

Grundlagen der linearen Algebra

ERIC KUNZE

21. September 2020

Dieses Werk ist lizenziert unter einer Creative Commons „Namensnennung – Nicht-kommerziell – Weitergabe unter gleichen Bedingungen 4.0 International“ Lizenz.



Mit dieser Zusammenfassung ist keine Garantie auf Vollständigkeit und/oder Korrektheit verbunden!

Inhaltsverzeichnis

1	Vektorräume	4
1.1	Definition	4
1.2	Lineare Unabhängigkeit	11
1.3	Basis und Dimension	13
1.4	Kern und Rang	15
1.5	Lösungen von linearen Gleichungssystemen	22

Kapitel 1

Vektorräume

1.1 Definition

Definition 1.1.1 (Vektorraum). Unter einem (K -)Vektorraum V verstehen wir das Tupel $(V, +, \cdot)$ bestehend aus einer Menge V und den Operationen "Addition" und "Skalarmultiplikation", die definiert sind durch

$$+: V \times V \rightarrow V \text{ mit } (x, y) \mapsto x + y$$

$$\cdot: K \times V \rightarrow V \text{ mit } (\lambda, x) \mapsto \lambda \cdot x$$

und folgende Eigenschaften erfüllen:

(V1) $(V, +)$ bildet eine abelsche Gruppe¹:

- (a) *Assoziativgesetz*: Für alle $u, v, w \in V$ gilt $u + (v + w) = (u + v) + w$
- (b) *neutrales Element*: es gibt $0_V \in V$ mit $0_V + v = v + 0_V = 0_V$ für alle $v \in V$
- (c) *inverses Element*: für alle $v \in V$ existiert ein Inverses $-v \in V$, sodass $v + (-v) = (-v) + v = 0_V$
- (d) *Kommutativgesetz*: für alle $u, v \in V$ gilt $u + v = v + u$

(V2) Addition und Skalarmultiplikation sind miteinander verträglich:

- (a) *Distributivgesetz 1*: Für alle $u, v \in V$ und $\lambda \in K$ gilt $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$
- (b) *Distributivgesetz 2*: Für alle $v \in V$ und $\lambda, \mu \in K$ gilt $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$
- (c) *Assoziativgesetz*: für alle $v \in V$ und $\lambda, \mu \in K$ gilt $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$
- (d) *neutrales Element*: für alle $v \in V$ gilt $1_K \cdot v = v$

¹Gruppen und auch abelsche Gruppen werden in der Vorlesung "Diskrete Strukturen" noch eingeführt werden.

Man beachte, dass wir nun sowohl im Körper K als auch im Vektorraum V eine Addition definiert haben, die wir der Einfachheit halber mit dem selben Symbol $+$ notieren. Ebenso benutzen wir das Symbol \cdot sowohl für die Multiplikation im Körper K als auch für die Skalarmultiplikation. Welche dieser Abbildungen jeweils gemeint ist, wird immer aus dem Kontext hervorgehen, und es ist wichtig, dass man sich der Unterscheidung immer bewusst ist! Zur Unterscheidung nennt man die Elemente von V oft auch *Vektoren* und die Elemente von K *Skalare*.

Insbesondere erlaubt die Definition in einem Vektorraum noch keine Multiplikation zweier Vektoren, wie man sie zum Beispiel als Skalar- und Kreuzprodukt aus der Schule kennt. Wir können einen Vektor $v \in V$ nur mit einem Skalar multiplizieren. Anschaulich entspricht dies der Streckung bzw. Stauchung eines Vektors im euklidischen Raum \mathbb{R}^3 .

Beispiel 1.1.2. ▪ Ein sehr anschauliches Beispiel für einen Vektorraum ist der aus der Schule bekannte euklidische Raum \mathbb{R}^3 über dem Körper \mathbb{R} . Die Vektoren $v \in \mathbb{R}^3$ haben dann die Gestalt $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. Die Einträge v_1, v_2, v_3 stammen stets aus dem zugrundeliegenden Körper, d.h. hier $v_1, v_2, v_3 \in \mathbb{R}$.

- In Verallgemeinerung des ersten Punktes ist \mathbb{R}^n über dem Körper \mathbb{R} ein Vektorraum für alle $n \in \mathbb{N}$. Die Vektoren $v \in \mathbb{R}^n$ haben dann genau n Einträge, d.h. sie sind von der Form² $v = (v_1, v_2, \dots, v_n)^\top$.
- Es gibt aber auch abstraktere Vektorräume, wie beispielsweise den Raum der Abbildungen von \mathbb{R} nach \mathbb{R} , den wir mit $\text{Abb}(\mathbb{R}, \mathbb{R}) := \{f: \mathbb{R} \rightarrow \mathbb{R}\}$ notieren. Als zugrundeliegenden Körper wählen wir wieder \mathbb{R} . Die Vektoren aus diesem Vektorraum haben jetzt aber nicht mehr die "typische" Gestalt aus den ersten beiden Punkten, sondern sind lediglich bestimmte Funktionen. So ist zum Beispiel die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $x \mapsto x^2$ ein "Vektor" in $\text{Abb}(\mathbb{R}, \mathbb{R})$.

Nun haben wir Mengen mit einer Struktur ausgestattet und wenn diese Mengen bestimmte Eigenschaften erfüllen, dann haben wir sie Vektorraum genannt. Nun kann man sich die Frage stellen, was denn mit Teilmengen passiert. Bilden diese wieder einen Vektorraum?

Definition 1.1.3 (Untervektorraum). Sei $(V, +, \cdot)$ ein K -Vektorraum. Wir nennen $U \subseteq V$ einen Untervektorraum von V , wenn folgenden Eigenschaften erfüllt sind:

²Um besonders große Vektoren platzsparend notieren zu können, schreiben wir Spaltenvektoren oft einfach als Zeilenvektoren und transponieren diese.

(UV1) $0 \in U$

(UV2) Für $x, y \in U$ ist auch $x + y \in U$ (Abgeschlossenheit unter $+$)

(UV3) Für $\lambda \in K$ und $x \in U$ ist auch $\lambda \cdot x \in U$ (Abgeschlossenheit unter \cdot)

Bemerkung 1.1.4. Die Eigenschaft (UV1) kann auch durch $U \neq \emptyset$ ersetzt werden, d.h. unser Untervektorraum muss einfach nichtleer sein. Das zeigt auch das folgende Lemma, was als kleines Beispiel für das Führen eines Beweises dienen soll.

Lemma 1.1.5. Sei $U \subseteq V$ und U erfülle die Bedingungen (UV2) und (UV3). Dann gilt

$$U \text{ ist Untervektorraum} \Leftrightarrow 0 \in U \Leftrightarrow U \neq \emptyset$$

Beweis. Die Voraussetzungen des Lemmas können wir als gegeben annehmen, d.h. $U \subseteq V$ und U erfüllt (UV2) und (UV3). Nun müssen wir die beiden Äquivalenzen zeigen. Äquivalenzen zeigt man meistens, indem man beiden Richtungen der Implikation beweist. Die erste Äquivalenz

$$U \text{ ist Untervektorraum} \Leftrightarrow U \neq \emptyset$$

folgt direkt aus der Definition: Ist U ein Untervektorraum, dann erfüllt er auch (UV1), d.h. es gilt $0 \in U$. Andersherum sei $0 \in U$, d.h. U erfüllt nun auch (UV1). Somit gelten dann die Eigenschaften (UV1) bis (UV3) und U ist ein Untervektorraum von V . Analog verfahren wir für die zweite Äquivalenz

$$0 \in U \Leftrightarrow U \neq \emptyset$$

Wir zeigen zuerst die Hinrichtung, d.h. wir nehmen an, dass $0 \in U$ gilt und müssen nun zeigen, dass $U \neq \emptyset$ ist. Dies ist aber schon offensichtlich, weil mit dem Nullvektor schon ein Element in U enthalten ist. Für die Rückrichtung ist ein wenig mehr Arbeit notwendig. Wir wissen, dass $U \neq \emptyset$ gilt. Somit muss es mindestens ein Element in U geben, bezeichnen wir dieses mit u , d.h. $u \in U$. Da U aber nach Voraussetzung bereits (UV3) erfüllt, dürfen wir mit beliebigen Skalaren multiplizieren, also insbesondere auch mit der Null des Körpers. Somit gilt dann $0 \cdot u \in U$ nach (UV3) und wegen $0 \cdot u = 0$ ist also auch $0 \in U$. Das ist die Behauptung, die wir zeigen wollten. \square

Beispiel 1.1.6 (triviale Untervektorräume). Sei V ein Vektorraum. Der kleinste Untervektorraum, den wir finden können, ist stets $U = \{0\}$, d.h. der Raum, der

nur den Nullvektor enthält. Damit ist offensichtlich (UV1) erfüllt und da wir nur ein Element haben, ist auch (UV2) und (UV3) schnell gezeigt: $0 + 0 = 0 \in U$ und $\lambda \cdot 0 = 0 \in U$. Der größte Untervektorraum ist dagegen immer der Vektorraum V selbst, weil er nach Definition (UV1) bis (UV3) erfüllt.

Beispiel 1.1.7 (Untervektorräume im \mathbb{R}^3). Wir betrachten den \mathbb{R} -Vektorraum \mathbb{R}^3 und wollen uns Untervektorräume von diesem ansehen. Untervektorräume können auf verschiedene Art und Weise angegeben werden. Wir betrachten hier die erste Möglichkeit, indem wir Bedingungen an die Einträge der Vektoren stellen. Betrachten wir

$$U := \left\{ \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \in \mathbb{R}^3 \mid u_3 = 0 \right\}$$

Offensichtlich ist $U \subseteq \mathbb{R}^3$. Außerdem ist $0 = (0, 0, 0)^\top \in U$. Nun wählen wir uns zwei Vektoren $u, v \in U$, d.h. $u_3 = 0 = v_3$. Es gilt

$$u + v = \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ 0 + 0 \end{pmatrix}$$

Nun ist für die Zugehörigkeit zu U nur die dritte Koordinate interessant, d.h. wir müssen nur schauen, wie der dritte Eintrag von $u + v$ aussieht. Da $u_3 + v_3 = 0 + 0 = 0$ ist, muss also auch $u + v \in U$ gelten, d.h. (UV2) ist gezeigt. Sei nun $\lambda \in \mathbb{R}$ ein beliebiger Skalar und $u \in U$, d.h. $u_3 = 0$. Dann gilt

$$\lambda \cdot u = \lambda \cdot \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda \cdot u_1 \\ \lambda \cdot u_2 \\ \lambda \cdot 0 \end{pmatrix}$$

Wieder ist die dritte Koordinate gleich Null³. Somit ist $\lambda \cdot u \in U$ für alle $\lambda \in \mathbb{R}$ und alle $u \in U$, d.h. (UV3) ist erfüllt. Schlussendlich sind (UV1) bis (UV3) erfüllt und U damit ein Untervektorraum des \mathbb{R}^3 .

Anschaulich gesprochen beschreibt U gerade die x - y -Ebene im dreidimensionalen Raum.

Um eine weitere Darstellung von Untervektorräumen zu bekommen, benötigen wir

³Die ersten beiden Koordinaten interessieren uns nicht, da an diese in U keine Bedingung gestellt wird.

zuerst noch ein neues Konzept: Linearkombinationen.

Definition 1.1.8 (Linearkombination). Sei V ein K -Vektorraum. Gegeben seien beliebig viele Vektoren $v_1, \dots, v_n \in V$ und eine gleiche Anzahl an Skalaren $\lambda_1, \dots, \lambda_n \in K$. Die Darstellung

$$\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n \in V \quad (1.1.1)$$

nennt man eine Linearkombination der v_1, \dots, v_n .

Bemerkung 1.1.9. Die Linearkombination in (1.1.1) lässt sich auch verkürzt schreiben als

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

Jede Linearkombination beschreibt wieder ein Element des Vektorraumes V , weil dieser abgeschlossen unter Skalarmultiplikation (d.h. jedes $\lambda_i \cdot v_i$ ist in V) und abgeschlossen unter Addition ist (d.h. auch die Summe verlässt V nicht).

Betrachten wir nun die folgende Situation: wir haben eine Menge von Vektoren v_1, v_2, \dots, v_n gegeben und wollen diese Menge so vervollständigen, dass sie die Struktur eines Vektorraums besitzt, d.h. ein Untervektorraum wird.

Fixieren wir dazu die v_1, \dots, v_n in (1.1.1) und variieren die Skalare $\lambda_1, \dots, \lambda_n \in K$ beliebig. Dann bekommen wir automatisch einen neuen Untervektorraum, den wir als Spannraum oder auch als lineare Hülle bezeichnen.

Definition 1.1.10 (lineare Hülle). Seien $v_1, v_2, \dots, v_n \in V$ gegeben und fixiert. Dann bezeichnet

$$\text{span}(\{v_1, v_2, \dots, v_n\}) := \{\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n \mid \lambda_1, \dots, \lambda_n \in K\}$$

die lineare Hülle bzw. den Spannraum von v_1, v_2, \dots, v_n .

Die lineare Hülle der Vektoren v_1, v_2, \dots, v_n ist der kleinste Untervektorraum von V , der v_1, v_2, \dots, v_n enthält. Wir haben uns gerade die Konstruktion "von unten" angesehen, d.h. aus den gegebenen Vektoren durch Linearkombination die Menge immer weiter vergrößert.

Eine andere Herangehensweise wäre die Konstruktion "von oben": wir nehmen uns zuerst alle Untervektorräume her, die die Vektoren v_1, v_2, \dots, v_n enthalten und bil-

den dann den Schnitt über all diese Vektorräume. Dann kann man zeigen, dass der Schnitt von Untervektorräumen wieder ein Untervektorraum bildet und wir erhalten die lineare Hülle als Schnitt aller Untervektorräume von V , die v_1, v_2, \dots, v_n enthalten. Mathematisch geschrieben erhalten wir also

$$\text{span}(\{v_1, v_2, \dots, v_n\}) = \bigcap \{U \subseteq V \mid U \text{ ist Untervektorraum, } v_1, v_2, \dots, v_n \in U\}$$

Beispiel 1.1.11. Wir betrachten nun den \mathbb{R} -Vektorraum \mathbb{R}^3 und die beiden Vektoren $u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ sowie $v = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$. Dann gilt

$$U := \text{span}(\{v\}) = \text{span}(v) = \left\{ \lambda \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} -\lambda \\ 2\lambda \\ \lambda \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$$

Somit haben wir eine Darstellung gefunden, die die Einträge in den zu U gehörigen Vektoren charakterisiert. Anschaulich beschreibt U nun eine Gerade, die durch den Ursprung gehen muss (weil $0 \in U$) und den Richtungsvektor $v = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$ besitzt. Betrachten wir nun

$$U' := \text{span}(\{u, v\}) = \text{span}(u, v) = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\}$$

Wir wissen wieder, dass U' ein Untervektorraum bildet, d.h. es ist $0 \in U'$. Dann können wir die Vektoren u und v als aufspannende Vektoren einer Ebene betrachten und den Nullvektor als Stützvektor nutzen.

Daran erkennt man nun schon, dass die Darstellung eines Untervektorraums als Spannraum mitunter schöner sein wird als die Charakterisierung der Vektoreinträge.

Beispiel 1.1.12 (Untervektorräume als Spannräume). Wir gehen zurück zu Untervektorräumen und wollen nun mithilfe der neuen Erkenntnisse untersuchen, ob gegebene Mengen Untervektorräume sind. Sei dazu $V = \mathbb{R}^2$ ein \mathbb{R} -Vektorraum und

$$U = \left\{ \begin{pmatrix} 2x - y \\ x + 4y \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

Nun haben wir hier also eine Menge gegeben, die durch Charakterisierung der Einträge festgelegt wird. Wir wollen diese Darstellung nun ein wenig zerlegen:

$$\begin{pmatrix} 2x - y \\ x + 4y \end{pmatrix} = \begin{pmatrix} 2x \\ x \end{pmatrix} + \begin{pmatrix} -y \\ 4y \end{pmatrix} = x \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y \cdot \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

Da die Einträge x und y Skalare aus dem Körper \mathbb{R} sind, steht nun eine Linearkombination der Vektoren $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 4 \end{pmatrix}$ da, d.h. jedes Element in U lässt sich als

Linearkombination dieser beiden Vektoren schreiben. Somit ergibt sich

$$U = \text{span} \left(\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \end{pmatrix} \right\} \right)$$

Da Spannräume immer Untervektorräume sind, ist auch U ein Untervektorraum von V .

1.2 Lineare Unabhängigkeit

Nun können wir den Spannraum $\text{span}(X)$ von einer relativ beliebigen Menge X von Vektoren betrachten. Allerdings kann diese Menge X eventuell sehr groß sein und vor allem auch Redundanzen enthalten, d.h. in X sind Elemente enthalten, die eigentlich nicht notwendig sind und ihr Fehlen den Spannraum nicht verändert. Dazu nutzen wir das Konzept der linearen Unabhängigkeit.

Definition 1.2.1 (lineare Unabhängigkeit). Sei V ein K -Vektorraum und $v_1, v_2, \dots, v_n \in V$. Die v_1, v_2, \dots, v_n heißen **linear abhängig**, wenn es $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ gibt, die nicht alle gleich Null sind⁴ und

$$\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n = 0 \quad (1.2.1)$$

erfüllen. Andernfalls heißen die v_1, v_2, \dots, v_n **linear unabhängig**.

Weil gerade der Fall der linearen Unabhängigkeit sehr wichtig sein wird, schreiben wir diesen Fall noch einmal explizit hin: Die v_1, v_2, \dots, v_n sind genau dann linear unabhängig, wenn die Darstellung (1.2.1) nur für $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ erfüllt ist.

Testen wir auf lineare (Un-)Abhängigkeit im \mathbb{R}^n , dann beginnen wir meistens mit der Darstellung (1.2.1). Die einzelnen Zeilen der Vektoren ergeben ein lineares Gleichungssystem, welches wir anschließend lösen. Wir brauchen keine exakte Lösung sondern vielmehr nur *eine* mögliche Belegung für die "Variablen" $\lambda_1, \lambda_2, \dots, \lambda_n$. Stellt sich heraus, dass mindestens eines der $\lambda_i \neq 0$ ist, dann liegt lineare Abhängigkeit vor, bekommen wir als einzige Lösung den Nullvektor, dann liegt lineare Unabhängigkeit vor.

⁴ $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ sind nicht alle gleich Null, wenn mindestens eines der λ_i ungleich Null ist.

Anschaulich kann man sich das Konzept wie folgt klar machen: Die Vektoren v_1, v_2, \dots, v_n seien linear abhängig, d.h. es gibt $\lambda_1, \lambda_2, \dots, \lambda_n$ (nicht alle gleich Null) mit $\lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n = 0$. Dann können wir eines der v_i als Linearkombination der anderen schreiben, z.B. $v_1 = \mu_2 \cdot v_2 + \mu_3 \cdot v_3 + \dots + \mu_n \cdot v_n$. Man stellt fest, dass dabei gerade $\mu_i = -\frac{\lambda_i}{\lambda_1}$ gilt.

Beispiel 1.2.2. Sei wieder $V = \mathbb{R}^2$ ein \mathbb{R} -Vektorraum.

- Ein einzelner Vektor $v \neq 0$ ist stets linear unabhängig, denn aus $\lambda \cdot v = 0$ muss wegen $v \neq 0$ schon $\lambda = 0$ gelten. Somit sind also alle Skalare gleich Null - v ist linear unabhängig. Sei dagegen $v = 0$ der Nullvektor, dann gilt für alle $\lambda \in \mathbb{R}$ schon $\lambda \cdot v = \lambda \cdot 0 = 0$. Beispielsweise ist die Gleichung auch für $\lambda = 5$ erfüllt. Somit existiert mindestens ein Nicht-Null-Skalar, also ist der Nullvektor stets linear abhängig. Aus dem gleichen Grund ist jede Menge von Vektoren, die den Nullvektor enthält, automatisch linear abhängig (man kann den Koeffizienten vor dem Nullvektor stets beliebig wählen).
- Betrachten wir die beiden Vektoren $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und $v = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$. Wir wollen wissen, ob u und v linear unabhängig sind. Also starten wir mit (1.2.1), d.h.

$$0 = \lambda_1 \cdot u + \lambda_2 \cdot v = \lambda_1 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot \lambda_1 + 5 \cdot \lambda_2 \\ 2 \cdot \lambda_1 + 6 \cdot \lambda_2 \end{pmatrix}$$

Nun müssen alle Einträge des Vektors auf der rechten Seite gleich Null sein, damit die Gleichung erfüllt ist. Somit können wir als Gleichungssystem formulieren

$$\begin{aligned} 0 &= 1 \cdot \lambda_1 + 5 \cdot \lambda_2 \\ 0 &= 2 \cdot \lambda_1 + 6 \cdot \lambda_2 \end{aligned} \quad \Longleftrightarrow \quad 0 = \begin{pmatrix} 1 & 5 \\ 2 & 6 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Addieren wir nun das -2 -fache der ersten Zeile zur zweiten Zeile, dann ergibt sich

$$\begin{aligned} 0 &= 1 \cdot \lambda_1 + 5 \cdot \lambda_2 \\ 0 &= -4 \cdot \lambda_2 \end{aligned} \quad \Longleftrightarrow \quad 0 = \begin{pmatrix} 1 & 5 \\ 0 & -4 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Aus der zweiten Zeile folgt nun $\lambda_2 = 0$ und nach Einsetzen in die erste Zeile liefert schließlich $\lambda_1 = 0$. Damit sind alle Koeffizienten gleich Null und u und v sind linear unabhängig. Anschaulich gesprochen heißt das, dass u nicht durch Strecken oder Stauchen aus v gewonnen werden kann.

1.3 Basis und Dimension

Wir haben uns bereits einige Linearkombinationen angesehen. Man kann nun versuchen, die Vektoren v_1, v_2, \dots, v_n in einem Vektorraum V so zu wählen, dass die Linearkombinationen eindeutig werden. Das führt uns zum Begriff der Basis.

Vorher wollen wir noch einen Begriff klären: Für einen K -Vektorraum V ist v_1, v_2, \dots, v_n ein **Erzeugendensystem**, wenn $V = \text{span}(\{v_1, v_2, \dots, v_n\})$ gilt, d.h. wenn jedes Element $v \in V$ geschrieben werden kann als $v = \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_n \cdot v_n$ für Skalare $\lambda_1, \dots, \lambda_n \in K$. Somit kann also jedes Element in V als Linearkombination der v_1, \dots, v_n geschrieben werden, man sagt auch v_1, \dots, v_n *erzeugen* V .

Definition 1.3.1 (Basis). Sei V ein K -Vektorraum. Wir nennen $\mathcal{B} = \{v_1, v_2, \dots, v_n\} \subseteq V$ eine **Basis** von V , wenn gilt:

- (1) \mathcal{B} ist ein Erzeugendensystem von V .
- (2) \mathcal{B} ist linear unabhängig.

Beispiel 1.3.2. Wir betrachten den Standardraum \mathbb{R}^3 über dem Körper \mathbb{R} . Dann ist eine einfache Basis die sogenannte Standardbasis $\mathcal{E} = \{e_1, e_2, e_3\}$ wobei $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Jedes $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in V$ lässt sich schreiben als $v = v_1 \cdot e_1 + v_2 \cdot e_2 + v_3 \cdot e_3$. \mathcal{E} ist also ein Erzeugendensystem von \mathbb{R}^3 . Außerdem sind e_1, e_2, e_3 linear unabhängig (jede Zeile ergibt, dass ein Koeffizient Null sein muss). Somit ist \mathcal{E} wirklich eine Basis von \mathbb{R}^3 .

Oftmals ist ein gegebener Untervektorraum schon als lineare Hülle (Spannraum) gegeben, woraus automatisch die Erzeugungseigenschaft folgt, d.h. wir müssen noch die lineare Unabhängigkeit der erzeugenden Vektoren beweisen.

Es gibt weitere nützliche Charakterisierungen von Basen:

Satz 1.3.3. Sei V ein K -Vektorraum und $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ eine Menge von Vektoren. Dann sind folgende Aussagen äquivalent:

- (1) \mathcal{B} ist eine Basis von V .
- (2) \mathcal{B} ist ein minimales Erzeugendensystem.
- (3) \mathcal{B} ist maximal linear unabhängig.

Wir verzichten an dieser Stelle auf einen Beweis und verdeutlichen uns lieber die Begriffe "minimal erzeugend" und "maximal linear unabhängig".

- Sei \mathcal{B} ein minimales Erzeugendensystem. Da \mathcal{B} ein Erzeugendensystem ist gilt $V = \text{span}(\{v_1, v_2, \dots, v_n\})$. \mathcal{B} ist minimal, wenn das Weglassen eines beliebigen v_i zum Verlust der Erzeugungseigenschaft führt, d.h. $\text{span}(\{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\}) \subsetneq V$. \mathcal{B} ist also minimal, wenn wir keine Redundanz, d.h. keine überflüssigen Elemente, mehr in unserer erzeugenden Menge haben.
- Sei \mathcal{B} nun maximal linear unabhängig. \mathcal{B} ist also insbesondere linear unabhängig und wenn wir einen beliebigen Vektor $v \in V$, der nicht schon in \mathcal{B} ist, hinzufügen, dann wird \mathcal{B} linear abhängig. Formal heißt das: \mathcal{B} ist linear unabhängig und $\mathcal{B} \cup \{v\}$ ist linear abhängig für alle $v \in V \setminus \mathcal{B}$.

Diese Charakterisierungen sind insbesondere dann nützlich, wenn wir einen Untervektorraum U als Spannraum von $B = \{v_1, \dots, v_n\}$ gegeben haben und nun eine Basis von U finden müssen. Da U als Spannraum gegeben ist, ist auch die Eigenschaft "Erzeugendensystem" klar. Jedoch ist zu erwarten, dass v_1, \dots, v_n linear abhängig sind. Wir müssen nun also Vektoren aus B entfernen (damit die lineare Abhängigkeit beseitigt wird), dürfen aber den Spannraum nicht verändern.

Durch das Streichen eines linear abhängigen Vektors ändern wir den Spannraum aber nicht, da jedes Vorkommen des linear abhängigen Vektors durch die Linearkombination ersetzt werden kann, die ohnehin schon im Spannraum enthalten sein muss. Korrekt heißt das also: Ist v_1, v_2, \dots, v_n linear abhängig und wir nehmen oBdA an, dass $\lambda_n \neq 0$. Dann existiert also eine Darstellung $v_n = \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_{n-1} \cdot v_{n-1} \in \text{span}(\{v_1, v_2, \dots, v_{n-1}\})$. Dementsprechend ist

$$\begin{aligned} \text{span}(\{v_1, v_2, \dots, v_n\}) &= \text{span}(\{v_1, v_2, \dots, v_{n-1}, \underbrace{\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_{n-1} \cdot v_{n-1}}_{\in \text{span}(\{v_1, v_2, \dots, v_{n-1}\})}\}) \\ &= \text{span}(\{v_1, v_2, \dots, v_{n-1}\}) \end{aligned}$$

Die Basis eines Vektorraumes ist nicht eindeutig bestimmt, aber wir können etwas über die Anzahl der Basisvektoren aussagen, denn diese ist eindeutig bestimmt.

Definition 1.3.4 (Dimension). Sei V ein endlich erzeugter⁵ K -Vektorraum und $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ eine Basis⁶ von V . Die Dimension von V bezüglich K ist dann definiert als die Mächtigkeit einer Basis, d.h.

$$\dim_K(V) = |\mathcal{B}| = n$$

⁵endlich erzeugt: es gibt ein endliches Erzeugendensystem

⁶Genauer sollten wir auch fordern, dass $v_i \neq v_j$ für $i \neq j$ gilt.

Bemerkung 1.3.5. *Existiert kein endliches Erzeugendensystem von V , d.h. V ist nicht endlich erzeugt, dann sagt man V hat "unendliche Dimension" und schreibt $\dim_K(V) = \infty$.*

Bemerkung 1.3.6. *Die gesamte Theorie, die wir bis hierher entwickelt haben (d.h. Linearkombinationen, lineare Unabhängigkeit, Basen und Dimension) sind stets abhängig vom Körper K , über dem wir den Vektorraum betrachten. Dies verdeutlichen wir auch im folgenden Beispiel.*

Beispiel 1.3.7. ▪ Der Standardraum \mathbb{R}^n hat Dimension $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$, denn die Standardbasis \mathcal{E} hat immer n Elemente.

- Der \mathbb{C} -Vektorraum \mathbb{C} hat Dimension $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, denn eine Basis ist $\mathcal{B} = \{1\}$. Wir können nämlich jedes $z \in \mathbb{C}$ darstellen als $z = \lambda \cdot 1$ für $\lambda \in \mathbb{C}$ und ein einzelner Vektor ist (sofern er nicht der Nullvektor ist) immer linear unabhängig.
- Betrachten wir dagegen \mathbb{C} als \mathbb{R} -Vektorraum, d.h. die Skalare dürfen nur aus \mathbb{R} stammen, dann ist $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ mit einer Basis $\mathcal{B}' = \{1, i\}$. Da wir nur Skalare aus \mathbb{R} zulassen, müssen wir den Imaginärteil mit einem eigenen Basisvektor beschreiben. Somit ist nun jedes $z \in \mathbb{C}$ darstellbar mit $z = \lambda_1 \cdot 1 + \lambda_2 \cdot i$. Dies entspricht der bekannten Darstellung einer komplexen Zahl mit $z = a + bi$.
- Der Vektorraum $\text{Abb}(\mathbb{R}, \mathbb{R})$ ist nicht endlich erzeugt, d.h. $\dim_{\mathbb{R}}(\text{Abb}(\mathbb{R}, \mathbb{R})) = \infty$.

1.4 Kern und Rang

Gegeben Sei nun eine Matrix $A \in \mathbb{R}^{n \times m}$. Nun kann man sich fragen, was A mit den Elementen von einem (Unter-)Vektorraum macht. Dafür interessieren uns zwei Dinge

- Welche Elemente lässt A verschwinden, d.h. für welche $x \in \mathbb{R}^n$ gilt $Ax = 0$.
- Was passiert mit den anderen Elementen?

Definition 1.4.1 (Kern und Bild einer Matrix). Sei $A \in \mathbb{R}^{m \times n}$ ist der Kern von A definiert als

$$\text{Ker}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$$

und das Bild von A als

$$\text{Im}(A) = \{Ax : x \in \mathbb{R}^n\}$$

Wollen wir den Kern bestimmen, dann müssen wir also das homogene lineare Gleichungssystem $Ax = 0$ lösen.

Beispiel 1.4.2. Wir betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 & -1 & 2 \\ -1 & 0 & -2 & 3 \\ -2 & -2 & -2 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 4}$$

und wollen den Kern von A bestimmen, d.h. alle $x \in \mathbb{R}^4$ mit $Ax = 0$ finden. Dazu lösen wir das (homogene) lineare Gleichungssystem mit dem Gauß-Jordan-Verfahren:

$$\begin{pmatrix} 0 & 1 & -1 & 2 \\ -1 & 0 & -2 & 3 \\ -2 & -2 & -2 & 2 \end{pmatrix} \xrightarrow{-1 \cdot II \rightsquigarrow I \leftrightarrow II} \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \\ -2 & -2 & -2 & 2 \end{pmatrix} \xrightarrow{2 \cdot I + III} \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \\ 0 & -2 & 2 & -4 \end{pmatrix} \xrightarrow{2 \cdot II + III} \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Nun haben wir eine Nullzeile erhalten und können nicht weiter "vereinfachen". Wir haben nun effektiv nur noch zwei beschreibende Gleichungen für vier Variablen (da $x \in \mathbb{R}^4$). Somit brauchen wir zwei freie Parameter und wählen hierfür s und t ⁷. Wir setzen also $x_4 = t$ und $x_3 = s$. Auch diese Wahl ist relativ frei, man beginnt aber üblicherweise "von hinten".

⁷Der Name der Parameter ist frei wählbar, üblich sind Buchstaben zwischen "s" und "w" oder griechische Buchstaben.

Nun können wir die ersten beiden Zeilen der umgeformten Matrix wieder als Gleichungen schreiben⁸ und ersetzen x_4 durch t und x_3 durch s .

$$0 = x_1 + 2x_3 - 3x_4 = x_1 + 2s - 3t$$

$$0 = x_2 - x_3 + 2x_4 = x_2 - s + 2t$$

Die zweite Gleichung können wir nach x_2 umstellen und erhalten

$$x_2 = s - 2t$$

Nun können wir die erste Gleichung nach x_1 umstellen, was uns

$$x_1 = -2s + 3t$$

liefert. Somit ergibt sich als Lösung des Gleichungssystems und damit auch als Elemente des Kerns von A

$$\begin{aligned} \text{Ker}(A) &= \left\{ \begin{pmatrix} -2s+3t \\ s-2t \\ s \\ t \end{pmatrix} : s, t \in \mathbb{R} \right\} = \left\{ s \cdot \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix} : s, t \in \mathbb{R} \right\} \\ &= \text{span} \left(\left\{ \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix} \right\} \right) \end{aligned}$$

Wollen wir nun noch eine Basis des Kerns bestimmen, so testen wir wieder die aufspannenden Vektoren auf lineare Unabhängigkeit: sind sie linear unabhängig, dann haben wir eine Basis gefunden; sind sie linear abhängig, müssen wir einen Vektor streichen und testen den Rest auf lineare Unabhängigkeit usw. Für unser Beispiel hier kann man nun nachrechnen und feststellen, dass beide Vektoren linear unabhängig sind und somit eine Basis von $\text{Ker}(A)$ gegeben ist durch

$$\mathcal{B} = \left\{ \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

und es gilt $\dim_{\mathbb{R}}(\text{Ker}(A)) = 2$.

⁸Da wir immer noch das homogene System $Ax = 0$ lösen, gilt immer "Zeile = 0".

Nun wollen wir uns der zweiten Frage widmen: Was passiert mit allen anderen Elementen? Wie verändert die Matrix also das "Aussehen" einer Menge von Vektoren bzw. eines Untervektorraums. Dazu bezeichnen wir im Folgenden mit s_1, \dots, s_n die Spalten und mit z_1, \dots, z_m die Zeilen von $A \in \mathbb{R}^{n \times m}$.

$$A = \begin{pmatrix} | & | & & | \\ s_1 & s_2 & \dots & s_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} - & z_1 & - \\ - & z_2 & - \\ & \vdots & \\ - & z_m & - \end{pmatrix}$$

Definition 1.4.3 (Spalten- und Zeilenraum einer Matrix). Sei $A \in \mathbb{R}^{m \times n}$. Der Zeilenraum von A ist

$$\text{Row}(A) := \text{span}(\{z_1, z_2, \dots, z_m\})$$

Der Spaltenraum von A ist

$$\text{Col}(A) := \text{span}(\{s_1, s_2, \dots, s_n\})$$

Für $A \in \mathbb{R}^{m \times n}$ und $x \in \mathbb{R}^n$ gilt

$$Ax = x_1 \cdot s_1 + x_2 \cdot s_2 + \dots + x_n \cdot s_n$$

und deswegen $\text{Im}(A) = \text{span}(\{s_1, s_2, \dots, s_n\}) = \text{Col}(A)$.

Erneut können wir uns die Frage stellen, ob in den jeweiligen Spannräume die minimale Anzahl an Vektoren steht. Dies können wir nun aber schon mit den alten Methoden behandeln, wenn wir $\text{Row}(A)$ bzw. $\text{Col}(A)$ als Untervektorräume auffassen. Wenn wir dann schließlich eine Basis von $\text{Row}(A)$ bzw. $\text{Col}(A)$ gefunden haben, können wir auch die Dimension bestimmen. Dabei stellt man fest, dass

$$\dim_K(\text{Row}(A)) = \dim_K(\text{Col}(A)) = \dim_K(\text{Im}(A)) \quad (1.4.1)$$

gilt. Man beachte aber, dass daraus im Allgemeinen nicht $\text{Col}(A) = \text{Row}(A)$ folgt.

Motiviert durch (1.4.1) definiert man nun den Begriff des Rangs einer Matrix.

Definition 1.4.4 (Rang einer Matrix). Sei $A \in \mathbb{R}^{m \times n}$. Der Rang von A ist definiert als

$$\text{rg}(A) := \dim_{\mathbb{R}}(\text{Im}(A))$$

Beispiel 1.4.5. Man bestimme das Bild und den Rang der Matrix

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

Wir haben gesehen, dass $\text{Im}(A) = \text{Col}(A) = \text{span}(s_1, s_2, s_3)$ gilt. Daher ist

$$\text{Im}(A) = \text{span} \left(\left\{ \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \right\} \right)$$

Nun wollen wir natürlich wieder eine Basis \mathcal{B} von $\text{Im}(A)$ bestimmen, d.h. wir prüfen auf lineare Unabhängigkeit.

$$0 = \lambda_1 \cdot \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

Somit müssen wir auch hier das homogene System $A \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = 0$ lösen. Also nutzen wir erneut das Gauß-Jordan-Verfahren:

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad (1.4.2)$$

Nun benötigen wir aber keine allgemeine Lösung (d.h. nicht alle Lösungen des LGS), sondern nur eine (das reicht ja schon für lineare Abhängigkeit). Wir setzen also der Einfachheit halber $\lambda_3 = 1$. Dann ergibt sich aus der zweiten Zeile

$$\lambda_2 + 2\lambda_3 = 0 \Leftrightarrow \lambda_2 = -2 \cdot \lambda_3 = -2 \cdot 1 = -2$$

Aus der ersten Zeile folgt nun

$$\lambda_1 + \lambda_2 + \lambda_3 = 0 \Leftrightarrow \lambda_1 = -\lambda_2 - \lambda_3 = -(-2) - 1 = 1$$

Testen wir dies, so ergibt sich mit

$$0 = 1 \cdot \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

eine wahre Aussage. Somit sind die Spalten also linear abhängig und wir können eine Spalte streichen, z.B. streichen wir s_3 . Dann ergibt sich

$$\text{Im}(A) = \text{span}(\{s_1, s_2, s_3\}) = \text{span}(\{s_1, s_2\})$$

Testen wir diese beiden Vektoren wieder auf lineare Unabhängigkeit, so ergibt sich analog zu (1.4.2) die umgeformte Matrix des Gleichungssystems zu

$$\begin{pmatrix} 3 & 2 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Aus der zweiten Zeile folgt nun aber schon $\lambda_2 = 0$ und mit der ersten Zeile schließlich auch schon $\lambda_1 = 0$. Somit sind s_1 und s_2 linear unabhängig und $\mathcal{B} = \{s_1, s_2\}$ eine Basis von $\text{Im}(A) = \text{Col}(A)$. Es gilt

$$\text{rg}(A) = \dim_{\mathbb{R}}(\text{Im}(A)) = \dim_{\mathbb{R}}(\text{Col}(A)) = 2$$

Bemerkung 1.4.6 (Zusammenfassung). *Den Kern $\text{Ker}(A)$ von A erhalten wir durch Lösen des homogenen Systems $Ax = 0$. Das Bild $\text{Im}(A)$ erhalten wir durch den Spaltenraum $\text{Col}(A)$. Eine Basis derer wird wie üblich durch den Test auf lineare Unabhängigkeit und eventuelles Streichen von Spalten berechnet.*

Wir haben im vorangegangenen Beispiel also die linear abhängige *Spalte* aus der Matrix gestrichen. Dies passiert nicht ohne Grund: Wir bestimmen den Rang als Dimension des Bildes bzw. des Spaltenraumes oder Zeilenraumes. Den Spaltenraum erhalten wir als Spannraum der Spalten. Finden wir dort lineare Abhängigkeiten zwischen den Spalten, dann streichen wir diese aus der potenziellen Basismenge. Diesen Streichvorgang kann man auch direkt anhand der Matrix vollziehen. Somit kann der Rang der Matrix auch beschrieben werden als die Anzahl linear unabhängiger Spalten. Alternativ kann man dies auch für Zeilen statt Spalten machen und erhält dann den Rang als Anzahl linear unabhängiger Zeilen.

Man kann damit auch feststellen, dass für $A \in \mathbb{R}^{m \times n}$ stets gilt, dass $\text{rg}(A) \leq \min\{m, n\}$. Also der Rang ist immer maximal so groß, wie die kleinste "Ausdehnung" der Matrix.

Während wir die Gleichungssysteme im Beispiel oben gelöst haben, ist uns aufgefallen, dass oft eine gesamte Zeile nach den Umformungen des Gauß-Jordan-Verfahrens nur Nullen enthält. Dies kann man ausnutzen um den Rang einer Matrix einfach anzugeben:

Dazu wendet man das Gauß-Jordan-Verfahren auf die Matrix an, d.h. man führt elementare Zeilenumformungen durch. Diese lassen den Rang der Matrix unverändert.

Abschließend zählt man die Nicht-Null-Zeilen, d.h. die Zeilen, die mindestens eine von Null verschiedene Zahl enthalten. Diese Anzahl ist dann der Rang der Matrix.

Lemma 1.4.7. Sei $A \in \mathbb{R}^{m \times n}$ in Zeilenstufenform. Dann ist der Rang von A gleich der Anzahl an Nicht-Null-Zeilen.

Beispiel 1.4.8. Wir betrachten erneut die Matrix

$$A = \begin{pmatrix} 0 & 1 & -1 & 2 \\ -1 & 0 & -2 & 3 \\ -2 & -2 & -2 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 4}$$

aus Beispiel 1.4.2. Dort hatten wir die Umwandlung in Zeilenstufenform

$$\begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \\ -2 & -2 & -2 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \\ 0 & -2 & 2 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} =: \bar{A}$$

In der Zeilenstufenform haben wir nur zwei Nicht-Null-Zeilen, d.h. es gilt $\text{rg}(\bar{A}) = 2$. Nun haben die elementaren Zeilenumformungen aber den Rang nicht verändert, d.h. es gilt auch

$$\text{rg}(A) = \text{rg}(\bar{A}) = 2$$

Wir haben eingangs schon zwei Fragen gestellt, die zwei gegensätzliche Szenarien beschrieben: Welche Elemente verschwinden und was passiert mit dem Rest? Dies spiegelt sich auch in der Struktur von Kern und Bild bzw. ihrer Dimensionen wieder.

Satz 1.4.9 (Dimensionsformel). Sei K ein Körper und $A \in K^{m \times n}$. Dann gilt

$$\text{rg}(A) + \dim_K(\text{Ker}(A)) = n$$

In der Situation von Beispiel 1.4.2 haben wir für

$$A = \begin{pmatrix} 0 & 1 & -1 & 2 \\ -1 & 0 & -2 & 3 \\ -2 & -2 & -2 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 4}$$

herausgefunden, dass $\dim_{\mathbb{R}}(\text{Ker}(A)) = 2$ und $\text{rg}(A) = 2$ nach Beispiel 1.4.8. Somit

gitl

$$\dim_{\mathbb{R}}(\text{Ker}(A)) + \text{rg}(A) = 2 + 2 = 4$$

was der horizontalen Ausdehnung der Matrix A entspricht.

Diese Formel hilft insbesondere auch zur Überprüfung der eigenen Rechnung.

1.5 Lösungen von linearen Gleichungssystemen

Wir wollen nun die Methoden aus dem letzten Abschnitt nutzen, um zu entscheiden, ob ein lineares Gleichungssystem lösbar ist und wenn ja, wieviele Lösungen existieren.

Im Folgenden sei K ein Körper. Man kann sich hier zuerst auch stets den Körper $K = \mathbb{R}$ vorstellen. Jedoch sind die Resultate auch für andere Körper wie die endlichen Körper $GF(2)$ oder $GF(3)$ richtig, deswegen verwenden wir hier K als stellvertretende "Variable".

Für eine Matrix $A \in K^{m \times n}$ und $b \in K^m$ betrachten wir das lineare Gleichungssystem $Ax = b$. Man nennt dieses **homogen**⁹, wenn $b = 0$ der Nullvektor ist, sonst **inhomogen**. Die Lösungsmenge des Systems bezeichnen wir mit $\mathcal{L}(A, b) := \{x \in K^n : Ax = b\}$.

Betrachten wir nun zuerst nur homogene lineare Gleichungssysteme, d.h. wir betrachten $Ax = 0$ für $A \in K^{m \times n}$. Man macht sich schnell klar, dass der Nullvektor stets eine Lösung des Systems ist, d.h. $0 \in \mathcal{L}(A, 0)$. Hat das System weitere Lösungen, so ist

$$\mathcal{L}(A, 0) = \text{Ker}(A) = \{x \in K^n : Ax = 0\}$$

Da der Kern einer Matrix stets ein Untervektorraum bildet, ist nun also auch die Lösungsmenge $\mathcal{L}(A, 0)$ eines homogenen Systems ein Untervektorraum des K -Vektorraums K^n .

Betrachten wir nun inhomogene Systeme, d.h. $Ax = b$ mit $b \neq 0$. Nun kann es sinnvoll sein sich trotzdem auch das zugehörige homogene System $Ax = 0$ anzusehen. Wir nehmen an, dass homogene System habe eine Lösung x_{hom} . Außerdem nehmen wir nun an, dass wir schon eine Lösung x_{inhom} des inhomogenen Systems

⁹Ein homogenes Gleichungssystem hat also stets die Form $Ax = 0$.

kennen (bspw. durch gezieltes "raten"). Dann ist auch $x_{\text{inhom}} + x_{\text{hom}}$ eine Lösung des inhomogenen Systems. Wir überprüfen die Behauptung durch Einsetzen:

$$A(x_{\text{inhom}} + x_{\text{hom}}) = \underbrace{A \cdot x_{\text{inhom}}}_{=b} + \underbrace{A \cdot x_{\text{hom}}}_{=0} = b + 0 = b$$

Somit erhalten wir also *alle* Lösungen des inhomogenen Systems, indem wir *eine* Lösung kennen und *alle* Lösungen des homogenen Systems addieren.

Um dies kompakt notieren zu können und eine algebraische Struktur zu erhalten, betrachten wir nun einen Verwandten des Untervektorraums: den affinen Teilraum.

Definition 1.5.1 (affiner Teilraum). Sei V ein K -Vektorraum und U ein Untervektorraum sowie $x \in V$. Dann ist

$$x + U = \{x + u : u \in U\}$$

ein affiner Teilraum von V .

Bemerkung 1.5.2. *Man beachte, dass $x+U$ im Allgemeinen kein Untervektorraum mehr ist, da beispielsweise die Bedingung $0 \in x + U$ verletzt ist, wenn $x \neq 0$.*

Beispiel 1.5.3. Wir betrachten den \mathbb{R} -Vektorraum \mathbb{R}^3 . Dann ist $U = \text{span}(\{(\frac{1}{2})\})$ ein Untervektorraum. Dieser beschreibt eine Gerade durch den Ursprung mit Anstieg 2. Nun können wir diese Gerade verschieben, beispielsweise um eine Einheit in y -Richtung. Dann erhalten wir die Geradengleichung $y = 2x + 1$ und diese beschreibt einen affinen Teilraum, nämlich $(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}) + U = (\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}) + \text{span}(\{(\frac{1}{2})\})$.

Affine Teilräume erlauben also quasi eine gewisse Verschiebung von Untervektorräumen.

Dies nutzen wir nun und können *alle* Lösungen eines inhomogenen Gleichungssystems durch *eine* Lösung dessen und dem Lösungsraum des zugehörigen homogenen Systems beschreiben.

Satz 1.5.4. Für die Lösungsmenge $\mathcal{L}(A, b)$ eines inhomogenen Systems $Ax = b$ gilt

$$\mathcal{L}(A, b) = x^* + \mathcal{L}(A, 0)$$

wobei x^* *eine* Lösung des inhomogenen Systems ist.

Beweis. Den Beweis dazu haben wir oben schon angeschnitten: Sei $x_{\text{hom}} \in \mathcal{L}(A, 0)$ eine beliebige Lösung des homogenen Systems und x^* eine Lösung des inhomogenen Systems. Wegen

$$A(x^* + x_{\text{hom}}) = \underbrace{Ax^*}_{=b} + \underbrace{Ax_{\text{hom}}}_{=0} = b + 0 = b$$

ist $x^* + x_{\text{hom}} \in \mathcal{L}(A, b)$ für alle $x_{\text{hom}} \in \mathcal{L}(A, 0)$, d.h. $x^* + \mathcal{L}(A, 0)$ sind alle Lösung des inhomogenen Systems. \square

Abschließend wollen wir noch ein Lösbarkeitskriterium für lineare Gleichungssysteme formulieren, mithilfe dessen wir entscheiden können, ob überhaupt eine Lösung vorliegen kann.

Satz 1.5.5. Das lineare Gleichungssystem $Ax = b$ mit $A \in K^{m \times n}$ und $b \in K^m$ ist genau dann lösbar, wenn

$$\text{rg}(A) = \text{rg}(A \mid b)$$

gilt.