_{A,B}$.

(3) $\Rightarrow$ (2): Nach Voraussetzung $\exists \lambda \in \mathbb{R}: \ B - A = A + \lambda(B - A) \iff (\lambda - 2)A + (1 - \lambda)B = O$. Da nicht gleichzeitig $\lambda - 2 = 0$ und $1 - \lambda = 0$ sein kann, gilt (2).

(2) $\Rightarrow$ (1): Für $A = 0$ oder $B = 0$ wurde die Behauptung bereits in Satz 12.1 gezeigt. Wir untersuchen den Fall $A \neq 0$ und $B \neq 0$: Die vorausgesetzte lineare Abhängigkeit zusammen mit $A \neq B$ bedeutet $B = \alpha A$ mit $\alpha \in \mathbb{R} \backslash \{0, 1\}$. Für $\lambda := \frac{1}{1-\alpha}$ folgt $A + \lambda(B - A) = A + \frac{1}{1-\alpha}(\alpha A - A) = \left(1 + \frac{\alpha - 1}{1 - \alpha}\right) A = O \in g_{A,B}$.

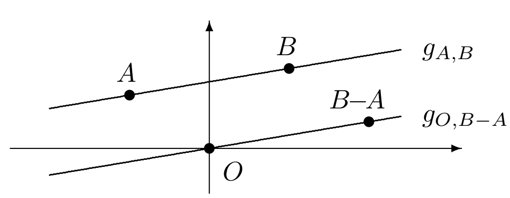
(1) $\Rightarrow$ (4): Dies wurde bereits in Kapitel 4.1 bewiesen.

Satz 12.3 Für alle Geraden $g_{A,B}$ und für alle Punkte C gilt

$$C \in g_{A,B} \iff B - A, \ C - A \text{ linear abhängig.}$$

Der **Beweis** ist entweder eine einfache Übungsaufgabe oder wird in der Vorlesung erledigt.

Wegen $g_{A,B} = \{A + \lambda(B - A) \mid \lambda \in \mathbb{R}\} = \{A + (O + \lambda((B - A) - O)) \mid \lambda \in \mathbb{R}\} =: A + g_{O,B-A}$ sagt man, dass die Gerade $g_{A,B}$ eine *Nebenklasse* des eindimensionalen Untervektorraumes $g_{O,B-A}$ ist.



Wir wollen die Geraden $g_{A,B}$ der Anschauungsebene aus einem anderen Blickwinkel betrachten, hierzu setzen wir $A = (a_1, a_2), B = (b_1, b_2)$. Es sind zwei Fälle möglich:

1. Fall: $a_1 = b_1$: Es ist $g_{A,B} = \{(a_1, a_2 + \lambda(b_2 - a_2)) \mid \lambda \in \mathbb{R}\}$ mit $a_2 \neq b_2$. In der zweiten Komponente steckt die bijektive Abbildung $\lambda \mapsto a_2 + \lambda(b_2 - a_2)$, zusammen mit den Umbenennungen $k := a_1 = b_1$ und $y = a_2 + \lambda(b_2 - a_2)$ erhalten wir

$$g_{A,B} = \{(k, y) \mid y \in \mathbb{R}\} =: g_k$$

2. Fall: $a_1 \neq b_1$: Es ist $g_{A,B} := \{(a_1 + \lambda(b_1 - a_1), a_2 + \lambda(b_2 - a_2)) \mid \lambda \in \mathbb{R}\}$. Die erste Komponente durchläuft alle reelle Zahlen (Bijektion!), wir setzen $x = a_1 + \lambda(b_1 - a_1) \iff \lambda = \frac{x-a_1}{b_1-a_1}$ und erhalten

$$\begin{aligned} g_{A,B} &= \left\{\left(x, a_2 + \frac{x - a_1}{b_1 - a_1}(b_2 - a_2)\right) \mid x \in \mathbb{R}\right\} \\ &= \left\{\left(x, \frac{b_2 - a_2}{b_1 - a_1}x + \frac{a_2 b_1 - a_1 b_2}{b_1 - a_1}\right) \mid x \in \mathbb{R}\right\} =: g_{m,b} \end{aligned}$$

wobei $m := \frac{b_2-a_2}{b_1-a_1}$ die *Steigung* und $b := \frac{a_2b_1-a_1b_2}{b_1-a_1}$ den Schnittpunkt $(0, b)$ der Geraden mit der y–Achse, genannt *Achsenabschnitt*, angibt.

Satz 12.4

a) Zu jeder Geraden $g_{A,B}$ gibt es genau ein $k \in \mathbb{R}$ oder genau ein $(m, b) \in \mathbb{R} \times \mathbb{R}$ mit $g_{A,B} = g_k$ oder $g_{A,B} = g_{m,b}$.

b) Zu jedem $k \in \mathbb{R}$ gibt es eine Gerade $g_{A,B} = g_k$.

c) Zu jedem $(m, b) \in \mathbb{R} \times \mathbb{R}$ gibt es eine Gerade $g_{A,B} = g_{m,b}$.

Beweis: a) wurde bereits bewiesen, b) und c) sind einfache Übungsaufgaben[9]

An Stelle von Geraden $g_{A,B}$ können wir daher ebenso von Geraden g_k und $g_{m,b}$ reden. Im Gegensatz zur Parameterdarstellung $g_{A,B}$ spricht man wegen der Abhängigkeit von den festen reellen Zahlen k bzw. m, b von einer *Koordinatenform*:

$$g_k = \{(x, y) \in \mathbb{R}^2 \mid 1 \cdot x + 0 \cdot y = k\}, \qquad g_{m,b} = \{(x, y) \in \mathbb{R}^2 \mid (-m) \cdot x + 1 \cdot y = b\}$$

Zum Glück stimmt unsere Anschauung mit der formalen Definition von Geraden überein. Auch von „parallel“ haben wir wahrscheinlich bereits eine feste Vorstellung.

Def 12.2 Geraden g und h der Anschauungsebene heißen parallel, geschrieben $g \parallel h$, $:\iff$

$g = h$ oder $g \cap h = \emptyset$.

Den folgenden Satz werden wir aus Zeitgründen nicht beweisen:

Satz 12.5 Es seien $g_k, g_l, g_{m,b}, g_{n,c}$ Geraden der Anschauungsebene. Dann gelten

a) Parallel ist eine Äquivalenzrelation auf der Geradenmenge.

b) $g_k \parallel g_l \qquad \forall k, l \in \mathbb{R}$

c) $g_{m,b} \parallel g_{n,c} \qquad \iff m = n$

d) $g_k \nparallel g_{m,b} \qquad \forall k, m, b \in \mathbb{R}$

Für mathematisch besonders interessierte Studierende stellen wir zwei Fragen im Zusammenhang mit der Koordinatenform von Geraden:

[9]Hinweis: Man zeichne g_k und $g_{m,b}$ und überlege, welche Punkte darauf liegen!

Frage: Für welche $u, v, w \in \mathbb{R}$ ist $g = \{(x, y) \in \mathbb{R}^2 \mid u \cdot x + v \cdot y = w\}$ eine Gerade?

Antwort: g Gerade $\iff$ $(u, v) \neq (0, 0)$

Zusatzfrage: Wie kann man aus dieser allgemeinen Koordinatenform den Richtungsvektor (die Steigung) erkennen?

Jetzt wenden wir uns den Punkten zu. Bei der folgenden Definition lasse man sich nicht durch die ungewohnte Schreibweise für eine Abbildung irritieren!

Def 12.3 Seien $X = (x_1, x_2), Y = (y_1, y_2) \in \mathbb{R}^2$. Dann heißt die Abbildung

$$< , >: \left\{ \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((x_1, x_2), (y_1, y_2)) & \mapsto & <X, Y> := x_1 y_1 + x_2 y_2 \end{array} \right. \qquad \textit{Skalarprodukt}$$

Beispiele: $<(1, 3), (1, 3)> = 10, \quad <(1, 3), (-3, 1)> = 0$

Wegen der Bijektion $\mathbb{R}^2 \to \mathbb{C}$, $(a, b) \mapsto a + ib$, können wir Punkte mit komplexen Zahlen identifizieren. In diesem Sinn ist $<X, X> = |X|^2$ bzw. $|X| = \sqrt{<X, X>}$.

Satz 12.6 Das Skalarprodukt ist

a) eine bilineare Abbildung, d.h., linear in jeder der beiden Komponenten.[10]

b) symmetrisch, d.h., $<X, Y> = <Y, X> \quad \forall X, Y \in \mathbb{R}^2$.

c) positiv definit, d.h., $<X, X> > 0 \quad \forall X \neq O$.

Der **Beweis** erfolgt in allen Teilen durch einfaches Nachrechnen.

Das Skalarprodukt aus Definition 12.3 ist nicht die einzige Abbildung $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, die Satz 12.6 erfüllt. In der Tat nennt man *jede* Abbildung mit diesen Eigenschaften Skalarprodukt oder auch *positiv definite symmetrische Bilinearform*. Weil „unser“ Skalarprodukt besonders einfach ist, dürfen wir es *kanonisches Skalarprodukt* nennen. Jeder reelle Vektorraum zusammen mit einem Skalarprodukt heißt *euklidischer Vektorraum*.

Wie wir in den Beispielen gesehen haben, kann für Vektoren $X, Y \neq O$ das zugehörige Skalarprodukt trotzdem den Wert 0 annehmen.

Def 12.4 Vektoren $X, Y \in \mathbb{R}^2$ heißen *orthogonal*, geschrieben $X \perp Y$, $:\iff <X, Y> = 0$.

Statt von orthogonalen Vektoren spricht man auch von Vektoren, die aufeinander senkrecht stehen. Stimmt dies mit unserer Anschauung überein?

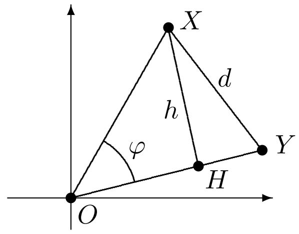
Satz 12.7 Seien $X, Y \in \mathbb{R}^2 \setminus \{(0, 0)\}$, sei φ der von X, Y eingeschlossene Winkel[11]. Dann gilt

$$<X, Y> = |X| \cdot |Y| \cdot \cos \varphi$$

Beweis: Es ist $<X, Y> = <|X| \cdot \frac{X}{|X|}, |Y| \cdot \frac{Y}{|Y|}> = |X| \cdot |Y| \cdot <\frac{X}{|X|}, \frac{Y}{|Y|}>$, daher behandeln wir nur den Spezialfall $|X| = |Y| = 1$ und beweisen hierfür $<X, Y> = \cos \varphi$.

[10] Ähnliches ist uns bereits bei der Determinantenfunktion begegnet!

[11] Wegen $\cos(2\pi - \varphi) = \cos \varphi$ ist es egal, ob man den Winkel von X nach Y oder von Y nach X misst.



Sei H der Höhenfußpunkt von X auf die Gerade $g_{O,Y}$. Ferner sei $c = |H|,\ h = |X - H|,\ d = |X - Y|$. Aus $d^2 = |X - Y|^2 = \,<X - Y, X - Y>\, = \,<X, X>\, -2 <X, Y> + <Y, Y>$ folgt $d^2 = 2 - 2 <X, Y>$. Wegen des Satzes von Pythagoras gilt $c^2 + h^2 = 1$ und $(1 - c)^2 + h^2 = d^2$, durch Subtraktion folgt $d^2 - 1 = 1 - 2c \iff d^2 = 2 - 2c$.

Insgesamt erhalten wir $2 - 2c = 2 - 2 <X, Y> \implies <X, Y> = c = \cos\varphi$.

Korollar $|<X, Y>| \le |X| \cdot |Y| \qquad \forall X, Y \in \mathbb{R}^2$.

Man hätte das Korollar, in der Fachliteratur als *Cauchy–Schwarzsche Ungleichung* bekannt, auch direkt mit Hilfe des Skalarproduktes beweisen und dann Satz 12.7 für eine elegante Winkeleinführung benutzen können. Wir wollen an dieser Stelle aber nicht näher auf Winkel eingehen und verweisen stattdessen auf die Geometrievorlesung.

In einem früheren Abschnitt haben wir uns mit linearen Abbildungen des $\mathbb{R}^2$ beschäftigt, dort ging es um Drehungen, Streckungen, Spiegelungen und Scherungen.

Def 12.5 Eine lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ heißt *orthogonal* $:\iff$

$$<X, Y> \; = \; <f(X), f(Y)> \quad \forall X, Y \in \mathbb{R}^2.$$

Man kann leicht nachweisen, dass orthogonale Abbildungen Längen und Winkel erhalten. Ohne Beweis halten wir fest, dass die Matrix einer orthogonalen Abbildung entweder von der Gestalt $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ oder $\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$ ist. Im ersten Fall liegt eine Drehung mit Determinante 1 vor, im anderen Fall handelt es sich um eine Geradenspiegelung mit Determinante -1.

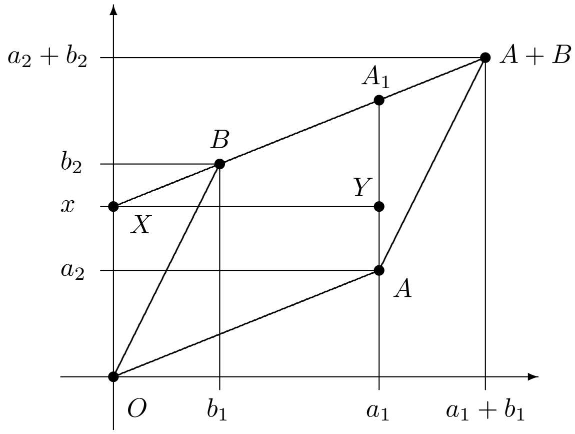
Fragen: 1) Zur Drehung: Welcher Punkt ist das Drehzentrum und um welchen Winkel wird gedreht?

2) Zur Geradenspiegelung: Wie hängt α aus der Matrix mit der Spiegelachse zusammen?

Zum Abschluss folgt eine Anwendung der Determinantenfunktion.

Satz 12.8 Seien $A = (a_1, a_2), B = (b_1, b_2) \in \mathbb{R}^2$. Dann ist $\left|\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}\right|$ der (mit Vorzeichen versehene) Flächeninhalt des von (a_1, a_2) und (b_1, b_2) aufgespannten Parallelogramms.

Je nach Größe der Zeitnot wird in der Vorlesung eine mehr oder weniger ausführlicher Beweis dieses Satzes gegeben, der nur Schulwissen wie Flächenvergleiche und eine Anwendung des Strahlensatzes benötigt.



13 Spezielle Eigenschaften des Anschauungsraumes $\mathbb{R}^3$

Vieles von dem, was wir im $\mathbb{R}^2$ behandelt haben, lässt sich problemlos auf höhere Dimensionen übertragen.

Def. 13.1 Seien $A, B \in \mathbb{R}^3$ verschiedene Punkte. Dann heißt

$g_{A,B} := \{X = A + \lambda(B - A) \in \mathbb{R}^3 \mid \lambda \in \mathbb{R}\}$ *Gerade* des $\mathbb{R}^3$.

Erneut spricht man von Aufpunkt und Richtungsvektor. Die Sätze 12.1 bis 12.3 gelten analog im $\mathbb{R}^3$. Satz 12.4 kann man nicht übertragen, da im $\mathbb{R}^3$ weitere Fälle möglich sind.

Punkte in der Ebene oder im Raum heißen *kollinear*, wenn sie gemeinsam auf einer Geraden liegen.

Def 13.2 Seien $A, B, C \in \mathbb{R}^3$ nicht kollinear. Dann heißt

$E_{A,B,C} := \{X = A + \alpha(B - A) + \beta(C - A) \in \mathbb{R}^3 \mid \alpha, \beta \in \mathbb{R}\}$ *Ebene* des $\mathbb{R}^3$.

Neben einem Aufpunkt besitzt eine Ebene zwei wegen Satz 12.3 linear unabhängige Richtungsvektoren. (in Def 13.2 sind dies $B - A$ und $C - A$). Die Sätze 12.1 und 12.2 für Geraden im $\mathbb{R}^2$ lassen sich auf Ebenen im $\mathbb{R}^3$ mit ähnlichen Beweisen übertragen:

Satz 13.1 Für alle Ebenen gilt $A, B, C \in E_{A,B,C}$ und $E_{A,B,C} = E_{A,C,B} = E_{B,A,C} = E_{B,C,A} = E_{C,A,B} = E_{C,B,A}$

Satz 13.2 Für alle Ebenen $E_{A,B,C}$ sind folgende Aussagen äquivalent:

(1) $O \in E_{A,B,C}$

(2) A, B, C linear abhängig

(3) $B - A, C - A \in E_{A,B,C}$

(4) $E_{A,B,C} \leq \mathbb{R}^3$ ($\leq$ bedeutet Untervektorraum)

Ebenen sind Nebenklassen der zweidimensionalen Untervektorräume, es gilt $E_{A,B,C} := A + E_{O,B-A,C-A}$.

In der Anschauungsebene sind Geraden entweder parallel oder sie haben genau einen Schnittpunkt. Bei Geraden im Anschauungsraum ist die Lage etwas komplizierter. Zwei Geraden ohne gemeinsamen Punkt heißen nur dann parallel, wenn es eine Ebene gibt, in der beide Geraden liegen, andernfalls spricht man von *windschief*. Zwei verschiedene Ebenen im Anschauungsraum heißen parallel, wenn sie keinen gemeinsamen Punkt besitzen.

Beispiel: Die Geraden g_{O,E_1}, g_{E_2,E_3} sind windschief. (Es ist $E_1 = (1,0,0)$, usw.)

Frage: Was weiß man über den Schnitt von nichtparallelen Ebenen E, F?

Wegen $E \cap F \neq \emptyset$ können wir oBdA $O \in E \cap F$ annehmen, d.h., E und F als zweidimensionale Untervektorräume des $\mathbb{R}^3$ voraussetzen. Aus der Dimensionsformel (Satz 3.6) erhalten wir

$$\dim(E \cap F) = \dim E + \dim F - \dim(E+F) \;=\; 2+2-3 = 1$$

Antwort: Zwei nichtparallele Ebenen schneiden sich stets in einer Geraden.

Das Skalarprodukt mit seinen Eigenschaften kann ebenfalls ohne Mühe vom $\mathbb{R}^2$ auf den $\mathbb{R}^3$ übertragen werden. Zur Erinnerung schreiben wir die Definition noch einmal auf:

Def 13.3 Seien $X = (x_1, x_2, x_3), Y = (y_1, y_2, y_3) \in \mathbb{R}^3$. Dann heißt die Abbildung

$$<\,,\,>: \begin{cases} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R} \\ ((x_1,x_2,x_3),(y_1,y_2,y_3)) & \mapsto & <X,Y> := x_1y_1 + x_2y_2 + x_3y_3 \end{cases} \qquad \textit{Skalarprodukt}$$

Selbstverständlich gelten Satz 12.6 (Eigenschaften des Skalarprodukts), Definition 12.4 (Orthogonalität), Satz 12.7, die Cauchy–Schwarzsche Ungleichung und Definition 12.5 (orthogonale lineare Abbildungen) analog. Orthogonale lineare Abbildungen im Raum sind Drehungen um eine Gerade, Spiegelungen an einer Ebene oder Kombinationen daraus.

Im ebenen Fall kann man die Fläche eines Parallelogramms mit Hilfe der Determinante bestimmen. Im Raum gilt

Satz 13.3 Seien $a = (a_1, a_2, a_3),\ b = (b_1, b_2, b_3),\ c = (c_1, c_2, c_3) \in \mathbb{R}^3$. Dann ist $\left| \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \right|$ das (mit Vorzeichen versehene) Volumen des von den drei Vektoren a, b, c aufgespannten *Parallelotops* (auch *Spat* genannt).

Beispiel: Im Spezialfall $(a_1, a_2, a_3) = (l, 0, 0), (b_1, b_2, b_3) = (0, b, 0), (c_1, c_2, c_3) = (0, 0, h)$ erhalten wir die bekannte Formel Länge · Breite · Höhe für das Volumen eines Quaders.

Fragen: 1) Wann hat ein Spat das Volumen $V = 0$?

2) Welches Volumen hat ein Tetraeder mit den Eckpunkten A, B, C, D?

3) Welches 4–dimensionale Volumen hat das 4–dimensionale Parallelotop mit den Seiten $(1,0,2,0)$, $(1,1,0,1)$, $(0,1,1,2)$, $(1,0,0,0)$?

Außer der skalaren Multiplikation ($\mathbb{R} \times V \to V$) und dem Skalarprodukt ($V \times V \to \mathbb{R}$) gibt es im Anschauungsraum eine weitere interessante Abbildung.

Def 13.4 Seien $X = (x_1, x_2, x_3),\ Y = (y_1, y_2, y_3) \in \mathbb{R}^3$. Dann heißt die Abbildung

$$\times : \begin{cases} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (X, Y) & \mapsto & X \times Y := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \end{cases}$$

Vektorprodukt oder *Kreuzprodukt*.

Beispiel: $e_1 \times e_2 = e_3$ (sogenannte Rechte Hand – Regel)

Als Merkhilfe für die Rechenvorschrift des Kreuzprodukts kann die„Determinante“ $\left|\begin{pmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}\right|$ dienen:

Man tut so, als seien die Einheitsvektoren e_i normale Zahlen und berechnet nach der Regel von Sarrus $e_1(x_2y_3 - x_3y_2) - e_2(x_1y_3 - x_3y_1) + e_3(x_1y_2 - x_2y_1)$. Dann erinnert man sich an die wahre Bedeutung der Einheitsvektoren und erhält das gewünschte Resultat $X \times Y$.

Satz 13.4 Seien $X, Y \in \mathbb{R}^3$ Dann gelten

a) $X \times Y = -(Y \times X)$ (Das Vektorprodukt $\times$ ist schiefsymmetrisch)

b) $X \perp (X \times Y), \quad Y \perp (X \times Y)$

c) $|X \times Y| = |X| \cdot |Y| \cdot \sin\varphi$, wobei φ der kleinere der von X und Y gebildeten Winkel ist.

Beweis: a) und b) erledigt man durch direktes Nachrechnen.

c): Nach Definition von Betrag und Skalarprodukt ist $|X \times Y|^2 = <X \times Y, X \times Y>$.

Elementare Rechnerei ergibt $<X \times Y, X \times Y> = |X|^2 \cdot |Y|^2 - <X, Y>^2$.

Zusammen mit Satz 12.7, der auch im $\mathbb{R}^3$ gilt, folgt die Behauptung aus

$<X \times Y, X \times Y> = |X|^2 \cdot |Y|^2 \cdot (1 - \cos^2\varphi) = |X|^2 \cdot |Y|^2 \cdot \sin^2\varphi$

Nachdem wir das Vektorprodukt kennengelernt haben, können wir das Volumen des Parallelotops (Satz 13.3) auch als $< a, b \times c >$ schreiben. Man nennt diese Kombination aus Skalar– und Vektorprodukt deshalb auch *Spatprodukt* .