

Praktische Hinweise für den Umgang mit Linearer Algebra

Sebastian Reichelt

13. April 2005

<http://www.stud.uni-karlsruhe.de/~uxau/>

Einleitung

Hallo! Ich bin Informatikstudent im 4. Semester und Tutor für Lineare Algebra. In diesem „Skript“ werde ich (meist im Voraus) in Kurzform die Hinweise aufschreiben, die ich im Tutorium geben will. Das hat den Vorteil, dass Ihr nicht unbedingt alles mitschreiben müsst und es trotzdem noch einmal nachlesen könnt. Die Übungsaufgaben, die ich während des Tutoriums vorrechne, werde ich natürlich nicht abtippen. Wer will, kann sich am Abend vorher die neueste Version des Skripts unter der oben angegebenen URL herunterladen, um während des Tutoriums mitlesen zu können.

Selbstverständlich dürft und sollt Ihr anderen Personen von diesem Skript erzählen. Der Inhalt der Linearen Algebra hängt schließlich nicht vom gewählten Tutorium ab, und im Großen und Ganzen auch nicht von dem Dozenten, der die Vorlesung hält. Was aber wiederum nicht heißt, dass dieses Skript ein Ersatz für die Vorlesung sein kann, wie Ihr schnell merken werdet.

Zum Schluss sei noch gesagt, dass **jegliche Haftung für den Inhalt dieses Skripts ausgeschlossen ist**. Wenn Ihr Fehler findet, schreibt bitte eine Mail an SebastianR@gmx.de.

Inhaltsverzeichnis

1 Grundlagen	1
1.1 Schul- und Uni-Mathematik	1
1.2 Existenz- und Allquantoren	1
1.3 Beweisführung	3
1.3.1 Voraussetzungen und Folgerungen	4
1.3.2 Einsetzen in Definitionen	4
1.3.3 Widerspruchsbeweise, Eingrenzen	5
1.3.4 Äquivalenz	5
1.3.5 Teilmengen	6
1.3.6 Gleichheit von Mengen	6
1.3.7 Gleichheit von Funktionen	6
1.3.8 Alternativen	6
1.3.9 Befreiung gebundener Variablen	7
1.3.10 Vollständige Induktion	8
1.3.11 Symmetrie	8
1.3.12 Gegenbeispiele	8
1.3.13 Benutzen aller Voraussetzungen	9
1.3.14 Themenüberschneidungen	9
1.3.15 Modelle	9
1.3.16 Spezialfälle	10
1.3.17 Formale Schreibweise	10
2 Mengenlehre	12
2.1 Mengen	12
2.2 Abbildungen	12
2.3 Relationen	14
2.4 Klassenbildung	15
3 Algebra	16
3.1 Gruppen	16
3.1.1 Definition	16
3.1.2 Erzeugnis	18
3.1.3 Homomorphismen	18
3.1.4 Der Homomorphiesatz	19

3.2	Ringe und Körper	21
3.2.1	Ringe	21
3.2.2	Körper	21
3.2.3	Unterringe und Homomorphismen	22
3.2.4	Matrizen	22
3.2.5	Polynome	24
3.3	Lineare Gleichungssysteme	26
3.3.1	Definition	26
3.3.2	Lösungsmethoden	26
3.3.3	Weiterführende Theorien	30
4	Vektorräume	31
4.1	Einführung	31
4.1.1	Ursprung	31
4.1.2	Axiome	32
4.1.3	Erzeugnis, Linearkombinationen	33
4.1.4	Lineare (Un-)Abhängigkeit	34
4.1.5	Basen und Dimension	35
4.2	Lineare Abbildungen	36
4.2.1	Definition	36
4.2.2	Multiplikation mit Matrizen	38
4.2.3	Definition über Basen	38
4.2.4	Basiswechsel	40
4.2.5	Dimensionsformel	40
4.3	Summen und Faktorräume	41
4.3.1	Summen von Vektorräumen	41
4.3.2	Faktorräume	41
4.4	Dualräume	42
4.4.1	Definition	42
4.4.2	Duale Basis	43
4.4.3	Duale Abbildung	43
4.4.4	Alternative Definition für Abbildungsmatrizen	45
4.5	Determinanten	45
4.5.1	Definition	45
4.5.2	Formeln	46

4.5.3	Aufteilung der Matrix	47
4.5.4	Laplace-Entwicklung	47
4.6	Eigenwerte	48
4.6.1	Definition	48
4.6.2	Einfache Fälle	50
4.6.3	Allgemeine Formeln	50
4.7	Jordan-Normalform	51
4.7.1	Beschreibung	51
4.7.2	Berechnung	52

1 Grundlagen

1.1 Schul- und Uni-Mathematik

Viele Erstsemester, mich eingeschlossen, stellen sich kurz nach Vorlesungsbeginn die Frage, warum sich die Mathematik an Schule und Uni so sehr unterscheiden. Denn wenn man an der Uni gleich zu Anfang mit völlig neuen Begriffen konfrontiert wird, mit deren Hilfe die Mathematik komplett und ohne Rückgriff auf bisher Gelerntes („axiomatisch“) aufgebaut werden soll, dann heißt das ja, dass nichts, was man in der Schule gelernt hat, für die Uni wichtig ist. Und umgekehrt hat man vielleicht das Gefühl, man kann viele Problemstellungen aus seinem Fachgebiet (wenn es nicht gerade Mathematik ist) mit den Mitteln der Schulmathematik lösen; wozu braucht man also neue Grundlagen?

Zum Glück wird schnell deutlich, dass das Wissen, das in der Schule vermittelt wurde, doch extrem wichtig ist. Und natürlich braucht man die neuen Begriffe, die man erst an der Uni hört! Ich sehe das folgendermaßen: Die Aufgaben, die man in der Schule löst, würden an der Uni unter den Begriff „Rechnen“ fallen. Grundlage für das Rechnen an der Schule sind die natürlichen, rationalen und reellen Zahlen und gewisse Begriffe, die darauf aufbauen. Bei der Einführung von Begriffen wird darauf geachtet, dass so wenige davon verwendet werden wie möglich und dass man sie nicht definiert, sondern erklärt und benutzt. Z.B. werden irgendwann Funktionen eingeführt. Da der Begriff aber schon recht abstrakt und schwer zu fassen ist, wird versucht, das Wesen der Funktionen zu vermitteln, indem man Wertetabellen aufstellt und Graphen zeichnet. Das geht, weil man sich auf reelle Zahlen beschränkt hat. Dann kann man sogar Stetigkeit und Differenzierbarkeit recht einfach und ohne Grenzwertbegriff anhand des Graphen erklären, auf eine sehr konkrete (wenig abstrakte) Art und Weise.

Dieses Konzept hat jedoch auch eine Schattenseite, nämlich dass man Voraussetzungen für Sätze und Definitionen größer angeben muss als nötig. An der Uni wird deshalb untersucht, welche Voraussetzungen man wirklich braucht. So kommt man zu den erwähnten neuen Begriffen, wie dem der „Gruppe“, den Ihr wahrscheinlich mittlerweile kennt. Man wird sogar Funktionen selbst in Funktionen einsetzen, was *sehr* unintuitiv und verwirrend sein kann; besonders dann, wenn man mit der jeweiligen Schreibweise noch nicht so vertraut ist. (Wer mir das jetzt nicht glaubt, soll bitte in ein paar Wochen noch einmal diese Zeilen lesen.)

Ein persönliches Ziel meines Tutoriums ist, die Frustration abzubauen, die sich durch die vielen neuen Begriffe erfahrungsgemäß am Anfang einstellt. Denn sie darf kein Grund sein, das Handtuch zu werfen, selbst wenn man die Vorlesung nicht mehr versteht und/oder die Übungsaufgaben zu schwer sind. Wenn man dabei bleibt, wird der Stoff irgendwann einfacher, weil sich die Begriffe eingepreßt haben und es wieder stärker darum geht, Zahlenbeispiele auszurechnen, bzw. die Lösungswege zu erlernen.

1.2 Existenz- und Allquantoren

Vielen von Euch werden die Symbole „ \exists “ und „ \forall “ wahrscheinlich noch nicht vertraut sein, und an der Uni kommen sie plötzlich fast in jedem Satz vor. Man kann sie als „es gibt“ und „für alle“ lesen; das macht die Sache in diesem Fall wesentlich einfacher. Aber weshalb kommen sie in der Schule kaum vor?

In der Schule könnte eine Aufgabe lauten: „Geben Sie die Lösungen der Gleichung $x^4 = 1$ an.“ Jeder Schüler erkennt sofort, dass „ $x \in \{1, -1\}$ “ die gesuchte Antwort ist. Etwas formaler ausgedrückt

könnte man schreiben: „ $x^4 = 1 \Rightarrow x \in \{1, -1\}$ “. Aber diese Aussage ist nicht besonders schön, denn ihr Wahrheitswert hängt davon ab, was „ x “ ist. („ x “ ist in diesem Fall eine so genannte „freie Variable“.) Für $x = 1 \in \mathbb{R}$ ist sie offensichtlich wahr, auch für $x = 2 \in \mathbb{R}$. (Erinnert Euch an die Definition von „ \Rightarrow “!) Sie ist eben für alle $x \in \mathbb{R}$ wahr. Aber für $x = i \in \mathbb{C}$ z.B. nicht! Also wäre die Aussage „ $\forall x : (x^4 = 1 \Rightarrow x \in \{1, -1\})$ “ in der Tat eine falsche Aussage (außerdem sind die Potenz und die Symbole „1“ und „-“ gar nicht für beliebige x definiert), während „ $\forall x \in \mathbb{R} : (x^4 = 1 \Rightarrow x \in \{1, -1\})$ “ wahr ist. Da das Ziel der Mathematik ist, wahre Aussagen aufzustellen, sollten in Sätzen eigentlich keine freien Variablen vorkommen. (Sie müssen allerdings nicht unbedingt durch Existenz- und Allquantoren gebunden sein. Auch „ $\sum_{i=1}^2 i = 3$ “ ist eine wahre Aussage.)

Etwas eigenartig ist, dass die Quantoren nie wirklich definiert werden. Offenbar fühlen sich Einige (nicht gerade Mathematiker) dadurch ermutigt, jede Menge Unsinn darüber zu behaupten. Ich habe deshalb im Internet recherchiert und dort wenigstens für endliche Mengen eine Art induktive Definition gefunden. Dazu zunächst ein analoges Beispiel: Die induktive Definition der Summe („ \sum “) sollte bekannt sein, nämlich in etwa so:

$$\sum_{i=1}^0 f(i) := 0 \text{ und } \sum_{i=1}^n f(i) := \sum_{i=1}^{n-1} f(i) + f(n) \text{ für } n \in \mathbb{N}_{>0}$$

Zumindest für endliche Indexmengen I kann man analog auch $\sum_{i \in I} f(i)$ definieren. Ebenso sollte das Produkt („ \prod “) bekannt sein als das Analogon der Multiplikation. Das führt zu folgender Tabelle, in der man auch die Existenz- und Allquantoren einordnen kann:

Symbol	Verknüpfung	Neutrales Element
\sum	+	0
\prod	\cdot	1
\forall	\wedge (und)	wahr
\exists	\vee (oder)	falsch

Z.B. heißt „ $\forall_{i \in \{1,2\}} (i < 3)$ “ (was üblicherweise geschrieben wird als „ $\forall i \in \{1,2\} : i < 3$ “) das Gleiche wie „ $(1 < 3) \wedge (2 < 3)$ “. Folgende Regeln ergeben sich ganz automatisch:

- $\neg(\forall x \in M : p(x)) \Leftrightarrow (\exists x \in M : \neg p(x))$
- $\neg(\exists x \in M : p(x)) \Leftrightarrow (\forall x \in M : \neg p(x))$

„ \neg “ bedeutet „nicht“. Es ist klar, dass diese Regeln auch für unendliche Mengen gelten sollen. Des Weiteren gilt für eine beliebige „größere“ Menge N ($M \subset N$):

- $(\forall x \in M : p(x)) \Leftrightarrow (\forall x \in N : (x \in M \Rightarrow p(x)))$
- $(\exists x \in M : p(x)) \Leftrightarrow (\exists x \in N : (x \in M \wedge p(x)))$

Man beachte den Unterschied bei „ \forall “ und „ \exists “!

Die gedankliche Übertragung auf unendliche und „allumfassende“ Mengen liefert dann die wichtigen Regeln:

1. $\neg(\forall x : p(x)) \Leftrightarrow (\exists x : \neg p(x))$
2. $\neg(\exists x : p(x)) \Leftrightarrow (\forall x : \neg p(x))$
3. $(\forall x \in M : p(x)) \Leftrightarrow (\forall x : (x \in M \Rightarrow p(x)))$
4. $(\exists x \in M : p(x)) \Leftrightarrow (\exists x : (x \in M \wedge p(x)))$

In Textform sollten sie jedem sofort einleuchten:

1. Dass $p(x)$ nicht für alle x gilt, heißt genau, dass es ein x gibt, für das $p(x)$ nicht gilt.
2. Dass es kein x gibt, für das $p(x)$ gilt, heißt genau, dass $p(x)$ für alle x falsch ist.
3. Dass $p(x)$ für alle x aus M gilt, heißt genau, dass $p(x)$ gilt, falls x aus M stammt (und für alle anderen x wird keine Aussage gemacht).
4. Dass es ein x aus M gibt, für das $p(x)$ gilt, heißt genau, dass es ein x gibt, das in M liegt und für das $p(x)$ gilt.

Wer also etwas Anderes behauptet (wie „ $(\forall x \in M : p(x)) \Leftrightarrow (\forall x : (x \in M \wedge p(x)))$ “), liegt damit auf jeden Fall falsch.

Die meisten Sätze fangen in etwa so an: „Sei $x \in \mathbb{R}$ (beliebig). Dann gilt: ...“. Statt dessen könnte man auch sagen: „Für alle $x \in \mathbb{R}$ gilt: ...“, bzw. „ $\forall x \in \mathbb{R} : \dots$ “. Auf diese Weise kann man die öde Sprache der Mathematik etwas auflockern, ohne dabei an Exaktheit einzubüßen. Wichtig ist, dass die Formulierung eine genau definierte formale Entsprechung hat, und dass einem geübten Leser sofort klar ist, wie die formale Entsprechung lautet. (Allerdings gibt es scheinbar in der Logik einen Satz, der besagt: „Wenn eine Aussage für ein beliebiges x gilt, dann gilt sie für alle x .“ Die Entsprechung gilt hier also wahrscheinlich nicht per *Definition*.)

So weit zu meinen persönlichen Theorien. Hoffentlich messt Ihr den Existenz- und Allquantoren genug Bedeutung bei, um meine langen Ausführungen zu entschuldigen. Ich werde Euch bei Gelegenheit dafür danken.

1.3 Beweisführung

Kommen wir jetzt zum ersten wirklich praktischen Teil. Während der gesamten Studienzeit werdet Ihr wohl oder übel Beweise führen müssen. Erfahrungsgemäß sagt einem aber niemand, wie das geht, sondern die Dozenten rechnen Beweise vor in der Hoffnung, dass man ihre Technik irgendwann erlernt. Keine Panik, es gibt durchaus Verfahren, die in den meisten Fällen irgendwann zum Ziel führen. Die ersten erscheinen wahrscheinlich trivial.

1.3.1 Voraussetzungen und Folgerungen

Ein Beweis hat normalerweise die Form eines Weges. Man fängt an einem bestimmten Punkt an, versucht, ein Ziel anzusteuern, und hofft darauf, nicht zwischendurch wieder an einen bekannten Platz zurückzukommen. Einen großen Unterschied gibt es schon: Auf dem gesamten Weg gewinnt man immer mehr Erkenntnisse, d.h. man kann sich dem Ziel eigentlich gar nicht weiter entfernen. Konkret sieht das so aus:

Zunächst schreibe ich alle Voraussetzungen auf, die mir auf den Weg gegeben wurden; das ist mein Ausgangspunkt: „Sei G eine Gruppe, ...“. Nun gehe ich in einzelnen Schritten in eine Richtung, die mir sinnvoll erscheint; die Schritte trenne ich durch einen Folgepfeil ($, \Rightarrow$). In jedem Schritt darf ich die gesamten Voraussetzungen benutzen, das dürfte jedem einleuchten. Ich darf aber auch *alle* Aussagen benutzen, die ich auf dem Weg schon gemacht habe. Am besten nenne ich mal ein Beispiel:

Sei G eine Gruppe, $x \in G$, y_1 und y_2 invers zu x . Dann gilt:

$$y_1 \cdot x = 1_G \wedge x \cdot y_2 = 1_G \Rightarrow (y_1 \cdot x) \cdot y_2 = 1_G \cdot y_2 \Rightarrow y_1 \cdot (x \cdot y_2) = y_2 \underset{x \cdot y_2 = 1_G}{\Rightarrow} y_1 \cdot 1_G = y_2 \Rightarrow y_1 = y_2$$

Solange man wie hier gegebenenfalls die benutzte Tatsache unter den Folgepfeil schreibt, sollte dies anerkannt werden, denn man hat vorher hergeleitet, dass sie unter den Voraussetzungen richtig ist.

Es ist übrigens eine Konvention, „ $A \Rightarrow B \Rightarrow C$ “ für „ $(A \Rightarrow B) \wedge (B \Rightarrow C)$ “ zu schreiben, genau wie man gerne „ $a = b = c$ “ für „ $(a = b) \wedge (b = c)$ “ schreibt. Das macht in beiden Fällen wegen der Transitivität Sinn (d.h. es gilt auch $A \Rightarrow C$ bzw. $a = c$). Etwas gefährlicher ist da schon „ $A \Leftrightarrow B \Rightarrow C$ “, was für „ $(A \Leftrightarrow B) \wedge (B \Rightarrow C)$ “ steht. Denn wenn man es übertreibt, z.B. „ $A \Leftarrow B \Leftrightarrow C \Rightarrow D$ “, dann kommt Unsinn dabei heraus. In diesem Beispiel könnte man weder D aus A folgern noch umgekehrt.

1.3.2 Einsetzen in Definitionen

Der Schritt von „ y_1 invers zu x “ nach „ $y_1 \cdot x = 1_G$ “ im obigen Beweis ist eigentlich trivial, denn es wird nur die Definition des „inversen Elements“ benutzt. Trotzdem lohnt es sich mehr als man glaubt, den Schritt explizit hinzuschreiben. Denn nur in der ausgeschriebenen Formel sieht man, wie man weiter verfahren muss; von einer Aufgabenstellung wie „Sei G eine Gruppe, $x \in G$, y_1 und y_2 invers zu x . Zeigen Sie: $y_1 = y_2$.“ wird man oft zunächst überrumpelt.

Es bietet sich manchmal an, nicht direkt die Definition zu benutzen, sondern einen Satz, der sich daraus ableitet. Das drängt sich besonders auf, wenn der Satz zu einer alternativen Definition führt.

Z.B. gibt es für eine Funktion f die Definition:

$$f \text{ ist bijektiv} :\Leftrightarrow f \text{ ist injektiv} \wedge f \text{ ist surjektiv}$$

Aber es gilt die Äquivalenz:

$$f \text{ ist bijektiv} \Leftrightarrow f^{-1} \text{ existiert}$$

Das wäre also eine alternative Definition. Ist als Voraussetzung gegeben, dass f bijektiv ist, dann braucht man viel öfter die Existenz von f^{-1} als die Injektivität oder Surjektivität, und fängt den Beweis lieber so an: „ f ist bijektiv $\Rightarrow f^{-1}$ existiert $\Rightarrow \dots$ “

Besondere Aufmerksamkeit verdienen Mengen, die über Bedingungen definiert sind: $M := \{x : p(x)\}$ mit einem Prädikat p . Für eine Variable m ist dann $m \in M$ äquivalent zu $p(m)$ und $m \notin M$ äquivalent zu $\neg p(m)$.

Ist z.B. zu zeigen, dass Kern $f \subset M$ ist, dann macht man das normalerweise so: „Sei $k \in \text{Kern } f$ beliebig. $f(k) = 0 \Rightarrow \dots \Rightarrow k \in M$ “. Hier wurde die Definition von Kern f benutzt, nämlich Kern $f := \{x \in \text{Def } f : f(x) = 0\}$. Das Prädikat p ist in diesem Fall $p(x) \equiv (x \in \text{Def } f \wedge f(x) = 0)$.

Manchmal soll man die Abgeschlossenheit einer Menge M bezüglich einer Verknüpfung „ \star “ zeigen. Ist M definiert durch $M := \{x : p(x)\}$ wie oben, dann geht das im Allgemeinen so: Seien $m_1, m_2 \in M$. $p(m_1) \wedge p(m_2) \Rightarrow \dots \Rightarrow p(m_1 \star m_2) \Rightarrow m_1 \star m_2 \in M$

1.3.3 Widerspruchsbeispiele, Eingrenzen

Wieder etwas, das in der Theorie völlig klar ist: Wenn ich „ $A \Rightarrow B$ “ beweisen will, kann ich statt dessen auch „ $\neg B \Rightarrow \neg A$ “ zeigen, denn das ist gerade äquivalent. Im weiteren Sinne wird auch diese Technik „Widerspruchsbeweis“ genannt. Man sollte stets beide Varianten ausprobieren, d.h. man fängt bei A an und folgert in Richtung B , und fängt außerdem bei $\neg B$ an und folgert in Richtung $\neg A$. Wenn eine der beiden Varianten einfacher scheint, dann führt sie vielleicht auch eher zum Ziel. Es lässt sich jedoch *jeder* Beweis komplett umkehren, indem man einfach jede Aussage negiert und die Folgepfeile umdreht.

Mit etwas mehr Erfahrung kann man Probleme auf diese Weise auch eingrenzen, obwohl man noch keinen vollständigen Lösungsweg kennt. Führt man nämlich *gleichzeitig* beide Varianten durch, kann man sich quasi in der Mitte treffen. Man setzt dazu *sowohl* A *als auch* $\neg B$ voraus und folgert daraus einen Widerspruch.

Widerspruchsbeweise beginnen immer mit einer besonders ausgezeichneten Annahme. In diesem Fall würde man schreiben: „Annahme: $\neg B$. Daraus folgt: ...“ Kann man direkt das Gegenteil der Voraussetzung A folgern, dann reicht das schon aus. Greift man auf die Eingrenz-Technik zurück, dann sollte man am Schluss noch unbedingt schreiben: „... im Widerspruch zur Annahme“. Ein Beispiel findet sich unter 1.3.17.

1.3.4 Äquivalenz

Das naheliegendste Verfahren, um die Äquivalenz von Aussagen zu zeigen, sind Äquivalenzumformungen, wie man sie aus der Schule kennt. Überraschenderweise ist es in der Linearen Algebra oft viel vorteilhafter, nur Folgerungen zu benutzen. Um „ $A \Leftrightarrow B$ “ zu zeigen, zeigt man sowohl „ $A \Rightarrow B$ “ als auch „ $B \Rightarrow A$ “, und zwar getrennt.

Soll man die Äquivalenz mehrerer Aussagen zeigen, dann kann man dies mit Hilfe eines Zirkelschlusses tun. Am besten notiert man sich in der Aufgabenstellung mit Pfeilen, welche Schlüsse machbar sind. Das ist ein Graph, bei dem jeder Zyklus bedeutet, dass die jeweiligen Aussagen äquivalent sind. Erhält man einen Zyklus, der nicht alle Aussagen umfasst, kann man immer noch die Äquivalenz der übrigen Aussagen zu den Aussagen im Zyklus zeigen.

Beispiel: Man soll „ $A \Leftrightarrow B \Leftrightarrow C \Leftrightarrow D$ “ zeigen. Sind die Schlüsse „ $A \Rightarrow B$ “, „ $B \Rightarrow C$ “ und „ $C \Rightarrow A$ “ machbar, dann hat man einen Zyklus „ $A \Rightarrow B \Rightarrow C \Rightarrow A$ “ erzeugt und damit „ $A \Leftrightarrow B \Leftrightarrow C$ “ gezeigt. Jetzt reicht es, D aus irgendeiner der anderen Aussagen zu folgern, und irgendeine Aussage aus D , also z.B. „ $A \Rightarrow D$ “ und „ $D \Rightarrow B$ “.

1.3.5 Teilmengen

Soll man „ $A \subset B$ “ zeigen, dann ist es *fast immer* sinnvoll, sich ein bestimmtes Element aus A „herauszunehmen“ und zu zeigen, dass es auch in B liegt: „Sei $a \in A$ beliebig. Dann: $\dots \Rightarrow a \in B$ “ Das reicht als Beweis auch schon völlig aus; jedem Korrektor sollte klar sein, dass damit „ $A \subset B$ “ gezeigt wurde. Der Grund, warum dies einfacher ist als Mengenumformungen, ist, dass Mengen immer über ihre Elemente charakterisiert werden. In manchen Fällen ist dies offensichtlich; unter 1.3.2 gab es ein solches Beispiel. Manchmal ist es aber auch nicht so leicht zu sehen:

„Sei G eine Gruppe, H_1 und H_2 Untergruppen mit $H_1 \cup H_2 = G$. Zeigen Sie: $G \subset H_1 \vee G \subset H_2$.“

Scheinbar werden hier nur Aussagen über die Mengen selbst gemacht. Aber Moment: Wie ist denn „ $H_1 \cup H_2$ “ definiert? Richtig: $H_1 \cup H_2 = \{h : h \in H_1 \vee h \in H_2\}$. Es hilft also doch, ein spezielles g aus G in der Hand zu haben, denn dann ist schon mal $g \in H_1$ oder $g \in H_2$. Das ist natürlich keine Überraschung.

1.3.6 Gleichheit von Mengen

Wie bei der Äquivalenz von Aussagen (1.3.4) bietet es sich fast immer an, die beiden Richtungen getrennt zu behandeln. Das Beispiel von 1.3.5, etwas abgeändert:

„Sei G eine Gruppe, H_1 und H_2 Untergruppen mit $H_1 \cup H_2 = G$. Zeigen Sie: $G = H_1 \vee G = H_2$.“

„ $H_i \subset G$ “ ($i \in \{1, 2\}$) ist aber trivial; es bleibt also noch „ $G \subset H_i$ “ für ein i , wie oben.

1.3.7 Gleichheit von Funktionen

Für zwei Funktionen $f_i : D \rightarrow R$ ($i = 1, 2$) ist die Gleichheit *definiert* durch $f_1 = f_2 :\Leftrightarrow \forall x \in D : f_1(x) = f_2(x)$. Um sie zu zeigen, greift man sich also wieder ein beliebiges $x \in D$ heraus und zeigt erst einmal $f_1(x) = f_2(x)$. Da x beliebig war, ist dann $f_1 = f_2$.

Oft soll man zeigen, dass eine bestimmte Funktion die Nullfunktion ist. Das ist eine kleine Falle, denn „Zeigen Sie: $f = 0$ “ sieht sehr unscheinbar aus. Man muss zeigen, dass f alle x aus dem Definitionsbereich auf 0 abbildet, also wieder so: „Sei $x \in \text{Def } f$ beliebig. Dann: $\dots \Rightarrow f(x) = 0$ “

1.3.8 Alternativen

Enthält die Aussage, die man zeigen soll, zwei durch „oder“ getrennte Alternativen, dann muss man nur die Negation von einer der Alternativen annehmen und die andere zeigen. Denn ist diese Annahme falsch, also die entsprechende Alternative richtig, dann stimmt die gesamte Aussage sowieso schon.

Um „ $G \subset H_1 \vee G \subset H_2$ “ im Beispiel von 1.3.5 zu zeigen, kann ich annehmen, dass $G \not\subset H_1$ ist (d.h., dass es ein Element in G gibt, das nicht in H_1 liegt), und unter dieser Annahme zeigen, dass $G \subset H_2$ gilt. Denn wenn $G \subset H_1$ ist, dann bin ich schon fertig.

Natürlich gibt es auch hier die Möglichkeit des Widerspruchsbeweises, d.h. man nimmt an, dass beide Alternativen falsch sind, und folgert, dass dann die Voraussetzungen nicht erfüllt sind.

1.3.9 Befreiung gebundener Variablen

„ $\exists x \dots$ “ und „ $\forall x \dots$ “ binden jeweils die Variable x . Außerhalb des Ausdrucks hat x eigentlich keine Bedeutung, man könnte sie sogar theoretisch neu belegen. Folgendes Beispiel wäre formal korrekt:

$$\exists y \in \text{Bild } f : z = y \Rightarrow \exists y \in \text{Def } f : z = f(y)$$

Hier ist y in der ersten Aussage das, was $f(y)$ in der zweiten Aussage ist.

So etwas ist aber überhaupt nicht schön und kann sogar Punktabzug geben! Aus Gründen der Übersichtlichkeit werden Variablen nämlich oft auch außerhalb ihres Geltungsbereichs benutzt. Hat man die Existenz eines Objekts bewiesen, das einer bestimmten Bedingung genügt, dann will man dieses oft benennen, um damit weiterrechnen zu können. Der Existenzquantor hat dem Objekt aber schon vorläufig einen Namen gegeben, den man nach Möglichkeit beibehalten sollte:

$\exists y \in \text{Bild } f : z = y$. Für dieses y gilt: $\exists x \in \text{Def } f : y = f(x)$, und damit für dieses x : $z = f(x)$

Das reißt den Beweis jedoch stark auseinander, und man verfällt leicht in immer mehr natürliche Sprache. Man könnte versuchen, es mit mathematischen Symbolen zu schreiben, also so:

$$\exists y \in \text{Bild } f : z = y \Rightarrow \exists x \in \text{Def } f : y = f(x) \Rightarrow z = f(x)$$

Es ist jedoch nicht klar, wie dies zu verstehen ist. Zunächst einmal gibt es jeweils zwei verschiedene Möglichkeiten, wie man Klammern setzen könnte:

1. $\exists y \in \text{Bild } f : (z = y \Rightarrow \exists x \in \text{Def } f : y = f(x))$
2. $(\exists y \in \text{Bild } f : z = y) \Rightarrow (\exists x \in \text{Def } f : y = f(x))$

Die erste Version drückt nicht das Richtige aus. Sie wäre z.B. auch dann wahr, wenn $z = y$ für kein $y \in \text{Bild } f$ gelten würde. Die zweite Version kann aber auch nicht richtig sein, denn y ist eigentlich außerhalb der ersten Klammer gar nicht mehr definiert.

Man braucht also eine neue Schreibweise. Mein persönlicher Vorschlag ist, die Folgerung versetzt unter den Term im Existenzquantor zu schreiben:

$$\begin{aligned} \exists \quad & \underbrace{y \in \text{Bild } f}_{\substack{\Rightarrow \exists x \in \text{Def } f : \\ y = f(x) \\ \Rightarrow_{z=y} z = f(x)}} & : z = y \end{aligned}$$

Das sind die gleichen Symbole wie oben, nur anders platziert. Nun ist etwas klarer, auf welche Weise gefolgert wird. Das Ergebnis soll übrigens sein: $\exists x \in \text{Def } f : z = f(x)$

Das Ganze funktioniert übrigens so *nur* mit Existenzquantoren und nur eingeschränkt mit Allquantoren. (Warum wohl?) Glücklicherweise liefert die Technik des Widerspruchsbeweises eine Möglichkeit, Allquantoren anhand der bekannten Regeln in Existenzquantoren umzuwandeln:

„Sei $f : D \rightarrow R$ eine injektive Funktion. Zeigen Sie: $\forall M \subset D : f^{-1}(f(M)) = M$.“

Die Injektivität ist definiert als: $\forall x_1, x_2 \in D, f(x_1) = f(x_2) : x_1 = x_2$. Die Negation davon ist: $\exists x_1, x_2 \in D, x_1 \neq x_2 : f(x_1) = f(x_2)$. Und auch die Negation von $\forall M \subset D : f^{-1}(f(M)) = M$ liefert einen Existenzquantor: $\exists M \subset D : f^{-1}(f(M)) \neq M$. Fangen wir also damit an. Wegen $M \subset f^{-1}(f(M))$ ist dann $f^{-1}(f(M)) \not\subset M$, also gibt es ein $x \in f^{-1}(f(M))$ mit $x \notin M$. Nach der Definition von f^{-1} ist $f(x) \in f(M)$. Für jedes Element z aus $f(M)$ (insbesondere für $z = f(x)$)

gibt es aber auch ein $y \in M$, so dass $z = f(y)$ ist, d.h. $f(x) = f(y)$. Wegen $x \notin M$ und $y \in M$ ist $x \neq y$. Damit ist f nicht injektiv.

Es ist hoffentlich klar, dass es in Wirklichkeit keine Rolle spielt, wie gebundene Variablen heißen. Die Aussage, die hier gezeigt wurde, war: „ $\exists x, y \in D, x \neq y : f(x) = f(y)$ “. Das ist äquivalent zu „ $\exists x_1, x_2 \in D, x_1 \neq x_2 : f(x_1) = f(x_2)$ “.

1.3.10 Vollständige Induktion

Das Prinzip der vollständigen Induktion möchte ich hier nur der Vollständigkeit halber erwähnen. Es eignet sich gut, um eine Aussage zu zeigen, die von einer natürlichen Zahl n abhängt. Man zeigt sie dann z.B. für $n = 1$ und sagt: „Wenn die Aussage für $n = n_0$ gilt ($n_0 \in \mathbb{N}$), dann gilt sie auch für $n = n_0 + 1$.“ Man kann auch voraussetzen, dass die Aussage für alle $n \leq n_0$ gilt, aber das braucht man sehr selten. Generell muss man *immer* die Induktionsvoraussetzung benutzen; darauf sollte man als Erstes hinarbeiten.

Vollständige Induktion übt man im Laufe der ersten Semester sehr oft. Wer sich an den Übungen beteiligt, bekommt in den Klausuren dadurch wichtige Punkte geschenkt.

1.3.11 Symmetrie

Ich denke, in der Zwischenzeit habt Ihr die Formulierung „o.B.d.A.“ („ohne Beschränkung der Allgemeinheit“ bzw. „ohne Bedenken des Autors“) kennen gelernt. Tatsächlich gibt es einige Situation, in denen es gerechtfertigt ist, Voraussetzungen zu treffen, die nicht in der Aufgabenstellung gegeben sind.

Beispiel: In der Aufgabenstellung ist eine endliche Menge reeller Zahlen ($\{a_1, a_2, \dots, a_n\}$) gegeben, deren Reihenfolge für das Resultat keine Rolle spielt. Man hätte sie gerne in aufsteigender Reihenfolge sortiert. Warum nicht? Denn wenn sie es nicht sind, kann man sie sortieren, den Beweis bzw. die Rechnung auf den sortierten Zahlen durchführen, und das Ergebnis auch auf die unsortierten Zahlen anwenden, da die Reihenfolge ja keine Rolle spielte.

Wichtig ist, dass es in der Aufgabenstellung eine Symmetrie im weitesten Sinne gibt. Was ich mit „Symmetrie“ meine, wird vielleicht eher an folgendem Beispiel deutlich:

„Sei entweder $a > 0$ und $b < 0$, oder $a < 0$ und $b > 0$. Zeigen Sie: $a \cdot b < 0$.“

Hier wäre es völlig legitim, am Anfang „o.B.d.A. $a > 0, b < 0$ “ zu schreiben. Am besten schreibt man am Ende noch etwas wie: „Der Fall $a < 0, b > 0$ folgt aus Symmetriegründen wegen der Kommutativität der Multiplikation.“

1.3.12 Gegenbeispiele

Ein Gegenbeispiel gibt man immer dann an, wenn man zeigen will, dass eine bestimmte Aussage, die in mathematischer Schreibweise ein Allquantor-Term wäre, falsch ist. Theoretisch gesehen wandelt man dabei den Allquantor in einen Existenzquantor um, nach der entsprechenden Regel.

$\forall x \in \mathbb{R} : x > 0$ ist falsch, denn $x = 0$ ist ein Gegenbeispiel. (Wow!)

1.3.13 Benutzen aller Voraussetzungen

In Übungs- und Klausuraufgaben werden Voraussetzungen normalerweise nur dann angegeben, wenn sie auch wirklich wichtig sind, d.h. wenn man sie irgendwo braucht. Die Aufgabe wird nicht von einer „abelschen Gruppe“ sprechen, wenn die Kommutativität für das Resultat keine Rolle spielt. Dementsprechend sollte man versuchen, „blind“ die Voraussetzungen anzuwenden, wenn man sonst keine Idee hat.

Wenn man sich nicht sicher ist, welchen Einfluss die einzelnen Voraussetzungen haben, könnte es helfen, sie wegzulassen und jeweils ein Gegenbeispiel zu finden. Um bei dem Beispiel zu bleiben: Man nehme z.B. die Gruppe S_3 , weil sie nicht abelsch ist, und schau, warum die zu zeigende Aussage dann nicht mehr gilt.

Es gibt eine Ausnahme von dieser Regel: Bei Vektorräumen wird oft vorausgesetzt, dass sie endlichdimensional sind. Das hilft beim Beweisen, weil man sich zunächst eine Basis basteln kann. Aber die Aussage könnte trotzdem für alle Vektorräume gelten; sie ist dann eben nur schwerer zu zeigen. Dementsprechend sollte man nie versuchen, ein unendlichdimensionales Gegenbeispiel zu finden; das wäre sowieso viel zu schwer.

Besondere Aufmerksamkeit sollte man auf die Ergebnisse der vorherigen Aufgabenteile richten. Teilaufgaben hängen meistens zusammen, wobei es nicht immer die direkt aufeinander folgenden Teile sein müssen.

1.3.14 Themenüberschneidungen

Manchmal überschneiden sich verschiedene Teilgebiete der Mathematik. Wo dies passiert, wird es auch an den Aufgaben deutlich. Es ist dann hilfreich, sich die wichtigsten Ergebnisse der Teilgebiete noch einmal ins Gedächtnis zu rufen.

Z.B. überschneiden sich die Theorien über Ringe, Vektorräume und Polynome beim charakteristischen Polynom einer Matrix. Ist nun eine Aufgabe gegeben, in der Matrizen potenziert, mit Skalaren multipliziert und addiert werden, sollte man schauen, ob man daraus nicht ein Polynom bilden kann; auch wenn sich der Stoff gerade komplett um Vektorräume dreht.

1.3.15 Modelle

Man kann für die verschiedenen Themen der Mathematik nur dann ein intuitives Verständnis bilden, wenn man sich Modelle überlegt, die ein abstraktes Objekt in ein anschauliches Gebilde übersetzen. Das geht in der Gruppentheorie nur sehr bedingt; man kann sich höchstens überlegen, was man als Gruppe betrachten kann und was nicht, welche Untergruppen es jeweils gibt, und wo es Analogien gibt. Die Ringtheorie bietet da schon mehr Spielraum, denn die zwei Verknüpfungen eines Rings ähneln auf irgendeine Art und Weise immer den aus der Schule bekannten. Spätestens bei den Vektorräumen bieten sich jede Menge Möglichkeiten, Dinge anschaulich zu machen.

Bei jedem Modell muss man sich aber auch überlegen, wo die Grenzen des Modells liegen, d.h. welche Sachverhalte vom Modell nicht sinnvoll wiedergegeben werden.

Einen endlichdimensionalen Vektorraum kann man sich oft in seine einzelnen Dimensionen aufgeteilt denken. Das macht Sinn, wenn man allgemeine lineare Abbildungen untersucht, denn sie bilden in gewisser Weise jede Dimension entweder wieder auf eine Dimension oder auf gar nichts ab. Auch für Faktorräume ist das Modell gut geeignet; damit kann man Dimensionen „entfernen“. Was das

Modell jedoch nicht leistet, ist eine sinnvolle Aussage über Vektoren selbst. Denn die identische Abbildung (als Endomorphismus) könnte man in diesem Modell z.B. nicht von irgendeinem anderen Automorphismus unterscheiden.

1.3.16 Spezialfälle

Ist eine Aufgabenstellung von allgemeiner Natur, und die Allgemeinheit der Grund dafür, dass sie schwierig ist, dann hilft es meistens, sich bestimmte Spezialfälle zu überlegen, um so eine Grundlage für die allgemeine Lösung zu erhalten. Das ist besonders hilfreich, wenn man die Lösung durch vollständige Induktion beweisen kann. Dann kann man anhand der ersten paar Spezialfälle meist die allgemeine Formel erraten. Normalerweise ist Raten in der Mathematik nicht erwünscht, aber wenn man anschließend mittels vollständiger Induktion die Richtigkeit beweist, hat man seine Pflicht getan.

Bei allgemeinen Vektorräumen bietet sich oft an, sich die Dimensionen 1, 2 und 3 besonders anzuschauen. Wenn es um Unterräume geht, hat dies allerdings einen Nachteil: Bei Dimension n kommen Unterräume nur in $n + 1$ verschiedenen Dimensionen vor, wovon auch noch 2 völlig uninteressant sind. Im 3-dimensionalen Anschauungsraum bleiben uns nur Geraden und Ebenen als Unterräume, die wir untersuchen können.

Ist ein bestimmter Mindestwert für eine Zahl gegeben (gilt z.B. eine Aussage nur für Vektorräume der Dimension 2 oder größer), dann lohnt es sich, diesen Mindestwert als Spezialfall genau zu untersuchen, und auch festzustellen, warum die Aussage für kleinere Werte noch nicht gilt (siehe 1.3.13).

1.3.17 Formale Schreibweise

Ausnahmsweise möchte ich noch einen Hinweis geben, der nicht dem Finden eines Lösungswegs dient, sondern nur der Korrektheit und Nachvollziehbarkeit, insbesondere für den Korrektor, also z.B. mich. :-). Wenn Ihr die Lösung vor Augen habt, solltet Ihr Euch unbedingt die Mühe machen, den Beweis so formal wie möglich zu führen. Mal wieder ein Beispiel, und zwar der klassische Beweis für das Prinzip der vollständigen Induktion, zunächst in Textform:

Sei $p(n)$ eine Aussage, die für jedes $n \in \mathbb{N}$ entweder wahr oder falsch ist. $p(1)$ sei wahr, und wenn $p(n)$ für ein $n \in \mathbb{N}$ wahr ist, dann sei auch $p(n + 1)$ wahr.

1. *Sei m die kleinste natürliche Zahl, für die $p(m)$ falsch ist.*
2. *Weil $p(1)$ wahr ist, ist $m \geq 2$.*
3. *$p(m - 1)$ ist wahr, weil m ja gerade die kleinste Zahl war, für die $p(m)$ falsch ist.*
4. *Dann ist nach Voraussetzung auch $p(m)$ wahr.*
5. *Dies ist ein Widerspruch, also kann es kein solches m geben.*
6. *Also ist $p(n)$ für alle n wahr.*

Auf den ersten Blick ist der Beweis einleuchtend, und wahrscheinlich würde er so akzeptiert. Aber eigentlich enthält er einige Lücken:

1. Wer sagt denn bitte, dass es eine kleinste natürliche Zahl, für die $p(m)$ falsch ist, überhaupt gibt?! Siehe „5.“
2. Eigentlich folgt daraus erst einmal $m \neq 1$, daher $m > 1$ oder $m < 1$ (was wegen $m \in \mathbb{N}$ unmöglich ist), und aus $m > 1$ und $m \in \mathbb{N}$ folgt $m \geq 2$. (Keine Sorge, so pingelig ist kein Korrektor. Das hier ist nur ein Beispiel.)
3. Im formalen Beweis folgt dies direkt aus der Definition des Minimums. Hier ist zu viel natürliche Sprache in Gebrauch, was sogar das Verständnis erschwert.
4. Das ist schon gravierender. Denn die Voraussetzung besagt: „Wenn $p(n)$ für ein $n \in \mathbb{N}$ wahr ist, dann ist auch $p(n + 1)$ wahr.“ Offenbar wird hier $n := m - 1$ gesetzt, denn dann ist $n + 1 = m$. Das sollte man nicht auslassen. Und noch wichtiger ist es, $n \in \mathbb{N}$ auch wirklich zu zeigen!
5. Bei „1.“ hat man von „der“ kleinsten Zahl gesprochen. Es ist extrem unsauber, dann zu zeigen, dass man vorher irgendwo eine falsche Voraussetzung benutzt hat. Hätte man gesagt, dass m eine Zahl sein soll, für die $p(m)$ nicht gilt, und dass außerdem für jede andere Zahl x , für die $p(x)$ ebenfalls falsch ist, $x \geq m$ gelten soll, dann hätte man diese Unsauberkeit vermieden. Oder man hätte wenigstens sagen sollen, dass man die Existenz des Minimums annimmt, um sie zu widerlegen.
6. „Für alle $n \in \mathbb{N}$ “, bitte schön. Wenn man „für alle n “ sagt, dann ist n durch den Satz gebunden; die Aussage „ $n \in \mathbb{N}$ “ von oben gilt nicht für dieses n .

Es fällt nicht schwer, den Beweis in die formale mathematische Sprache zu übersetzen:

1. Die „kleinste Zahl“ heißt in der Mathematik „Minimum“. Ein Minimum bezieht sich jedoch immer auf eine Menge (oder Funktion, aber nicht in der Linearen Algebra). Also brauchen wir erst einmal die Menge M , von der wir das Minimum $\min M$ bilden. Dabei können wir auch gleich die Definition von „min“ in Bezug auf natürliche Zahlen nachschlagen, um festzustellen, wann es denn nun existiert: $\forall M \subset \mathbb{N}, M \neq \emptyset : \exists! m \in M : \forall n \in M : n \geq m$; und $\min M := m$.
2. Hier fehlen bloß noch die Symbole.
3. Es bietet sich an, die Voraussetzung $m = \min M$ hier noch einmal zu erwähnen.
4. Mit einem kleinen Trick werden die beschriebenen Lücken geschlossen, aber die Zeile bleibt trotzdem kurz.
5. Wir haben gezeigt, dass das Minimum von M nicht existiert. Ein ganz normaler Widerspruchsbeweis. Denn angewendet auf den Satz über das Minimum, den ich bei „1.“ mehr oder weniger formal hingeschrieben habe, folgt daraus rein formal $M = \emptyset$.
6. Auch hier fehlen wieder nur Symbole.

Das Ergebnis sieht dann so aus:

1. Sei $M := \{n \in \mathbb{N} : p(n) \text{ ist falsch}\}$. Annahme: $M \neq \emptyset$. Dann existiert $m := \min M$.

2. $m \in M \Rightarrow p(m)$ ist falsch $\underset{p(1) \text{ ist wahr}}{\Rightarrow} m \neq 1 \Rightarrow m \geq 2 \Rightarrow m - 1 \in \mathbb{N}$
3. $m = \min M \Rightarrow m - 1 \notin M \Rightarrow p(m - 1)$ ist wahr
4. $\underset{m-1 \in \mathbb{N}}{\Rightarrow} p(m)$ ist wahr (nach Voraussetzung)
5. Widerspruch zu $m \in M$, also Annahme $M \neq \emptyset$ falsch $\Rightarrow M = \emptyset$
6. $\Rightarrow \forall n \in \mathbb{N} : p(n)$

Dieser Beweis ist zwar vielleicht schwerer zu lesen, aber er lässt keine Zweifel aufkommen. Wer will, kann wieder etwas mehr natürliche Sprache einbauen, z.B. Folgepfeile durch „daraus folgt“, „also“, „deshalb“, „daher“, „weil“, „wegen“ usw. ersetzen. Wichtig ist, dass ein Leser jeden einzelnen Schritt ohne den Rest des Beweises nachvollziehen kann. Wenn Ihr Euch sicher seid, dass Eure Beweise dies auch ohne Formalismus leisten, könnt Ihr schließlich wieder fast die Textversion hinschreiben. Aber übt bitte am Anfang das formale Beweisen.

2 Mengenlehre

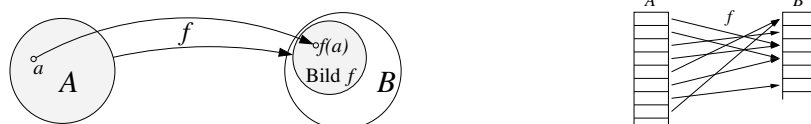
2.1 Mengen

Mengen und ihre Verknüpfungen sollten aus der Schule bekannt sein. Neu ist vielleicht die Schreibweise „ $\{x, p(x)\}$ “, „ $\{x \mid p(x)\}$ “ oder „ $\{x : p(x)\}$ “ für die Menge aller x , für die $p(x)$ gilt. Das sollte man nicht nur wissen, sondern auch damit umgehen können. Siehe z.B. 1.3.2.

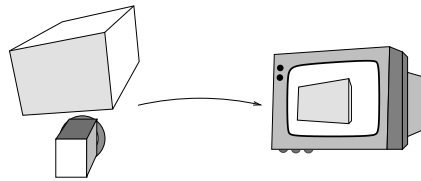
Außerdem wird an der Uni die Potenzmenge $\mathcal{P}(M)$ einer Menge M eingeführt; das ist die Menge aller Teilmengen von M . Aussagen über $\mathcal{P}(M)$ erscheinen oft kompliziert, aber werden sofort einfacher, wenn man daran denkt, dass $A \in \mathcal{P}(M)$ nichts Anderes heißt als $A \subset M$. Zum Beispiel bildet eine Funktion (s.u.) $f : A \rightarrow \mathcal{P}(B)$ einfach jedes Element $a \in A$ auf eine bestimmte Teilmenge $f(a) \subset B$ ab.

2.2 Abbildungen

„Abbildung“ ist nur ein anderes Wort für „Funktion“, das auch in der Schule behandelt wird. Informatikern fällt es nicht schwer, sich Funktionen so vorstellen, dass sie eine „Eingabe“ entgegennehmen und zu diesem Wert eine „Ausgabe“ „berechnen“. An den vielen Anführungszeichen merkt man schon, dass es nicht ganz so einfach ist; z.B. muss man den Funktionswert nicht unbedingt wirklich berechnen können, damit die Funktion existiert. Trotzdem möchte ich gerne eine Funktion $f : A \rightarrow B$ manchmal auf diese zwei Arten darstellen:



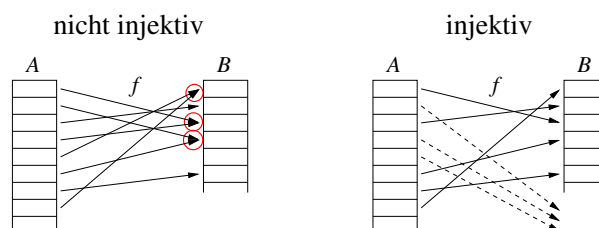
Dass in der Linearen Algebra häufiger der Begriff „Abbildung“ benutzt wird, könnte daran liegen, dass sich die Lineare Algebra in erster Linie mit Anschauungsobjekten beschäftigt. Z.B. ist die Vorschrift, die angibt, wie in einem 3D-Computerprogramm ein Objekt auf dem Bildschirm angezeigt werden soll, sicher eine Funktion bzw. Abbildung (wobei ich hier zugegebenermaßen wichtige Details außer Acht lasse):



Hier werden vielleicht auch die Begriffe „Bild“ und „Urbild“ deutlich: Sei O das Objekt; es kann z.B. eine Teilmenge des Definitionsbereichs der Abbildung, genannt f , sein. Dann ist das Bild von O unter f , also $f(O) = \text{Bild } f|_O$, tatsächlich die entsprechende Menge im Wertebereich, d.h. auf dem Bildschirm. Umgekehrt ist O gerade das Urbild seiner Darstellung auf dem Bildschirm.

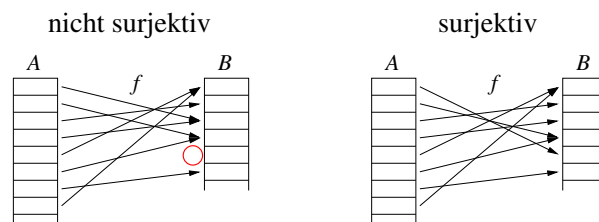
Eine Abbildung heißt „injektiv“, wenn es zu jedem Bildelement nur ein Urbildelement gibt, bzw. zu jedem Element des Wertebereichs *höchstens ein* Urbildelement. (Die genaue Definition könnt Ihr Euch vielleicht noch einmal selbst überlegen; es ist nur eine mathematische Formulierung dieses Satzes.) Im Beispiel oben hätte der Benutzer vielleicht gerne eine injektive Abbildung, dann wüsste er genau, wie die Szenerie wirklich beschaffen ist. Aber wer weiß, ob hinter dem Klotz jemand hockt...

Hier noch ein Vergleich von injektiven und nicht injektiven Abbildungen:



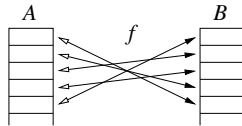
Man sieht, dass bei endlichen Mengen $|B| \geq |A|$ sein muss, und bei unendlichen Mengen wird diese Relation sogar über die Existenz einer injektiven Abbildung definiert.

Eine Abbildung heißt „surjektiv“, wenn jedes Element des Wertebereichs tatsächlich erreicht werden kann, d.h. im Bild der Abbildung liegt bzw. *mindestens ein* Urbildelement hat. Anders ausgedrückt ist das Bild der Abbildung der gesamte Wertebereich. Die obige Abbildung ist surjektiv, denn jeder Punkt auf dem Bildschirm kann ein Objekt enthalten. Genauer gesagt gibt es zu jeder Bildschirmkoordinate als Element des Wertebereichs eine Position, an der ein Objekt stehen kann, um darauf abgebildet zu werden. (Aber natürlich mehr als eine einzige Position, s.o.) Man kann jede Abbildung durch Einschränken des Wertebereichs „surjektiv machen“.



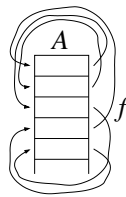
Hier muss nun $|A| \geq |B|$ sein.

Eine injektive und surjektive Abbildung heißt „bijektiv“ oder „invertierbar“ („umkehrbar“). Genau dann existiert nämlich die Umkehrabbildung, denn Bijektivität bedeutet, dass es zu jedem Element des Wertebereichs *genau ein* Urbildelement:

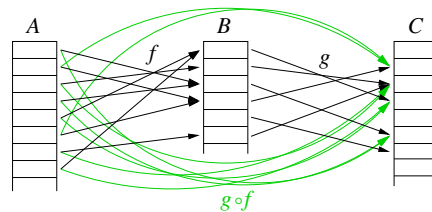


Wegen $|B| \geq |A|$ und $|A| \geq |B|$ ist dann sogar $|A| = |B|$.

Eine bijektive Selbstabbildung (d.h. eine bijektive Abbildung $f : A \rightarrow A$) nennt man „Permutation“. Wem dieser Begriff in einem anderen Zusammenhang bekannt ist, nämlich als das Ändern der Reihenfolge von geordneten Elementen, dem fällt vielleicht auf, dass es genau darum geht: f kann die Elemente aus A nur mischen, aber weder Elemente verschwinden lassen noch mehrere Elemente auf die selbe Stelle abbilden:



Zwei Abbildungen $f : A \rightarrow B$ und $g : B \rightarrow C$ kann man zu einer Abbildung $A \rightarrow C$ verketteten, die mit „ $g \circ f$ “ bezeichnet wird. Daran, dass die Schreibweise „rückwärts“ ist, muss man sich leider gewöhnen. Aber die Reihenfolge stimmt praktisch in allen Fällen, die damit etwas zu tun haben, überein. Wäre sie andersherum, müsste z.B. auch die Matrizenmultiplikation, die ihr noch kennen lernt, umgekehrt verlaufen. Daran kann man sich später vielleicht auch orientieren. Außerdem ist $(g \circ f)(x) = g(f(x))$, d.h. eigentlich ist die Schreibweise „ $f(x)$ “ schon „rückwärts“. Grafisch sieht das so aus:



2.3 Relationen

Eine Relation vergleicht zwei Elemente einer Menge. Man kann sie z.B. als Verknüpfung betrachten, die je zwei Werten entweder „wahr“ oder „falsch“ zuordnet, oder als Teilmenge des cartesischen Produkts der Menge mit sich selbst. Die meisten Symbole, die in Aussagen vorkommen („=“, „<“, „≤“, „C“, „⇒“, „⇔“, ...) sind Relationen. Zur Abgrenzung: „∈“ ist z.B. keine Relation, weil die linke und rechte Seite nicht zur selben Kategorie gehören. „+“ ist auch keine Relation, weil die Verknüpfung nicht „wahr“ oder „falsch“ als Ergebnis liefert, sondern eine Zahl.

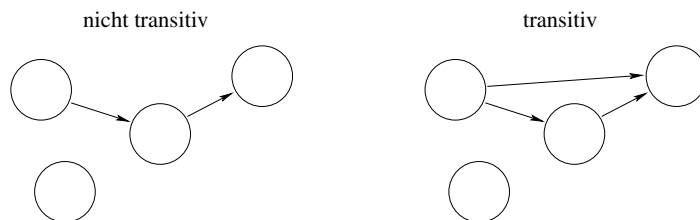
Man klassifiziert Relationen anhand bestimmter Eigenschaften. Am wichtigsten sind Reflexivität (jedes Element steht mit sich selbst in Relation), Symmetrie (es kommt nicht auf die Reihenfolge an) und Transitivität („ $x < y < z \Rightarrow x < z$ “, siehe auch 1.3.1). Sind nämlich diese drei Eigenschaften erfüllt, kann man die Mengenelemente mit Hilfe der Relation in Klassen („Äquivalenzklassen“) einteilen, so dass jedes Element genau mit den Elementen seiner Klasse in Relation steht (siehe 2.4). Die Relation heißt dann entsprechend „Äquivalenzrelation“.

Ist die Menge endlich, dann kann man für die Relation (wie für jede Verknüpfung) eine Verknüpfungstafel aufstellen. Zum Beispiel könnte man eine Relation „ \prec “ auf der Menge $\{a, b\}$ definieren durch $a \prec a, a \prec b, b \not\prec a, b \prec b$. Die dazugehörige Verknüpfungstafel sieht so aus:

\prec	a	b
a	w	w
b	f	w

Per Definition gilt: Um festzustellen, ob $x \prec y$ gilt, sucht man x auf der linken Seite und y auf der oberen Seite; dann kann man das Ergebnis der Verknüpfung im entsprechenden Kästchen ablesen.

Reflexivität und Symmetrie sieht man dann sofort, Transitivität leider nicht. In der Informatik werden Relationen auf endlichen Mengen statt dessen häufig als Graphen dargestellt, indem man die Elemente der Menge in der Zeichenebene verteilt und Pfeile zwischen den Elementen einzeichnet, die in Relation stehen. Bei symmetrischen Relationen ersetzt man die Pfeile durch einfache Linien. In Graphen sieht man die Transitivität besser:



2.4 Klassenbildung

Wenn man eine Menge in Teilmengen aufteilt, so dass jedes Element in genau einer dieser Teilmengen vorkommt, dann nennt man diese Teilmengen „Klassen“. Z.B. könnte man die Menge $\{1, 2, 3, 4, 5\}$ in die Klassen $\{1, 4\}$, $\{2, 3\}$ und $\{5\}$ aufteilen, wenn man Lust dazu hat. Die Menge der Klassen ist dann $\{\{1, 4\}, \{2, 3\}, \{5\}\}$. Die Klasse eines Elements x wird normalerweise mit $[x]$ bezeichnet. In diesem Fall ist z.B. $[1] = [4] = \{1, 4\}$. Auch unendliche Mengen kann man in Klassen aufteilen, z.B. die Menge der ganzen Zahlen in gerade und ungerade Zahlen (dann gibt es die zwei Klassen $[0]$ und $[1]$).

Hat man eine Menge in Klassen unterteilt, kann man über die Zugehörigkeit zur selben Klasse eine Relation „ \approx “ festlegen: $x \approx y \Leftrightarrow [x] = [y]$. D.h.: Zwei Elemente stehen genau dann in Relation, wenn sie in der selben Klasse liegen. Dies ist eine Äquivalenzrelation. Umgekehrt (und das ist der wichtigere Fall) kann man mit einer beliebigen Äquivalenzrelation „ \sim “ eine Menge M in Klassen aufteilen, deren Elemente jeweils miteinander in Relation stehen (siehe 2.3). Die Menge dieser Klassen wird mit M/\sim bezeichnet, eine einzelne Klasse mit $[x]_\sim$ (wenn man die Relation hervorheben will).

Beispiel: Um die ganzen Zahlen in gerade und ungerade Zahlen aufzuteilen, kann man auch erst eine Äquivalenzrelation definieren:

$$x \sim y \Leftrightarrow \exists z \in \mathbb{Z} : x - y = 2 \cdot z$$

Dann gilt z.B. $0 \sim 2 \sim 4 \sim \dots, 1 \sim 3 \sim 5 \sim \dots, 0 \not\sim 1, 1 \not\sim 2$, usw. Das bedeutet $[0]_\sim = [2]_\sim = [4]_\sim = \dots, [1]_\sim = [3]_\sim = [5]_\sim = \dots$ und $[0]_\sim \neq [1]_\sim$. Damit hat man genau zwei Klassen, d.h. $\mathbb{Z}/\sim = \{[0]_\sim, [1]_\sim\}$, wie oben.

Wenn man eine Menge in Äquivalenzklassen einteilt, dann tut man es normalerweise aus dem Grund, dass sich in bestimmten Situationen jedes Element aus der Klasse gleich verhält. Dann möchte man nicht mehr nur mit den einzelnen Elementen rechnen, sondern mit den Klassen selbst. Konkret heißt das, man will Abbildungen und Verknüpfungen für die Klassen definieren.

Z.B. könnte man eine Abbildung

$$f_1 : \mathbb{Z}/\sim \rightarrow \{0, 1\}, k \mapsto \begin{cases} 0, & \text{falls } k = [0]_{\sim} \\ 1, & \text{falls } k = [1]_{\sim} \end{cases}$$

definieren (wobei „ \sim “ die Relation aus dem vorherigen Beispiel ist). Weil es in \mathbb{Z}/\sim eben nur die beiden Elemente $[0]_{\sim}$ und $[1]_{\sim}$ gibt, ist die Abbildung damit vollständig definiert.

Aber wenn die Anzahl der Klassen unendlich groß ist, dann funktioniert das nicht mehr so einfach. Deshalb darf man sich *ausnahmsweise* bei der Definition der Abbildung ein Element der Klasse („Vertreter“ genannt) herausnehmen, ungefähr so:

$$f_2 : \mathbb{Z}/\sim \rightarrow \{0, 1\}, [x]_{\sim} \mapsto x$$

Auf den ersten Blick ist dies tatsächlich eine Abbildung, nämlich die gleiche wie f_1 , denn $f_2([0]_{\sim}) = 0$ und $f_2([1]_{\sim}) = 1$. Aber 0 und 1 sind ja nicht die einzigen Elemente aus $[0]_{\sim}$ bzw. $[1]_{\sim}$! Es ist $[0]_{\sim} = [2]_{\sim}$, und deshalb muss $f_2([0]_{\sim}) = f_2([2]_{\sim})$ sein. (Wenn $x = y$ ist, dann ist immer auch $f(x) = f(y)$, sonst wäre $f(x)$ ja gar nicht eindeutig bestimmt.) Auf der anderen Seite sagt die Definition von f_2 aber $f_2([0]_{\sim}) = 0$ und $f_2([2]_{\sim}) = 2$.

Also kann man die Abbildung gar nicht so definieren. Man sagt, die Abbildung ist nicht „wohldefiniert“. Das Problem der Wohldefiniertheit ergibt sich erst dann, wenn man sich die Ausnahme zunutze macht, dass man die Abbildung über einen Vertreter der Klasse und nicht über die Klasse selbst definieren darf. Dann muss man eben explizit dafür sorgen, dass das Resultat der Abbildung nicht davon abhängt, welchen Vertreter der Klasse man benutzt. Korrekt wäre:

$$f_3 : \mathbb{Z}/\sim \rightarrow \{0, 1\}, [x]_{\sim} \mapsto \begin{cases} 0, & \text{falls } x \text{ gerade ist} \\ 1, & \text{falls } x \text{ ungerade ist} \end{cases}$$

Diese Abbildung ist wohldefiniert, denn nimmt man sich aus $[x]_{\sim}$ zusätzlich zu x einen weiteren Vertreter x' , dann ist x' gerade, falls x gerade ist, und ungerade, falls x ungerade ist. Um dies formal zu beweisen, müsste man ausnutzen, dass $x \sim x'$ ist, und dies in die Definition von „ \sim “ einsetzen.

3 Algebra

3.1 Gruppen

3.1.1 Definition

Eine Halbgruppe ist eine Menge mit assoziativer innerer Verknüpfung (d.h. sie verknüpft zwei Elemente der Menge zu einem neuen Element der gleichen Menge). Eine Gruppe ist eine Halbgruppe mit neutralem Element und inversen Elementen. Eine Untergruppe einer Gruppe ist eine Untermenge, die selbst wieder Gruppe ist.

Diese Definition ist erst einmal ziemlich abstrakt; man kann sich unter einer Gruppe schwer etwas vorstellen. Also sollte man zunächst untersuchen, was eine Gruppe ist und was nicht. Dabei bekommt

man ziemlich viele verschiedene Ergebnisse, die außer den (abstrakten) Gruppeneigenschaften nicht viel miteinander zu tun haben. Lohnenswerter ist es, die Eigenschaften von Gruppen zu untersuchen, um wenigstens zu wissen, wofür man Gruppen braucht. Leider (oder glücklicherweise?) werden sie erst in der Algebra wirklich wichtig. Ein Beispiel für ein gruppentheoretisches Ergebnis, das Ihr wahrscheinlich kennen lernen werdet, ist der Satz von Fermat-Euler.

Normalerweise muss man beim Beweis, dass eine Menge M mit Verknüpfung „ \circ “ Gruppe ist, sämtliche Gruppenaxiome testen:

- Das Ergebnis der Verknüpfung ist tatsächlich immer ein Element der Menge: $\forall m_1, m_2 \in M : m_1 \circ m_2 \in M$. Man spricht von der „Abgeschlossenheit“ der Verknüpfung.
- Die Verknüpfung ist assoziativ: $\forall m_1, m_2, m_3 \in M : (m_1 \circ m_2) \circ m_3 = m_1 \circ (m_2 \circ m_3)$.
- Es gibt ein Element, das sowohl rechts- als auch linksneutral ist: $\exists e \in M : \forall m \in M : m \circ e = e \circ m = m$. Oft findet man ein Element, das auf einer Seite neutral ist. Falls die Verknüpfung nicht zufällig kommutativ ist, muss man nachweisen, dass es auch auf der anderen Seite neutral ist.
- Zu jedem Element gibt es ein Element, das sowohl rechts- als auch linksinvers ist: $\forall m \in M : \exists m^{-1} \in M : m \circ m^{-1} = m^{-1} \circ m = e$. Auch hier tritt oft der Fall auf, dass man zu einem beliebigen Element ein anderes findet, das auf einer Seite invers ist, und man muss noch zeigen, dass es auch auf der anderen Seite invers ist.

Theoretisch reichen für neutrale und inverse Elemente auch etwas schwächere Bedingungen, aber da man im Allgemeinen diese Elemente direkt angeben kann, sollte man die Axiome so zeigen, wie sie hier stehen. Für Untergruppen U muss man die Assoziativität nicht mehr zeigen. Auch weiß man genau, welche Elemente neutral und jeweils invers sind. Das heißt aber nicht, dass man gar nichts mehr zeigen muss; vielmehr ergeben sich andere Probleme:

- Die Verknüpfung ist nicht unbedingt abgeschlossen, wenn man sie auf die Untermenge einschränkt. D.h. man muss zeigen, dass für alle $u_1, u_2 \in U$ auch wirklich $u_1 \circ u_2 \in U$ ist. Im Allgemeinen ist das Produkt nur ein Element aus M .
- Das neutrale Element muss auch in der Untergruppe liegen: $e \in U$.
- Zu jedem Element $x \in U$ muss man zeigen, dass $x^{-1} \in U$ ist.

Auch hier kann man es sich theoretisch wieder einfacher machen, indem man nur zeigt, dass U nicht leer ist, und dass für alle $u_1, u_2 \in U$ das Produkt $u_1 \circ u_2^{-1} \in U$ ist.

Dabei muss man sich häufig erst klar machen, was überhaupt „ $\in U$ “ bedeutet. Für $U = \{x \in M : p(x)\}$ weist man am besten $p(e)$ nach und schreibt dann: „Seien $u_1, u_2 \in U$, d.h. $p(u_1)$ und $p(u_2)$. Zu zeigen: $u_1 \circ u_2^{-1} \in U$, d.h. $p(u_1 \circ u_2^{-1})$“ (Siehe auch 1.3.2.) Ist p ein etwas komplizierteres Prädikat, wird das Ganze sonst ziemlich unübersichtlich.

3.1.2 Erzeugnis

Viele Gruppen, darunter alle endlichen, werden von endlich vielen Elementen „erzeugt“. Das muss man sich so vorstellen, dass man diese Elemente und die Gruppenverknüpfung vorgibt und dann die Elemente sozusagen „arbeiten lässt“, um eine Gruppe zu bilden. D.h. man fordert die Existenz von neutralem Element und Inversen, verknüpft jedes Element mit jedem anderen, und erhält damit eine Gruppe. Für das Erzeugnis einer Menge M schreibt man $[M]$ oder $\langle M \rangle$. M selbst nennt man „Erzeugendensystem“.

Z.B. ist $(\mathbb{Z}, +) = [\{1\}]$, denn jedes Element lässt sich als $1 + 1 + \dots + 1$ darstellen, oder als Inverses einer solchen Zahl.

Ich schreibe dies aus folgendem Grund: Kennt man ein Erzeugendensystem einer Gruppe, dann gibt dies Auskunft über die Struktur der Gruppe. Z.B. kann man Untergruppen finden, indem man aus dem Erzeugendensystem Elemente weglässt.

Jede Permutation lässt sich als Produkt (Verkettung) von Transpositionen schreiben, d.h. von Permutationen, die nur zwei Elemente vertauschen. Für die Gruppe S_3 aller bijektiven Selbstabbildungen von $\{1, 2, 3\}$ gilt also z.B.:

$$S_3 = \left[\left\{ \begin{pmatrix} 1 \mapsto 2 \\ 2 \mapsto 1 \\ 3 \mapsto 3 \end{pmatrix}, \begin{pmatrix} 1 \mapsto 1 \\ 2 \mapsto 3 \\ 3 \mapsto 2 \end{pmatrix}, \begin{pmatrix} 1 \mapsto 3 \\ 2 \mapsto 2 \\ 3 \mapsto 1 \end{pmatrix} \right\} \right]$$

Nimmt man nur das Erzeugnis der ersten Transposition, also $\left[\left\{ \begin{pmatrix} 1 \mapsto 2 \\ 2 \mapsto 1 \\ 3 \mapsto 3 \end{pmatrix} \right\} \right]$, hat man eine Untergruppe, die S_2 entspricht.

3.1.3 Homomorphismen

Abbildungen zwischen Gruppen sind immer dann interessant, wenn sie die Gruppenstrukturen erhalten; man nennt sie dann „Homomorphismen“. Aber was heißt das genau? Wichtige Eigenschaften sind z.B.:

- Das neutrale Element der einen Gruppe wird auf das neutrale Element der anderen abgebildet.
- Das Inverse eines Elements wird auf das Inverse des Bildelements abgebildet.
- Untergruppen werden auf Untergruppen abgebildet. Insbesondere sind Bild und Kern eines Homomorphismus Untergruppen.
- Man kann bei Rechnungen zuerst den Homomorphismus anwenden, im Bild rechnen, und das Ergebnis wieder in die ursprüngliche Gruppe zurück transformieren.
- Die Verkettung von Homomorphismen ist wieder ein Homomorphismus. Ist ein Homomorphismus bijektiv, ist auch die Umkehrabbildung ein Homomorphismus.

Alles lässt sich in der Eigenschaft $f(x \circ y) = f(x) \star f(y)$ zusammenfassen, wobei f der Homomorphismus und „ \circ “ und „ \star “ die jeweiligen Gruppenverknüpfungen sind. Wichtiger sind aber eigentlich

die anderen Eigenschaften. Soll man z.B. nachweisen, dass eine Teilmenge einer Gruppe eine Untergruppe ist, dann kann man sich überlegen, ob man nicht einen Homomorphismus kennt, dessen Kern oder Bild gerade diese Teilmenge ist.

Einen bijektiven Homomorphismus nennt man „Isomorphismus“. Das Besondere an Isomorphismen ist, dass man Rechnungen vollständig in eine andere Gruppe übertragen kann, also gewissermaßen die beiden Gruppen als äquivalent betrachtet. Das spielt später bei Vektorräumen (statt Gruppen) eine sehr große Rolle, und zwar beim so genannten Basiswechsel. Es kann aber auch schon bei Gruppen hilfreich sein:

Soll man z.B. in einer Gruppe einen Term x^k ausrechnen, doch die k -fache Multiplikation von x ist schwierig, dann ist es vielleicht einfacher, ein y zu finden, so dass $z := y^{-1} \cdot x \cdot y$ eine einfache Form hat. Dann kann man z^k ausrechnen und das Ergebnis wieder zurückzutransformieren (also $y \cdot z^k \cdot y^{-1}$). Man kann zwar leicht nachrechnen, dass dies das Gleiche ist (die $y \cdot y^{-1}$ kürzen sich jeweils weg), aber man sieht es sogar ohne Rechnung und kommt vielleicht eher darauf, wenn man weiß, dass die Abbildung $f : G \rightarrow G, x \mapsto y^{-1} \cdot x \cdot y$ ein Isomorphismus ist.

Man nennt zwei Gruppen „isomorph“, wenn es einen Isomorphismus zwischen ihnen gibt. D.h. um zu zeigen, dass zwei Gruppen isomorph sind, sollte man diesen Isomorphismus angeben. Aber wie findet man eigentlich Homomorphismen? Denn überlegt man sich erst eine Abbildung, dann ist sie wahrscheinlich noch kein Homomorphismus. Die einzige Eigenschaft, auf die man sofort achten kann, ist, dass das neutrale Element der einen Gruppe auf das neutrale Element der anderen Gruppe abgebildet werden muss.

Praktischer ist es, ein Erzeugendensystem M der Gruppe G zu kennen (siehe 3.1.2), wenn man einen Homomorphismus $f : G \rightarrow H$ finden will. Denn dann muss man nur angeben, worauf die Elemente von M abgebildet werden sollen. Der Rest ergibt sich automatisch, weil sich jedes $g \in G$ als Produkt von Elementen aus M sowie deren Inversen schreiben lässt; und die Eigenschaft $f(x \circ y) = f(x) \star f(y)$ definiert dann gerade jedes $f(g)$.

Muss man zeigen, dass zwei Gruppen nicht isomorph sind, d.h., dass es *keinen* Isomorphismus zwischen ihnen gibt, dann hört sich das erst einmal schwer an. (Es sei denn, sie sind endlich und haben unterschiedlich viele Elemente.) Aber mit dem Ergebnis aus dem vorherigen Abschnitt sieht man, dass es gar nicht so schlimm ist, wenn man ein Erzeugendensystem der einen Gruppe kennt. Sind die Gruppen klein, dann bildet man die einzelnen Elemente des Erzeugendensystems einfach nacheinander auf alles ab, das es gibt, rechnet damit evtl. noch ein paar mehr Bildelemente aus, und sucht jeweils eine Stelle, an der die Injektivität verletzt ist. Oft kann man es sich aber auch viel einfacher machen, wenn man weiß, worin sich die Struktur der Gruppen unterscheidet. Gibt es z.B. in der einen Gruppe ein Element $g \neq e$ (Neutralelement), das sein eigenes Inverses ist, in der anderen Gruppe aber nicht, dann kann es keinen Isomorphismus geben. Denn es müsste ja $f(g) = f(g^{-1}) = (f(g))^{-1}$ sein, also $f(g) = e$. Dann wäre f nicht mehr bijektiv.

3.1.4 Der Homomorphiesatz

Obwohl dem Homomorphiesatz für Gruppen in der Linearen Algebra keine besonders große Bedeutung zukommt, kann es nicht schaden, ihn verstanden zu haben. Denn es gibt einen analogen Homomorphiesatz für Vektorräume, und der ist an einigen Stellen unersetzlich. Außerdem verrät er Einiges über das Wesen von Homomorphismen.

Um zu erklären, worum es geht, fange ich erst einmal wieder bei ganz normalen Mengen und Abbildungen an, also *nicht* bei Gruppen. Ich hatte bereits geschrieben, dass man eine Abbildung „surjektiv machen“ kann, indem man den Wertebereich auf das Bild einschränkt. Mathematisch

korrekt formuliert kann man eine Abbildung $f : A \rightarrow B$ aufteilen in eine surjektive Abbildung $\hat{f} : A \rightarrow \text{Bild } f$, $a \mapsto f(a)$ und eine Inklusionsabbildung $i : \text{Bild } f \rightarrow B$, $b \mapsto b$, so dass $f = i \circ \hat{f}$ ist. Die Frage ist nun, ob man eine Abbildung auch „injektiv machen“ kann, und damit insgesamt sogar bijektiv.

Natürlich ist das kein Problem. Und zwar muss man nur die Menge A in Klassen einteilen, so dass alle Elemente einer Klasse jeweils auf das gleiche Element aus B abgebildet werden. Mit der Äquivalenzrelation „ \sim “, die durch $a_1 \sim a_2 :\Leftrightarrow f(a_1) = f(a_2)$ definiert wird, ist A/\sim gerade die Menge dieser Klassen. Die Abbildung $\bar{f} : A/\sim \rightarrow B$, $[a]_{\sim} \mapsto f(a)$ ist dann wohldefiniert und „entspricht“ f . Genauer gesagt ist mit der kanonischen Abbildung $k : A \rightarrow A/\sim$, $a \rightarrow [a]_{\sim}$, die einfach jedes Element in seine Klasse abbildet, $f = \bar{f} \circ k$.

Insgesamt bekommt man also ohne Weiteres eine bijektive Abbildung $\tilde{f} : A/\sim \rightarrow \text{Bild } f$, $[a]_{\sim} \mapsto f(a)$, so dass $f = i \circ \tilde{f} \circ k$ ist. Ein kleines Diagramm zeigt die einzelnen Schritte:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ k \downarrow & \text{//} & \uparrow i \\ A/\sim & \xrightarrow{\tilde{f}} & \text{Bild } f \end{array}$$

Die drei Striche bedeuten, dass das Diagramm kommutiert, d.h. es ist egal, welchen Weg man geht. Noch einmal: Die Abbildung \tilde{f} ist bis auf die Formalitäten identisch mit f , aber bijektiv. Um das zu erreichen, braucht man die Klassen sowie die sehr einfachen Abbildungen k und i . Bis hier hin ist übrigens noch nichts Besonderes passiert.

Jetzt möchte ich das Ganze auf Gruppen und deren Homomorphismen übertragen. Natürlich könnte f auch ein Homomorphismus sein. Die Frage ist, ob auch \tilde{f} dadurch ein Homomorphismus wird; wenn nicht, dann macht dieses Verfahren für Gruppen keinen Sinn.

Dafür muss man erst einmal klären, wie man den Mengen A/\sim und $\text{Bild } f$ überhaupt eine Gruppenstruktur verleihen will. Bei $\text{Bild } f$ ist es kein Problem, denn es ist eine Untergruppe von B . Wenn man sich „ \sim “ genauer ansieht, dann fällt vielleicht auf, dass zwei Elemente genau dann in Relation stehen, wenn sie sich nur um ein Element aus dem Kern von f unterscheiden; man schreibt für „ A/\sim “ dann auch „ $A/\text{Kern } f$ “. Auf A/\sim versucht man, eine vertreterweise Verknüpfung $[x]_{\sim} \bullet [y]_{\sim} := [x \circ y]_{\sim}$ einzuführen, wobei „ \circ “ die Gruppenverknüpfung von A sein soll. Der Homomorphiesatz sagt nun: Die Verknüpfung „ \bullet “ ist wohldefiniert, $(A/\sim, \bullet)$ ist eine Gruppe, und k , \tilde{f} und i sind tatsächlich Gruppenhomomorphismen.

Eigentlich gehört nur die letzte Behauptung zum Homomorphiesatz. Für die beiden anderen ist streng genommen die Normalteilereigenschaft des Kerns verantwortlich; auf Normalteiler möchte ich aber hier nicht eingehen. Auf jeden Fall sollte man sich merken, dass der Kern eines Homomorphismus genau das ist, was ihn an der Injektivität hindert. Es ist nicht nur so, dass ein Homomorphismus genau dann injektiv ist, wenn der Kern nur aus dem neutralen Element besteht, sondern man kann den Kern sogar „entfernen“ und den Homomorphismus damit „injektiv machen“.

Wobei man dazu sagen muss, dass durch dieses „Entfernen“ die Gruppe mehr Elemente verliert als nur die, die im Kern liegen (und nicht neutral sind). Das Entfernen findet eben nicht auf der Elementebene statt, sondern auf der Ebene der Untergruppen; man faktorisiert A nach dem Kern. Es ist eher ein „Teilen“ als ein „Subtrahieren“; deshalb schreibt man auch „ A/\sim “ bzw. „ $A/\text{Kern } f$ “.

Und zum Schluss möchte ich noch ein kleines Anwendungsbeispiel liefern. Übrigens sind praktisch alle Aufgaben zum Homomorphiesatz (für Gruppen oder Vektorräume) von ähnlicher Art. Und wenn man den Homomorphiesatz verstanden hat, dann kann man dabei ohne viel Aufwand wichtige Punkte sammeln.

„Seien G, H_1 und H_2 Gruppen, $f_1 : G \rightarrow H_1$ und $f_2 : G \rightarrow H_2$ Gruppenhomomorphismen mit $\text{Kern } f_1 = \text{Kern } f_2 =: K$, und f_1 surjektiv. Man finde einen Homomorphismus $h : H_1 \rightarrow H_2$ mit $h \circ f_1 = f_2$.“

Die Aufgabe wäre ja sehr leicht, wenn doch nur f_1 bijektiv wäre. Dann könnte man einfach schreiben: „ $h := f_2 \circ f_1^{-1}$ “. Da f_1 nicht injektiv ist, muss man es „injektiv machen“, also G entsprechend faktorisieren. Nach dem Homomorphiesatz existieren Homomorphismen $\bar{f}_1 : G/K \rightarrow H_1$ (bijektiv) und $\bar{f}_2 : G/K \rightarrow H_2$, so dass mit der kanonischen Abbildung $k : G \rightarrow G/K$ gilt: $f_1 = \bar{f}_1 \circ k$ und $f_2 = \bar{f}_2 \circ k$. Mit $h := \bar{f}_2 \circ \bar{f}_1^{-1}$ ist $h \circ f_1 = \bar{f}_2 \circ \bar{f}_1^{-1} \circ \bar{f}_1 \circ k = \bar{f}_2 \circ k = f_2$.

3.2 Ringe und Körper

3.2.1 Ringe

Ein Ring ist eine Menge M mit zwei Verknüpfungen „+“ und „ \cdot “, so dass $(M, +)$ eine abelsche Gruppe ist, (M, \cdot) eine Halbgruppe, und die üblichen Distributivgesetze erfüllt sind. Das neutrale Element von $(M, +)$ wird mit „0“ bezeichnet. Besitzt (M, \cdot) ein neutrales Element, wird es mit „1“ bezeichnet. Man spricht dann von einem „Ring mit 1“. In der Linearen Algebra kommen fast nur Ringe mit 1 vor.

Mit Ringen kann man in vielen Fällen so rechnen wie mit Zahlen; insbesondere ganzen Zahlen. Man muss allerdings aufpassen, denn die Multiplikation muss nicht unbedingt kommutativ sein. Es stellt sich heraus, dass die Addition meistens etwas mit der Addition von Zahlen zu tun hat, die Multiplikation aber nicht unbedingt. Z.B. wird sich später zeigen, dass spezielle Abbildungen (also keine Zahlen) einen Ring bilden, wenn man die Verkettung als Multiplikation benutzt. Weil die Addition in einem Ring schon ziemlich festgelegt ist, beziehen sich übrigens Eigenschaften von Ringen (wie Kommutativität, Inverse, Nullteilerfreiheit, Teilbarkeit mit Rest, usw.) fast immer auf die Multiplikation.

Es ist wichtig zu wissen, was man in Ringen tun darf. Die Distributivgesetze geben einem die Möglichkeit, Faktoren auszuklammern. Bestimmte Formeln, die man normalerweise mit reellen Zahlen assoziiert, gelten daher auch in Ringen. Z.B. kann man leicht nachrechnen, dass die binomischen Formeln in allen kommutativen Ringen mit 1 gelten, wenn man $2 := 1 + 1$ setzt.

\mathbb{Z} , \mathbb{Q} und \mathbb{R} sind natürlich Ringe. Interessanterweise sind auch die $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$ Ringe. (Insbesondere sind die Verknüpfungen auf den Klassen wohldefiniert.) Es werden bald noch mehr Ringe folgen.

In einem Ring M mit 1 wird die Menge der Elemente, die ein (multiplikatives) Inverses besitzen, mit „ M^\times “ bezeichnet. (M^\times, \cdot) ist immer eine Gruppe (die so genannte „Einheitengruppe“); dies ist eine ziemlich wichtige Eigenschaft von Ringen. Z.B. ist $\mathbb{Z}^\times = \{1, -1\}$ und $\mathbb{Z}_4^\times = \{[1], [3]\}$.

3.2.2 Körper

Ein Körper ist ein kommutativer Ring mit 1, dessen Einheitengruppe die gesamte Menge M außer der 0 umfasst ($M^\times = M \setminus \{0\}$). D.h. auch $(M \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

In einem Körper kann man nun endgültig so rechnen wie mit den Zahlen, die man aus der Schule kennt. Da in einem Körper jedes Element außer der 0 ein (multiplikatives) Inverses besitzt, darf man in einem Körper insbesondere teilen. \mathbb{Q} und \mathbb{R} sind Körper, und auch die komplexen Zahlen \mathbb{C} als Erweiterung von \mathbb{R} bilden einen Körper. (Wäre das nicht so, dann wäre \mathbb{C} wahrscheinlich keine

sinnvolle Erweiterung von \mathbb{R} .) Die \mathbb{Z}_m bilden genau dann einen Körper, wenn m eine Primzahl ist, und werden dann häufig auch \mathbb{F}_m genannt.

Das Inverse eines Elements $0 \neq a + i \cdot b \in \mathbb{C}$ ist übrigens:

$$\begin{aligned} \frac{1}{a + i \cdot b} &= \frac{1}{a + i \cdot b} \cdot \frac{a - i \cdot b}{a - i \cdot b} = \frac{a - i \cdot b}{a^2 - (i \cdot b)^2} = \frac{a - i \cdot b}{a^2 - i^2 \cdot b^2} = \frac{a - i \cdot b}{a^2 - (-1) \cdot b^2} = \frac{a - i \cdot b}{a^2 + b^2} = \\ &= \frac{a}{a^2 + b^2} + i \cdot \frac{-b}{a^2 + b^2} \end{aligned}$$

Die Inversen in einem der Primkörper \mathbb{F}_m (oder überhaupt in einem \mathbb{Z}_m) zu finden, ist nicht so einfach. Natürlich ist $[1]^{-1} = [1]$ und $[m-1]^{-1} = [-1]^{-1} = [-1] = [m-1]$, d.h. $[1]$ und $[m-1]$ sind immer selbstinvers. Die anderen Inversen bekommt man entweder durch Probieren oder mit dem euklidischen Algorithmus: Möchte man das Inverse der Klasse $[z]$ bestimmen, berechnet man den ggT von z und m mit dem euklidischen Algorithmus. Das Inverse existiert genau dann, wenn dieser ggT 1 ist. Durch Rücksubstitution erhält man eine Darstellung von 1 als $k \cdot z + j \cdot m$. Dann ist $[z]^{-1} = [k]$.

Beispiel: Um das Inverse von $[4] \in \mathbb{F}_{11}$ zu bestimmen, führt man den euklidischen Algorithmus mit 4 und 11 durch:

$$\begin{aligned} 11 &= 2 \cdot 4 + 3 \\ 4 &= 1 \cdot 3 + 1 \end{aligned}$$

Damit hat man $1 = 4 - 1 \cdot 3 = 4 - 1 \cdot (11 - 2 \cdot 4) = 3 \cdot 4 + j \cdot 11 \Rightarrow [4]^{-1} = [3]$.

3.2.3 Unterringe und Homomorphismen

Um nachzuweisen, dass eine Teilmenge Unterring oder -körper ist, muss man nur die vorkommenden (Halb-)Gruppen betrachten. Denn die Distributivgesetze bleiben natürlich auch weiterhin erfüllt. Bei Ringen mit 1, die keine Körper sind, darf man nicht vergessen nachzuweisen, dass die 1 in der Teilmenge enthalten ist.

Ein Homomorphismus zwischen Ringen oder Körpern ist einfach eine Abbildung $f : R \rightarrow S$, die Homomorphismus bezüglich der entsprechenden (Halb-)Gruppen ist. Ein Homomorphismus von Halbgruppen ist dabei genauso definiert wie ein Homomorphismus von Gruppen. Bei Ringen mit 1, die keine Körper sind, muss man speziell noch nachweisen, dass $f(1_R) = 1_S$ ist, denn bei Halbgruppenshomomorphismen ist dies nicht selbstverständlich.

3.2.4 Matrizen

Matrizen sind gewissermaßen zweidimensionale Tupel über einem Ring R mit 1, für die neben der komponentenweisen Addition noch eine spezielle Multiplikation eingeführt wird. Diese funktioniert folgendermaßen:

Seien Matrizen $A := ((a_{ij})) \in R^{k \times l}$, $B := ((b_{ij})) \in R^{l \times m}$, $C := ((c_{ij})) := A \cdot B \in R^{k \times m}$ gegeben. Dann berechnet sich ein beliebiges c_{ij} wie folgt:

$$\begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{il} \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} \cdots & b_{1j} & \cdots \\ \cdots & b_{2j} & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & b_{lj} & \cdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \cdots & c_{ij} = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + \dots + a_{il} \cdot b_{lj} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Es gibt den schönen Trick, die rechte Matrix höher zu schreiben, so dass das Ergebnis darunter passt. Dann sieht man sofort, welche Zeile von A und Spalte von B man benutzen muss:

$$\begin{pmatrix} \dots & b_{1j} & \dots \\ \dots & b_{2j} & \dots \\ & \vdots & \\ \dots & b_{lj} & \dots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{il} \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots \\ \dots & c_{ij} & \dots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

Was viele nicht wissen: Eine Matrix kann man beliebig in Teilmatrizen unterteilen. In den Fällen, in denen die Matrizenmultiplikation dann noch definiert ist, ist das Ergebnis genau das Gleiche. Z.B. sieht man sofort, dass das Ergebnis für c_{ij} gleich bleibt, wenn man A in ihre Zeilen und/oder B in ihre Spalten unterteilt: Wird A in Zeilen unterteilt, dann erhält man im Prinzip eine $k \times 1$ -Matrix, deren Einträge wiederum Matrizen aus $R^{1 \times l}$ sind. Streng genommen sind diese Matrizen natürlich keine Ringelemente; die Hauptsache ist aber, dass die Multiplikation funktioniert. Das Produkt von $X := ((x_j)) \in R^{1 \times l}$ mit $Y := ((y_i)) \in R^{l \times 1}$ ist ja definiert, und zwar so:

$$(x_1 \quad x_2 \quad \dots \quad x_l) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{pmatrix} = (x_1 \cdot y_1 + x_2 \cdot y_2 + \dots + x_l \cdot y_l)$$

Das ist ein Spezialfall, den man sich leicht merken kann, und daraus lässt sich die allgemeine Matrizenmultiplikation komplett ableiten:

$$\begin{pmatrix} \dots & \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{lj} \end{pmatrix} & \dots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots \\ (a_{i1} \quad a_{i2} \quad \dots \quad a_{il}) \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots \\ \dots & (c_{ij}) & \dots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

Für die Zukunft ist eher der Fall wichtig, dass nur die rechte Matrix aus einer einzelnen Spalte besteht. D.h. man sollte sich die allgemeine Matrizenmultiplikation vielleicht so einprägen, dass man nur die rechte Matrix B in ihre Spalten unterteilt, die linke aber so lässt. Das Produkt von $A = ((a_{ij})) \in R^{k \times l}$ mit $Y = ((y_i)) \in R^{l \times 1}$ ist auch noch relativ übersichtlich hinzuschreiben:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2l} \\ a_{31} & a_{32} & \dots & a_{3l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kl} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{pmatrix} = \begin{pmatrix} a_{11} \cdot y_1 + a_{12} \cdot y_2 + \dots + a_{1l} \cdot y_l \\ a_{21} \cdot y_1 + a_{22} \cdot y_2 + \dots + a_{2l} \cdot y_l \\ a_{31} \cdot y_1 + a_{32} \cdot y_2 + \dots + a_{3l} \cdot y_l \\ \vdots \\ a_{k1} \cdot y_1 + a_{k2} \cdot y_2 + \dots + a_{kl} \cdot y_l \end{pmatrix}$$

Vielleicht ist dem Einen oder Anderen schon aufgefallen, dass man damit ein lineares Gleichungssystem für die Variablen y_1, \dots, y_l sehr schön und einfach als $A \cdot Y = B$ mit $B \in R^{k \times 1}$ schreiben kann, wobei A und B fest sind. Das ist aber keineswegs der einzige Anwendungsfall.

Übrigens macht es in diesem einen Fall manchmal Sinn, die linke Matrix A in Spalten zu unterteilen (anstatt in Zeilen). Dann erhält man gewissermaßen eine $1 \times l$ -Matrix, deren Einträge Elemente aus $R^{k \times 1}$ sind:

$$\begin{aligned} \left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{pmatrix} \quad \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix} \quad \cdots \quad \begin{pmatrix} a_{1l} \\ a_{2l} \\ \vdots \\ a_{kl} \end{pmatrix} \right) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{pmatrix} &= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{pmatrix} \cdot y_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix} \cdot y_2 + \cdots + \begin{pmatrix} a_{1l} \\ a_{2l} \\ \vdots \\ a_{kl} \end{pmatrix} \cdot y_l = \\ &= \begin{pmatrix} a_{11} \cdot y_1 + a_{12} \cdot y_2 + \cdots + a_{1l} \cdot y_l \\ a_{21} \cdot y_1 + a_{22} \cdot y_2 + \cdots + a_{2l} \cdot y_l \\ \vdots \\ a_{k1} \cdot y_1 + a_{k2} \cdot y_2 + \cdots + a_{kl} \cdot y_l \end{pmatrix} \end{aligned}$$

Ich hoffe, das zeigt, dass Matrizen mit dieser Multiplikation für viele verschiedene Zwecke geeignet sind. Aber der Hauptgrund, warum die Multiplikation gerade auf diese Weise definiert ist, dürfte sein, dass die quadratischen Matrizen $R^{n \times n}$ mit der komponentenweisen Addition und dieser Multiplikation selbst wieder einen Ring mit 1 bilden. Das Einselement (die sogenannte „Einheitsmatrix“) ist, wie man leicht nachrechnen kann, die Matrix:

$$E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Übrigens gelten die Assoziativ- und Distributivgesetze sowie die Neutralität von E_n auch für nicht-quadratische Matrizen in den Fällen, in denen die Multiplikation jeweils definiert ist.

Die Einheitengruppe des Matrizenrings wird mit $GL(n, R)$ oder $GL_n(R)$ bezeichnet. Normalerweise ist $GL(n, R) \cup \{0\}$, wobei 0 die Nullmatrix bezeichnet, kein Körper. Allerdings kann man die komplexen Zahlen sehr schön als spezielle invertierbare reelle 2×2 -Matrizen definieren, so dass das Produkt zweier komplexer Zahlen dem Matrizenprodukt entspricht. Dabei wird $1 \in \mathbb{R}$ mit E_2 identifiziert und $i := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ gesetzt. Die komplexe Zahl $a + i \cdot b$ ist dann also die Matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Als „transponierte“ Matrix A^T bezeichnet man die Matrix, die aus A durch Spiegelung an der Diagonalen hervorgeht. Matrizen mit $A^T = A$ heißen „symmetrisch“. Falls der Grundring R kommutativ ist, gilt $(A \cdot B)^T = B^T \cdot A^T$ für alle multiplizierbaren Matrizen A und B über R . Außerdem überträgt sich die Inversenbildung: $(A^T)^{-1} = (A^{-1})^T$, falls A invertierbar ist.

3.2.5 Polynome

Die Polynome, wie sie in der Linearen Algebra behandelt werden, sind eine Erweiterung der aus der Schule bekannten. Die erste Erweiterung ist, dass man Polynome über beliebigen Ringen definieren kann. Die zweite Erweiterung ist eine Abstraktion: Aus der Polynomfunktion $p(x) = a_0 \cdot x^0 + a_1 \cdot x^1 + a_2 \cdot x^2 + \cdots + a_n \cdot x^n$ ($n \in \mathbb{N}_0, a_i \in R \forall i \in \{0, \dots, n\}, x \in R$) wird dabei ein abstraktes Objekt

$p = a_0 \cdot X^0 + a_1 \cdot X^1 + a_2 \cdot X^2 + \dots + a_n \cdot X^n$, in dem auch die Größe X abstrakt ist. Diese p bilden einen kommutativen Ring $R[X]$ mit 1. Ist $a_n \neq 0$, heißt n der „Grad“ von p . Ist spezieller $a_n = 1$, heißt p „normiert“.

Hintergrund dieser Abstraktion ist zum Einen, dass es im Allgemeinen keine Bijektion zwischen Polynomfunktionen und Polynomen gibt (in \mathbb{R} schon), und zum Anderen, dass man dann in Polynome mit Koeffizienten in R nicht nur Elemente aus R , sondern z.B. auch Matrizen und bestimmte Abbildungen über R einsetzen kann. Es gibt bestimmte Bedingungen, die dafür erfüllt sein müssen; diese gelten aber z.B. für Matrizen über dem selben Ring automatisch.

Das klingt vielleicht ein bisschen kompliziert, aber man muss sich eigentlich nur merken, dass man z.B. auch Matrizen in Polynome über dem Grundring einsetzen kann. Dabei muss man allerdings aufpassen: Die Größe X^0 wird oft weggelassen, d.h. mit 1 identifiziert. Für jede quadratische Matrix A ist A^0 die Einheitsmatrix, d.h. man muss den absoluten Term a_0 noch mit der Einheitsmatrix multiplizieren (wie erwartet).

Eine wichtige Eigenschaft des Polynomrings über einem Körper ist es, dass man eine Division mit Rest durchführen kann. Auf diese Weise kann man für jede Nullstelle einen Faktor aus dem Polynom abspalten. Die Polynomdivision wird häufig in der Schule behandelt und funktioniert im Prinzip genauso wie die Division ganzer Zahlen, deshalb bringe ich hier nur ein kleines Beispiel:

Sei $p := 2 \cdot X^3 + X^2 + X + 1 \in \mathbb{R}[X]$. Dieses Polynom soll durch $q := X^2 - X + 1$ geteilt werden:

$$\begin{array}{r} (2 \cdot X^3 + X^2 + X + 1) / (X^2 - X + 1) = 2 \cdot X + 3 + r/q \\ - (2 \cdot X^3 - 2 \cdot X^2 + 2 \cdot X) \\ \hline 3 \cdot X^2 - X + 1 \qquad \qquad \qquad \text{mit } r := 2 \cdot X - 2 \\ - (3 \cdot X^2 - 3 \cdot X + 3) \\ \hline 2 \cdot X - 2 \end{array}$$

Es bleibt also ein Rest r . Wäre dieser gleich 0, dann könnte man den Faktor q komplett vom Polynom abspalten. Übrigens gilt die Beziehung $p = (X^2 - X + 1) \cdot (2 \cdot X + 3) + 2 \cdot X - 2$ allgemein für das Polynom; d.h. auch dann, wenn für X eine Matrix mit Koeffizienten in \mathbb{R} eingesetzt wird. Solche Ergebnisse sind gerade der Inhalt der Theorie über formale Polynome.

Nach diesem Prinzip kann man auf Polynome über Körpern auch den euklidischen Algorithmus anwenden und den ggT bestimmen. Der ggT ist (analog zu \mathbb{N}) das größte Polynom, welches gleichzeitig Teiler von zwei bestimmten Polynomen ist. Die Größe wird hierbei am Grad gemessen. Außerdem ist es wichtig zu wissen, dass der ggT von zwei Polynomen nicht eindeutig bestimmt ist. Man kann einen ggT immer mit einer Einheit, d.h. einem invertierbaren Element, multiplizieren. Einheiten im Polynomring sind gerade die konstanten Polynome außer dem Nullpolynom, denn die Multiplikation mit einem solchen Polynom lässt sich gerade rückgängig machen.

Beispiel: Es soll der (bzw. ein) ggT von $p := X^4 + X^3 + X^2 + 2 \cdot X + 3$ und $q := X^3 - X^2 + 2$ aus $\mathbb{R}[X]$ bestimmt werden. Offensichtlich ist p das größere Polynom, d.h. man muss zuerst p durch q teilen (mit Polynomdivision). Man erhält $p/q = (X+2) + (X^2-1)/q$, d.h. $p = (X+2) \cdot q + (X^2-1)$. Jetzt geht es weiter wie beim euklidischen Algorithmus für natürliche Zahlen:

$$\begin{aligned} X^4 + X^3 + X^2 + 2 \cdot X + 3 &= (X + 2) \cdot (X^3 - X^2 + 2) + (X^2 - 1) \\ X^3 - X^2 + 2 &= (X - 1) \cdot (X^2 - 1) + (X + 1) \\ X^2 - 1 &= (X - 1) \cdot (X + 1) \end{aligned}$$

Da im letztes Schritt die Division von $X^2 - 1$ durch $X + 1$ ohne Rest aufging, ist offenbar $X + 1$ ein Teiler sowohl von p als auch von q , und zwar ein ggT.

3.3 Lineare Gleichungssysteme

3.3.1 Definition

Ein lineares Gleichungssystem (LGS) ist eine Menge von Variablen (die Elemente eines Körpers K sind) und Gleichungen, die nur aus Summen dieser Variablen und Konstanten bestehen (wobei die Variablen noch mit Körperelementen, so genannten Koeffizienten, multipliziert werden dürfen). Diese Gleichungen sollen alle gleichzeitig erfüllt werden. Bringt man alle Konstanten auf die rechte Seite und alle Variablen auf die linke, und fügt man Nullen als Koeffizienten ein, so dass jede Variable in jeder Gleichung vorkommt, dann erhält man immer die folgende Form für ein LGS:

$$\begin{aligned}a_{11} \cdot x_1 + a_{12} \cdot x_2 + \cdots + a_{1l} \cdot x_l &= b_1 \\a_{21} \cdot x_1 + a_{22} \cdot x_2 + \cdots + a_{2l} \cdot x_l &= b_2 \\&\vdots \\a_{k1} \cdot x_1 + a_{k2} \cdot x_2 + \cdots + a_{kl} \cdot x_l &= b_k\end{aligned}$$

Dabei sind x_j die Variablen, a_{ij} die Koeffizienten, und b_i die Konstanten.

Schaut man sich noch einmal genau den Abschnitt über Matrizenmultiplikation an (3.2.4), dann fällt auf, dass ein solches LGS sehr viel mit Matrizen gemeinsam hat, besonders mit der Multiplikation einer Matrix mit einer Spaltenmatrix. In der Tat: Man kann leicht aus den k Gleichungen eine einzige Gleichung machen, indem man zwei Spaltenmatrizen gleichsetzt:

$$\begin{pmatrix} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \cdots + a_{1l} \cdot x_l \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + \cdots + a_{2l} \cdot x_l \\ \vdots \\ a_{k1} \cdot x_1 + a_{k2} \cdot x_2 + \cdots + a_{kl} \cdot x_l \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

Jetzt ist aber die linke Seite gerade das Matrizenprodukt von $((a_{ij})) \in K^{k \times l}$ mit $((x_j)) \in K^{l \times 1}$, bzw. ausgeschrieben:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kl} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

Man kann also jedes LGS als eine Matrixgleichung $A \cdot x = b$ schreiben, wobei $A \in K^{k \times l}$, $x \in K^{l \times 1}$ und $b \in K^{k \times 1}$ ist.

3.3.2 Lösungsmethoden

Diese Schreibweise legt in einem Spezialfall eine Lösungsmethode nahe, die man nicht erwarten würde, wenn man nur die Gleichungen betrachtet. Und zwar gibt es ja invertierbare Matrizen, und falls A invertierbar ist dann ist die Gleichung $A \cdot x = b$ äquivalent zu $x = A^{-1} \cdot b$. Hat man A^{-1} einmal ausgerechnet, kann man damit schnell für verschiedene b eine Lösung für x bestimmen. Leider kann man A^{-1} im Normalfall nicht direkt aus A ablesen. Es gibt aber durchaus Fälle, in denen das geht.

Ansonsten gibt es ja gewisse Umformungen, mit denen die Lösungsmenge eines LGS nicht verändert wird. Dazu zählen:

1. Umsortieren der Gleichungen

2. Umsortieren der Variablen
3. Multiplikation einer Gleichung mit einer Konstanten $\neq 0$
4. Addition einer Gleichung zu einer anderen

Diese Umformungen lassen sich natürlich direkt auf die Matrizen übertragen:

1. Das Umsortieren von Gleichungen entspricht dem (gleichzeitigen) Umsortieren der Zeilen von A und b .
2. Da das Umsortieren von Variablen erlaubt ist, kann man in A also auch die Spalten umsortieren, wenn man die Zeilen von x mit sortiert. Also ist z.B.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \\ a_{k1} & a_{k2} & \cdots \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

äquivalent zu

$$\begin{pmatrix} a_{12} & a_{11} & \cdots \\ a_{22} & a_{21} & \cdots \\ \vdots & \vdots & \\ a_{k2} & a_{k1} & \cdots \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ x_1 \\ \vdots \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

(Vertauschung der ersten beiden Spalten).

3. Man darf Zeilen von A und b gleichzeitig mit einer Konstanten multiplizieren, die nicht 0 ist (was leider oft vergessen wird, wenn die Konstante von einem Parameter abhängt).
4. Man darf eine Zeile von A und b zu einer anderen addieren. Dadurch, dass man sie vorher mit einer Konstanten multiplizieren darf, kann man sogar ein beliebiges Vielfaches der Zeile benutzen, d.h. man darf insbesondere auch subtrahieren statt addieren.

Da fast alle dieser Umformungen den Term x in der Gleichung $A \cdot x = b$ fest lassen, bietet es sich an, beim Durchführen der Umformungen das LGS nur als $(A|b)$ zu schreiben. Dann muss man nicht aufpassen, dass man die Umformungen gleichzeitig bei A und b machen muss.

Mit diesen Umformungen kann man nun das LGS in eine Form bringen, bei der man alle Lösungen für x (d.h. die Lösungsmenge) direkt ablesen kann:

- Unterhalb der Diagonalen stehen nur Nullen:

$$\left(\begin{array}{cccc|c} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \end{array} \right)$$

- Es gibt keine Stufe, die höher als eine Zeile ist:

$$\left(\begin{array}{cccc|c} * & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} * & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{array} \right)$$

- Am Anfang jeder Stufe steht eine 1:

$$\left(\begin{array}{cccc|c} * & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \end{array} \right)$$

- Über jeder 1 stehen nur Nullen:

$$\left(\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & * & 0 & 0 & * \\ 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 1 & * \end{array} \right)$$

Dabei erhält man die Nullen durch geschickte Vertauschungen und indem man andere Zeilen geeignet multipliziert und addiert/subtrahiert. Die Einsen bekommt man natürlich durch Multiplikation/Division der entsprechenden Zeilen.

Jetzt kann man die Lösungsmenge ablesen:

- Gibt es eine Zeile, in der links nur Nullen stehen, aber rechts $b_i \neq 0$, dann hat das LGS keine Lösung. Denn die entsprechende Gleichung ist $0 = b_i$.
- Für jede Zeile, in der links nur eine 1 steht, kann man rechts den Wert der entsprechenden Variable x_j direkt ablesen. Denn die Gleichung lautet $1 \cdot x_j = b_i$.
- In den anderen Zeilen gibt es Wahlmöglichkeiten. Am besten setzt man die Variablen, die den „*“-Spalten entsprechen, beliebig, und löst die Gleichung nach der Variable auf, die der 1 vorne entspricht. Zum Beispiel könnte man im LGS

$$\left(\begin{array}{cccc|c} 0 & 1 & 0 & a_{14} & b_1 \\ 0 & 0 & 1 & a_{24} & b_2 \end{array} \right)$$

die Variable x_4 frei wählen, und bekommt dann aus $1 \cdot x_3 + a_{24} \cdot x_4 = b_2$ die Formel $x_3 = b_2 - a_{24} \cdot x_4$, sowie aus $1 \cdot x_2 + a_{14} \cdot x_4 = b_1$ die Formel $x_2 = b_1 - a_{14} \cdot x_4$.

- Die Variablen, die überhaupt nicht aufgetaucht sind, sind beliebig.

Wie gibt man also die Lösungsmenge $L := \{x \in K^{l \times 1} : A \cdot x = b\}$ an? Man schreibt x einfach wieder als

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_l \end{pmatrix}$$

und setzt für jedes x_j , das nicht beliebig ist, die nach x_j aufgelöste Gleichung ein. Die x_j , die beliebig sind, müssen auch so in der Menge gekennzeichnet werden. Dann kann man diese eine Spaltenmatrix noch als Summe von Spaltenmatrizen schreiben, die eventuell mit einem der x_j multipliziert werden, so dass die Matrizen selbst konstant sind.

An einem Beispiel wird das vielleicht deutlicher:

Das folgende reelle LGS soll gelöst werden:

$$\begin{aligned} 2 \cdot x_1 + 1 \cdot x_2 &= 3 \\ -2 \cdot x_1 - 4 \cdot x_2 + 2 \cdot x_3 &= 2 \end{aligned}$$

Dies bringt man zunächst auf die Form:

$$\begin{pmatrix} 2 & 1 & 0 \\ -2 & -4 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

In der Matrix gibt es bereits eine 0, aber an der falschen Stelle. Am besten wäre es, wenn sie in der ersten Spalte stehen würde, und dahinter eine 1. Also vertauscht man die erste und dritte Spalte, d.h. x_1 und x_3 :

$$\begin{pmatrix} 0 & 1 & 2 \\ 2 & -4 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Nun kann man in der abgekürzten Schreibweise weitermachen:

$$\left(\begin{array}{ccc|c} 0 & 1 & 2 & 3 \\ 2 & -4 & -2 & 2 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 2 & -4 & -2 & 2 \\ 0 & 1 & 2 & 3 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 1 & 2 & 3 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 3 \end{array} \right)$$

Im letzten Schritt wurde die zweite Zeile zwei mal zur ersten addiert, um die 0 über der 1 zu bekommen. Man kann x_1 (jetzt die dritte Variable) beliebig wählen und bekommt dann $x_2 = 3 - 2 \cdot x_1$ und $x_3 = 7 - 3 \cdot x_1$. Die Lösungsmenge L kann man also schreiben als:

$$L = \left\{ \begin{pmatrix} x_1 \\ 3 - 2 \cdot x_1 \\ 7 - 3 \cdot x_1 \end{pmatrix} : x_1 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 0 \\ 3 \\ 7 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix} \cdot x_1 : x_1 \in \mathbb{R} \right\}$$

Um eine solche Darstellung der Lösungsmenge zu erhalten, kann man zum Schluss auch den sogenannten „-1-Trick“ benutzen: Man macht durch Streichen und Einfügen von Nullzeilen die linke Hälfte der Matrix zu einer quadratischen Matrix, bei der die erste 1 in jeder Zeile auf der Diagonalen steht (außer bei den Nullzeilen natürlich). Dann schreibt man in den Nullzeilen -1 in die Diagonale. Die Spalten, in denen dies geschieht, sind gerade die Matrizen, die in der Lösungsmenge mit einem beliebigem Körperelement multipliziert werden, und rechts steht der absolute Term.

Einfügen der Nullzeile:

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 3 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Ersetzen von 0 durch -1 auf der Diagonalen:

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{array} \right), \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}$$

Jetzt müssen wir uns noch daran erinnern, dass wir die Variablen x_1 und x_3 vertauscht hatten, und erhalten:

$$L = \left\{ \begin{pmatrix} 0 \\ 3 \\ 7 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \cdot r : r \in \mathbb{R} \right\}$$

3.3.3 Weiterführende Theorien

Muss man das LGS $A \cdot x = b$ für ein A und verschiedene b lösen, dann hilft ein Satz, der besagt, dass man die allgemeine Lösung für $A \cdot x = b$ (d.h. die Lösungsmenge) erhält, indem man zuerst die allgemeine Lösung für $A \cdot x = 0$ bestimmt und eine spezielle Lösung für $A \cdot x = b$ addiert. Selbst wenn nur ein b im Spiel ist, kann dies einfacher sein, wenn man schon eine spezielle Lösung kennt.

In obigem Beispiel ist also

$$\left\{ \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \cdot r : r \in \mathbb{R} \right\}$$

die Lösungsmenge von $A \cdot x = 0$, und

$$\begin{pmatrix} 0 \\ 3 \\ 7 \end{pmatrix}$$

eine spezielle Lösung von $A \cdot x = b$. Addition ergibt das bekannte Ergebnis.

Der vorgestellte Algorithmus (der Gaußalgorithmus) manipuliert offenbar die Matrizen A und b , ändert aber an der Lösungsmenge für $A \cdot x = b$ nichts. Es liegt also nahe, dass sich alle erlaubten Operationen durch die Multiplikation mit invertierbaren Matrizen C_i ausdrücken lassen, denn $C_i \cdot A \cdot x = C_i \cdot b$ ist eine dazu äquivalente Gleichung. In der Tat kann man diese Matrizen konkret angeben. Dies liefert ein Verfahren, um für eine quadratische Matrix $A \in K^{n \times n}$ die Inverse zu bestimmen, falls es sie gibt:

Ist nämlich A invertierbar, dann hat das LGS $A \cdot x = 0$ nur die triviale Lösung $x = 0$, denn man kann die Gleichung von links mit A^{-1} multiplizieren. Das heißt für die Matrix $C \cdot A$, die nach Anwenden des Gaußalgorithmus entsteht, dass

$$C \cdot A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = E_n$$

ist, also $C = A^{-1}$ (nach Multiplikation der Gleichung von rechts mit A^{-1}). Man muss also nur C finden, um A^{-1} zu bestimmen. Dazu wendet man einfach die gleichen Operationen auf die Einheitsmatrix E_n an, und erhält damit die Matrix $C \cdot E_n = C = A^{-1}$.

Konkret sieht das so aus, dass man $(A|E_n)$ betrachtet und entsprechend der Regeln umformt, so dass man am Schluss auf der linken Seite die Einheitsmatrix bekommt. Dann ist die rechte Seite A^{-1} .

Vielleicht ist es aufgefallen, dass man die Form der Lösungsmenge schon sehen kann, wenn man den Gaußalgorithmus fertig gerechnet hat. Und zwar ist die Anzahl der Variablen, die nicht beliebig sind, gerade die Anzahl der Zeilen, die am Schluss übrig bleiben. (Diese Zahl nennt man den „Rang“ der Matrix.) Die Anzahl der Variablen, die beliebig sind, und damit die Form der Lösungsmenge, ergibt sich daraus sofort.

Man kann sich leicht klar machen, dass der Rang von A gleich dem Rang von A^T ist. D.h. um den Rang zu bestimmen, darf man statt Zeilenumformungen auch Spaltenumformungen durchführen. Dies liefert bei einigen Matrizen eine einfachere Aussage darüber, wie die Lösungsmenge aussieht.

Z.B. kann man dem reellen LGS

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \\ 3 & 2 & 6 \\ 4 & 3 & 8 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

direkt ansehen, dass die Lösungsmenge die Form

$$\left\{ \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \cdot r : r \in \mathbb{R} \right\}$$

hat. Denn zieht man die erste Spalte zweimal von der dritten ab, dann bekommt man dort eine Nullspalte. Die zweite Spalte hat mit der einen 0 oben bereits die Form, die nach Anwendung des Gaußalgorithmus (mit Spalten statt Zeilen) entsteht. Also ist der Rang 2, und $3 - 2 = 1$ Variable ist beliebig.

Hier hat das sogar direkt einen praktischen Nutzen. Denn man sieht, dass $x_1 = 2, x_2 = 0, x_3 = -1$ eine Lösung ist. (Wenn nicht, dann bitte einfach mal diese Werte einsetzen und die Matrizenmultiplikation durchführen.) Also gibt es nur eine Möglichkeit für die Lösungsmenge:

$$\left\{ \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \cdot r : r \in \mathbb{R} \right\}$$

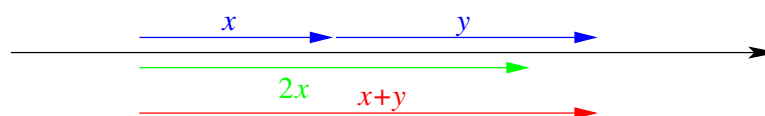
4 Vektorräume

4.1 Einführung

4.1.1 Ursprung

Der abstrakte Begriff des Vektorraums orientiert sich stark an den Vektoren, die man in ein reelles Koordinatensystem einzeichnen kann. Ein Vektor ist dabei ein Pfeil, den man beliebig verschieben kann. Die Addition von zwei Vektoren ist dadurch erklärt, dass man die Pfeile aneinander hängt; damit bekommt man einen neuen Vektor. Die Multiplikation mit einer reellen Zahl ändert die Länge des Pfeils.

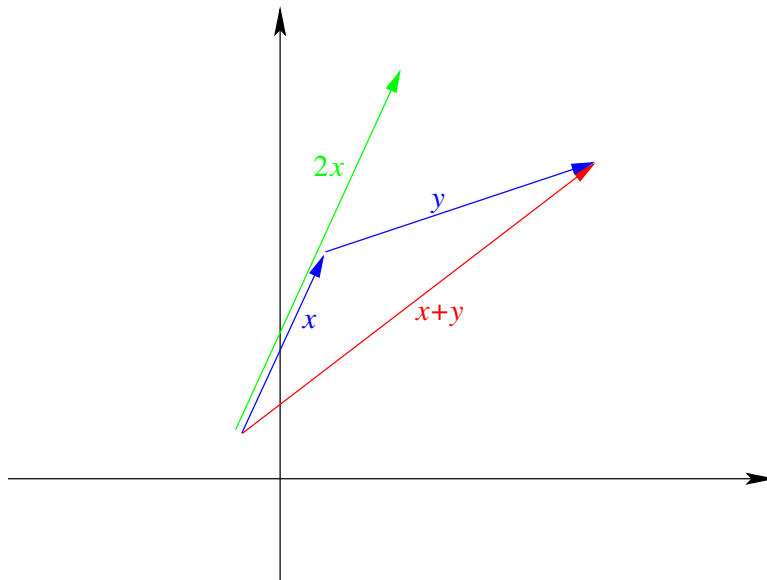
Bekanntermaßen lässt sich $\mathbb{R}^1 = \mathbb{R}$ durch eine Zahlengerade darstellen. Die Addition und Multiplikation von Vektoren entspricht dann genau der von reellen Zahlen:



Wichtig bei der Betrachtung einer solchen Darstellung ist, dass zwei Vektoren gleich sind, wenn die Pfeile gleiche Länge und Richtung haben. Es ist unerheblich, wo man den Pfeil einzeichnet. Möchte man die wirklichen reellen Zahlen als Vektoren betrachten, dann zeichnet man die Pfeile so ein, dass

sie im Nullpunkt beginnen. Die Pfeilspitze liegt dann auf der Zahlengeraden bei der Zahl, die man durch den Pfeil ausdrücken will.

Die übliche Darstellung von \mathbb{R}^2 ist ein zweidimensionales Koordinatensystem. Darin sieht Addition und Multiplikation entsprechend so aus:



Betrachtet man einen Vektor als ein Tupel (x_1, x_2) , dann funktionieren Addition und Multiplikation komponentenweise. Es ist wahrscheinlich aus der Schule bekannt, dass man auch die Punkte im Koordinatensystem als solche Tupel darstellt. Sie sind gerade die Pfeilspitzen, wenn man die Pfeile im Ursprung ansetzt. Man nennt die Vektoren dann „Ortsvektoren“.

Mathematiker möchten aber gerne von solchen konkreten Anschauungsobjekten in eine abstrakte Definition übergehen, die man auch für andere Zwecke gebrauchen kann. Dabei kann man erst einmal nüchtern feststellen, dass wir es offensichtlich mit einer Addition von zwei Vektoren und einer Multiplikation von einem Vektor mit einem so genannten „Skalar“ zu tun haben. Dann sucht man bestimmte Eigenschaften heraus, die für die Arbeit mit solchen Pfeil-Diagrammen wesentlich sind.

Z.B. soll auf jeden Fall die Multiplikation mit 1 den Vektor nicht verändern. Die Multiplikation mit 0 dagegen sollte ihn auf einen neutralen Vektor schrumpfen. Außerdem sollte die Multiplikation ungefähr so funktionieren, wie man sich eine Multiplikation vorstellt, d.h. z.B. $2 \cdot x = x + x$. Dies kann man allgemeiner in einem Distributivgesetz zusammenfassen. Insgesamt erhält man einige Gesetze, wobei sie sich teilweise auseinander ableiten lassen:

4.1.2 Axiome

Ein Vektorraum ist eine Menge M mit zwei Verknüpfungen $+$: $M \times M \rightarrow M$ und \cdot : $K \times M \rightarrow M$, wobei K ein Körper ist, so dass für alle $x, y \in M$ und $a, b \in K$ die folgenden Gesetze gelten:

- $(M, +)$ ist eine abelsche Gruppe.
- $a \cdot (x + y) = a \cdot x + a \cdot y$ (dies ist die Verknüpfung „+“ des Vektorraums)
- $(a + b) \cdot x = a \cdot x + b \cdot x$ (dies ist die Verknüpfung „+“ des Körpers)

- $a \cdot (b \cdot x) = (a \cdot b) \cdot x$ (dies ist zweimal die Verknüpfung „ \cdot “ des Vektorraums, einmal die des Körpers)
- $1 \cdot x = x$
- $0 \cdot x = 0$ (das neutrale Element von $(M, +)$)
- $a \cdot 0 = 0$
- $(-1) \cdot x = -x$ (hierbei ist $-x$ das Inverse zu x in $(M, +)$)

Es stellt sich heraus, dass die ersten fünf Gesetze ausreichen und sich die anderen daraus ableiten lassen. Es ist aber wichtig, sich die Bedeutung aller dieser Gesetze anhand des Spezialfalls der Pfeile im Koordinatensystem klar zu machen.

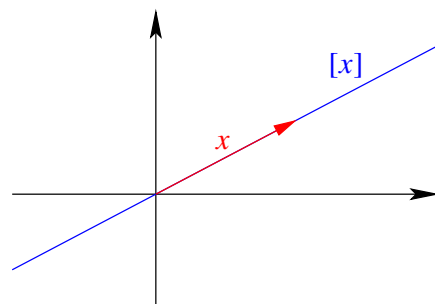
Anhand der benutzten Verknüpfungen sieht man, dass zumindest die Einführung von Ringen wesentlich für die Vektorraumtheorie war. Für grundlegende Untersuchungen über Vektorräume braucht man sehr bald Eigenschaften von Körpern; deshalb werden Vektorräume grundsätzlich nur über Körpern definiert. Dies ist immer noch eine sehr allgemeine Definition; es gibt viele verschiedene Beispiele für Vektorräume, die nichts mit Pfeildiagrammen zu tun haben (außer dass beide eben Vektorräume bilden).

Trotzdem ist es sinnvoll, sich zu fragen, wie einschränkend diese Definition bereits ist. Dazu kann man z.B. betrachten, welche Unterräume es gibt (d.h. Teilmengen von M , die mit den darauf eingeschränkten Verknüpfungen selbst einen Vektorraum bilden). Das Ergebnis ist, dass man in vielen Fällen alle Unterräume kennt:

4.1.3 Erzeugnis, Linearkombinationen

Das Erzeugnis $[M]$ einer Menge M von Vektoren ist analog zum Erzeugnis von Gruppenelementen definiert (siehe 3.1.2): Man lässt die Elemente „arbeiten“, um einen vollständigen Vektorraum zu bilden. Übrigens schreibt man für $\{x_1, x_2, \dots, x_n\}$ auch $[x_1, x_2, \dots, x_n]$.

Betrachten wir zunächst die Multiplikation. Laut Definition muss für jedes $x \in M$ und $a \in K$ das Produkt $a \cdot x$ in dem Vektorraum liegen. Das Erzeugnis $[M]$ muss also alle Vielfachen von Vektoren aus M enthalten. Geometrisch ist z.B. das Erzeugnis eines Ortsvektors $x \neq 0$ eine Gerade durch den Ursprung:



Für die Verknüpfung „ $+$ “ gilt, dass $[M]$ das Gruppenerzeugnis der so gewonnenen Vektoren enthalten muss. Man bildet also erst alle Geraden durch den Ursprung und addiert dann jeweils die Ortsvektoren von Punkten auf den Geraden. Z.B. ist das Erzeugnis von zwei verschiedenen Geraden

eine Ebene, denn die Punkte, die durch Addition zweier Ortsvektoren von Geraden erreicht werden können, liegen gerade auf der einen Ebene, die beide Geraden enthält.

Insgesamt ist also $[x_1, x_2, \dots, x_n] = \left\{ \sum_{i=1}^n a_i \cdot x_i : a_1, \dots, a_n \in K \right\}$. Eine solche Summe nennt man eine „Linearkombination“ der Vektoren x_1 bis x_n .

Über das Erzeugnis von Mengen kann man alle Untervektorräume eines Vektorraums angeben, denn für einen Unterraum U eines Vektorraums V ist natürlich immer $[U] = U$. Das Interessante ist, dass sie viele Unterräume durch einige wenige Vektoren erzeugen lassen.

4.1.4 Lineare (Un-)Abhängigkeit

Die Frage nach der linearen Abhängigkeit einer Menge M ist letztlich die Frage, ob es beim Bilden des Erzeugnis $[M]$ Vektoren in M gibt, die dafür keine Rolle spielen. Einige einfache Fälle sind:

- Einer der Vektoren aus M ist der Nullvektor. Da die Addition mit dem Nullvektor keine neuen Vektoren erzeugt, ist M immer linear abhängig, falls $0 \in M$ ist.
- Zwei Vektoren sind Vielfache voneinander. Da die Vektoren beim Bilden des Erzeugnis mit beliebigen Skalaren multipliziert werden können (und die zugrunde liegende Menge ein Körper ist, also immer Inverse besitzt), erhält man immer noch das gleiche Erzeugnis, wenn man einen der beiden Vektoren weglässt.
- Einer der Vektoren ist die Summe oder Differenz von zwei anderen Vektoren. Dann liegt dieser Vektor bereits im Gruppenerzeugnis dieser beiden anderen. Wenn man ihn weglässt, ändert sich das Erzeugnis ebenfalls nicht.

Allgemein ist M genau dann linear abhängig, wenn es einen Vektor $x \in M$ gibt, der in $[M \setminus \{x\}]$ liegt, sich also mit den anderen Vektoren aus M erzeugen lässt. D.h. er lässt sich als Linearkombination $\sum_{i=1}^n a_i \cdot x_i$ mit Vektoren $x_i \in M, x_i \neq x$ schreiben.

Daraus lässt sich ein etwas einfacheres Kriterium ableiten, damit man nicht jeden einzelnen Vektor aus M überprüfen muss. Und zwar ist M genau dann linear abhängig, wenn man den Nullvektor auf eine nichttriviale Art als Linearkombination schreiben kann, d.h. als Summe $\sum_{i=1}^n a_i \cdot x_i$ mit $x_i \in M$, so dass die a_i nicht alle 0 sind. Denn lässt sich o.B.d.A. der Vektor x_1 als Linearkombination $x_1 = \sum_{i=2}^n a_i \cdot x_i$ darstellen, dann ist mit $a_1 := -1$ die Gleichung $\sum_{i=1}^n a_i \cdot x_i = 0$ erfüllt. Ist umgekehrt

o.B.d.A. $a_1 \neq 0$, dann ist die Gleichung $\sum_{i=1}^n a_i \cdot x_i = 0$ äquivalent zu $x_1 = -\frac{\sum_{i=2}^n a_i \cdot x_i}{a_1}$.

Um dieses Kriterium möglichst einfach überprüfen zu können, sollte man sich noch einmal die Abschnitte über Matrizen (3.2.4) und lineare Gleichungssysteme (3.3) anschauen. Dann sieht man eventuell, dass man die Summe $\sum_{i=1}^n a_i \cdot x_i$ als Produkt von zwei „Matrizen“ schreiben kann:

$$\sum_{i=1}^n a_i \cdot x_i = (x_1 \quad x_2 \quad \cdots \quad x_n) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Dabei darf man nicht vergessen, dass die Vektoren x_i oft Tupel sind, die Pfeile in einem Koordinatensystem darstellen; und es wird bald auch noch eine Möglichkeit eingeführt, praktisch *jeden* Vektor als Tupel darstellen zu können. Es erweist sich als sinnvoll, diese Tupel aus K^m immer als Spaltenmatrizen aus $K^{m \times 1}$ zu schreiben. Sei also $x_i = (x_{1i}, x_{2i}, \dots, x_{mi})$, dann wird die Gleichung $\sum_{i=1}^n a_i \cdot x_i = 0$ zu einer Matrixgleichung:

$$\left(\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{m2} \end{pmatrix} \cdots \begin{pmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{mn} \end{pmatrix} \right) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Diese Gleichung ist ein lineares Gleichungssystem für die Skalare a_i , welches mit den üblichen Mitteln gelöst werden kann. M ist genau dann linear abhängig, wenn dieses LGS eine nichttriviale Lösung besitzt. Ist $n = m$, dann ist dies genau dann der Fall, wenn die Matrix regulär (invertierbar) ist. Ist $n > m$, dann besitzt das LGS immer eine nichttriviale Lösung, d.h. die Vektoren sind dann immer linear abhängig.

Z.B. sind 3 Vektoren im \mathbb{R}^2 immer linear abhängig. Sie können ja nicht mehr als die Ebene \mathbb{R}^2 selbst erzeugen, aber dazu braucht man nur 2 Vektoren.

4.1.5 Basen und Dimension

Eine Basis eines Vektorraums ist eine Teilmenge, die den ganzen Vektorraum erzeugt (ein „Erzeugendensystem“), und in der es keine überflüssigen Vektoren gibt. D.h. nach dem vorherigen Abschnitt genau, dass die Teilmenge linear unabhängig ist. Das ist äquivalent dazu, dass die Teilmenge maximal linear unabhängig ist, d.h. fügt man nur einen einzigen Vektor hinzu, dann wird die Menge linear abhängig.

Mit Hilfe einer Basis lässt sich jeder Vektor auf eindeutige Art als Linearkombination der Basisvektoren schreiben. Damit kann man später z.B. Abbildungen so definieren, dass man die Bilder der Basisvektoren festlegt und das Bild eines beliebigen anderen Vektors über seine Linearkombination errechnet.

Ein wichtiger Satz der Vektorraumtheorie ist, dass jede Basis eines Vektorraums gleich viele Elemente besitzt (vorausgesetzt, es gibt überhaupt eine Basis mit endlich vielen Elementen). Die Anzahl der Elemente nennt man die „Dimension“ des Vektorraums.

Z.B. ist $\{(1, 0, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)\}$ eine Basis von K^n mit genau n Elementen. Also ist $\dim K^n = n$.

Die Polynome über K vom Grad $\leq n$ bilden einen Vektorraum P , der $\{1, X, X^2, \dots, X^n\}$ als Basis hat. Also ist $\dim P = n + 1$.

Um aus einer beliebigen endlichen Menge M von Vektoren eine Basis des Vektorraums zu machen, muss man dafür sorgen, dass sie die beiden Bedingungen erfüllt:

1. M muss linear unabhängig werden. D.h. man muss für jeden Vektor prüfen, ob er sich als Linearkombination der anderen darstellen lässt, und ihn dann entfernen. Manchmal kann man dies für einzelne Vektoren schon sehen, oder man sieht statt dessen, dass M linear unabhängig ist. Im Allgemeinen gibt es zwei Verfahren, mit denen man auf einmal eine linear unabhängige Teilmenge von M bilden kann, die den gleichen Untervektorraum $[M]$ erzeugt:

- (a) Ersetzt man einen Vektor aus M durch ein Vielfaches davon oder addiert man einen Vektor zu einem anderen, dann ändert sich das Erzeugnis $[M]$ nicht. Diese Operationen reichen aus, um mit den Vektoren den Gaußalgorithmus durchzuführen. Da Vektoren üblicherweise als Spalten geschrieben werden, muss man allerdings aufpassen, denn die Operationen sind keine Zeilen-, sondern *Spaltenoperationen*.
 - (b) Schreibt man die Vektoren als Spalten in eine Matrix (wie man allgemeine Vektoren als Spalten schreibt, wird später noch erklärt), dann kann man auf dieser Matrix den Gaußalgorithmus mit *Zeilenoperationen* durchführen. In der Treppenform der Matrix schaut man sich die Spalten an, bei denen eine neue Treppenstufe beginnt. Die Vektoren, die vorher in diesen Spalten gestanden haben, sind linear unabhängig und erzeugen $[M]$. Diese Methode ist zwar nicht sofort ersichtlich, aber sie ist vorteilhaft, wenn man ohnehin die Matrix in Treppenform braucht. Man muss aber aufpassen, dass man die Methoden nicht miteinander vermischt.
2. Jetzt hat man eine Basis von $[M]$. Nach dem Basisergänzungssatz kann man sie zu einer Basis des ganzen Vektorraums ergänzen, indem man genügend linear unabhängige Vektoren hinzufügt. Aber wie findet man solche Vektoren? Das hängt davon ab, welches der beiden Verfahren man gewählt hat:
- (a) Hier ist es relativ einfach, denn die Vektoren, die man erhalten hat, bilden eine transponierte Treppenmatrix. Sie kann durch Vektoren, die nur an einer Stelle eine 1 haben, zu einer vollständigen Diagonalmatrix ergänzt werden.
 - (b) Bei der zweiten Methode kann man nicht direkt Vektoren finden, die die Menge zu einer Basis ergänzen. Man muss statt dessen die Treppenmatrix zu einer Diagonalmatrix machen und rückwärts verfolgen, wie sich die jeweiligen Spaltenvektoren dadurch ändern. Allerdings muss man aufpassen, denn es kann sich auf den gesamten Spaltenvektor auswirken.

Eigentlich ist es sogar gar nicht schwer, eine linear unabhängige Menge zu einer Basis zu ergänzen. Denn an Vektoren, die mit der Menge linear unabhängig sind, mangelt es nie. Man sieht es schon im \mathbb{R}^2 : Hat man einen beliebigen Vektor, dann ist ein zweiter Vektor nur dann linear abhängig mit dem ersten, wenn er ein Vielfaches von ihm ist, d.h. in die gleiche oder entgegengesetzte Richtung zeigt. Für die meisten Vektoren ist dies natürlich nicht der Fall. Also kann man die Vektoren auch raten, muss dann allerdings die lineare Unabhängigkeit noch nachweisen (z.B. mit einem LGS).

4.2 Lineare Abbildungen

4.2.1 Definition

Für Vektorräume führt man wie für Gruppen Abbildungen ein, die „strukturerhaltend“ sind, und nennt sie „Homomorphismen“ oder hier auch „lineare Abbildungen“ (siehe 3.1.3). Wenn sie die Vektorraumstruktur erhalten sollen, dann müssen sie zumindest Gruppenhomomorphismen bezüglich der Verknüpfung „+“ des Vektorraums sein. Aber um sich genau zu merken, wie die Definition eines Vektorraumhomomorphismus aussehen muss, sollte man sich überlegen, was Homomorphismen im Allgemeinen sind. Wenn man weiß, was ein Homomorphismus eigentlich ist, dann ist die genaue Definition Nebensache.

Dazu ist es am sinnvollsten, sich zunächst über den Spezialfall der Isomorphismen (der bijektiven Homomorphismen) Gedanken zu machen. Es wurde bereits erwähnt, dass man zwei Gruppen (bzw.

Vektorräume), zwischen denen ein Isomorphismus existiert, als „isomorph“ bezeichnet. Offensichtlich ist dies also eine wichtige Eigenschaft. Sie bedeutet, dass jedes Element der einen Gruppe eine genaue Entsprechung in der anderen Gruppe besitzt. D.h. es handelt sich eigentlich um die gleichen Gruppen, mit dem kleinen Unterschied, dass die Elemente andere Namen haben.

Z.B. sind die Gruppen $G = (\{a, b\}, \circ)$ und $H = (\{c, d\}, \star)$ mit den folgenden Verknüpfungstabellen isomorph:

\circ	a	b
a	a	b
b	b	a

\star	c	d
c	c	d
d	d	c

H entsteht aus G durch Umbenennen der Elemente. Der (einzige) Isomorphismus von G nach H bildet a auf c und b auf d ab.

Diese Eigenschaft, dass die beiden Gruppen/Ringe/Vektorräume gleich sind bis auf die Namen der Elemente, könnte man auch als Definition für die Isomorphie benutzen. Aber was bedeutet das genau? Es muss ja irgendeine Zuordnung zwischen den Elementen der beiden Gruppen geben, nennen wir sie f . f muss bijektiv sein, denn die Umbenennung muss in beiden Richtungen funktionieren. Und die Gruppen sind genau dann gleich, wenn alle Rechnungen das gleiche Ergebnis liefern, denn die Gruppen sind durch die Gruppenverknüpfung bereits vollständig charakterisiert. Es muss also z.B. egal sein, ob man zwei Elemente in der einen Gruppe verknüpft oder ob man sie erst umbenennt, in der anderen Gruppe verknüpft, und dann die Umbenennung rückwärts durchführt. Insgesamt bedeutet das, dass das folgende Diagramm kommutiert:

$$\begin{array}{ccc} G & \xrightarrow{\circ} & G \\ f \uparrow & \text{//} & \downarrow f \\ H & \xrightarrow{\star} & H \end{array}$$

(Dass ein Diagramm kommutiert, bedeutet, dass man im Diagramm von einem Punkt zu einem anderen einen beliebigen Weg wählen kann und immer das gleiche Ergebnis erhält.)

Jetzt kann man sich überlegen, wie man die Bedingung so abschwächen kann, dass f nicht mehr bijektiv sein muss. Wenn f nicht mehr bijektiv ist, darf der Pfeil nur noch nach unten zeigen:

$$\begin{array}{ccc} G & \xrightarrow{\circ} & G \\ f \downarrow & \text{//} & \downarrow f \\ H & \xrightarrow{\star} & H \end{array}$$

In diesem Diagramm gibt es aber überhaupt nur zwei Wege, die gleichen Anfangs- und Endpunkt haben, nämlich die Hintereinanderausführung von \circ und f sowie f und \star . Also ist die Bedingung folgendermaßen: Bildet man für zwei Elemente $x, y \in G$ den Wert $x \circ y$ und setzt ihn in f ein, d.h. $f(x \circ y)$, dann muss das Ergebnis übereinstimmen, wenn man erst $f(x)$ und $f(y)$ bildet und dann $f(x) \star f(y)$. Also $f(x \circ y) = f(x) \star f(y)$ wie gehabt.

Bei Vektorräumen V und W haben wir jeweils zwei Verknüpfungen „+“ und „ \cdot “, die neue Elemente aus V bzw. W liefern. Ein entsprechendes Diagramm sieht also so aus:

$$\begin{array}{ccc} V & \xrightarrow{+} & V \\ \Phi \downarrow & \text{//} & \downarrow \Phi \\ W & \xrightarrow{\cdot} & W \end{array}$$

D.h. es gibt eigentlich zwei Bedingungen. Zum Einen muss $\Phi(x + y) = \Phi(x) + \Phi(y)$ für alle $x, y \in V$ gelten. Zum Anderen muss aber auch für die Verknüpfungen „ \cdot “ von V und W egal sein, ob man zuerst \cdot und dann Φ ausführt oder zuerst Φ und dann \cdot . V und W müssen also erst einmal Vektorräume über dem selben Körper K sein, und dann muss $\Phi(a \cdot x) = a \cdot \Phi(x)$ für alle $a \in K$ gelten.

Diese beiden Bedingungen kann man noch zu $\Phi(a \cdot x + y) = a \cdot \Phi(x) + \Phi(y)$ zusammenfassen, denn mit $a = 1$ bzw. $y = 0$ kann man die beiden einzelnen Bedingungen erzeugen.

Die Begriffe „Kern“ und „Bild“ werden auf die jeweiligen Gruppen mit „+“ bezogen. Es gilt auch hier, dass Kern und Bild Untervektorräume sind. D.h. sie haben z.B. eine bestimmte Dimension, die höchstens so groß wie die Dimension von V bzw. W sein kann.

4.2.2 Multiplikation mit Matrizen

Eine spezielle Art der linearen Abbildungen, die man zwischen den Standardräumen K^n und K^m definieren kann, ist die Multiplikation mit einer Matrix $A \in K^{m \times n}$, d.h. $\Phi : K^n \rightarrow K^m, x \mapsto A \cdot x$. Schreibt man die Multiplikation aus, kann man leicht nachrechnen, dass dies eine lineare Abbildung definiert. Was sind Kern und Bild?

Es gilt Kern $\Phi = \{x \in K^n : \Phi(x) = A \cdot x = 0\}$. Das ist die Lösungsmenge des homogenen LGS $A \cdot x = 0$, also mit den bekannten Mitteln zu berechnen (siehe 3.3.2). Die Menge, die man dabei erhält, ist bereits in einer Form, in der jedes Element als Linearkombination bestimmter Vektoren geschrieben ist. Diese Vektoren sind immer linear unabhängig; das sieht man, indem man sich überlegt, wie sie gebildet werden. Also bilden sie eine Basis des Kerns.

Das Bild ist sogar noch einfacher zu finden. Wenn x jeden Vektor aus K^n durchläuft, bekommt man nach der letzten Darstellung im Abschnitt über Matrizenmultiplikation (3.2.4) alle Vielfachen der Spaltenvektoren von A sowie Summen davon. Also bilden die Spalten von A ein Erzeugendensystem von Bild Φ , und eine Basis lässt sich leicht daraus bilden, indem man linear abhängige Vektoren entfernt. Hat man den Kern bereits ausgerechnet, dann bilden nach dem Abschnitt über das Finden von Basen (4.1.5) die Vektoren eine Basis des Bildes, bei denen in der Treppenform eine neue Stufe anfängt.

4.2.3 Definition über Basen

Die wichtigste Eigenschaft einer Basis $B = \{b_1, \dots, b_n\}$ eines Vektorraums V ist, dass man jeden Vektor $x \in V$ eindeutig als Linearkombination $x = \sum_{i=1}^n a_i \cdot b_i = a_1 \cdot b_1 + \dots + a_n \cdot b_n$ schreiben kann.

Mit dieser Eigenschaft kann man im Zusammenhang mit linearen Abbildungen sehr viel anfangen. Ist $\Phi : V \rightarrow W$ eine lineare Abbildung, und sind die Bilder $\Phi(b_i)$ für alle i festgelegt, dann ergibt sich für ein beliebiges $x \in V$ wie oben:

$$\Phi(x) = \Phi(a_1 \cdot b_1 + \dots + a_n \cdot b_n) = a_1 \cdot \Phi(b_1) + \dots + a_n \cdot \Phi(b_n)$$

Da alle $\Phi(b_i)$ festgelegt wurden, ist Φ damit vollständig definiert. Sie ist auch wohldefiniert, denn die Darstellung jedes Vektors als Linearkombination ist eindeutig.

Beispiel: Im \mathbb{R}^2 kann eine lineare Abbildung $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert werden durch $\Phi\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 3$ und $\Phi\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 5$. Um ein bestimmtes Bild auszurechnen, z.B. $\Phi\left(\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right)$, schreibt man $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ als

Linearkombination der Basisvektoren, d.h. als $2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und rechnet aus: $\Phi\left(\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right) = \Phi\left(2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 2 \cdot \Phi\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) - 3 \cdot \Phi\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 2 \cdot 3 - 3 \cdot 5 = -9$.

Allgemeiner ist $\Phi\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1 \cdot 3 + x_2 \cdot 5 = (3 \ 5) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. D.h. die Abbildung lässt sich ausdrücken durch $\Phi(x) = (3 \ 5) \cdot x$, also durch die Multiplikation mit einer Matrix. Dies lässt sich verallgemeinern.

Dazu muss allerdings erst einmal die Basis B eine Ordnung bekommen, aber Mengen sind nicht geordnet. Auch wenn Basen oft als Mengen aufgefasst werden, schlage ich vor, sich eine Basis eher als eine 1-Zeilen-Matrix $B = (b_1 \ b_2 \ \dots \ b_n)$ vorzustellen. Im dem Fall, dass die b_i Spaltenvektoren sind, wird daraus dann sogar eine richtige Matrix, deren Matrixeigenschaften später noch wichtig werden. Jedenfalls kann man dann für ein $x \in V$, das als $\sum_{i=1}^n a_i \cdot b_i$ dargestellt wurde, den „Koordinatenvektor“

$$D_B(x) := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

(bezüglich B) definieren. Dann gilt mit der üblichen Matrizenmultiplikation:

$$B \cdot D_B(x) = (b_1 \ \dots \ b_n) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = b_1 \cdot a_1 + \dots + b_n \cdot a_n = x$$

Das Schöne ist, dass diese Koordinatenvektoren Elemente aus K^n sind, für die man wie im Beispiel eine Matrix $D_{SB}(\Phi)$ konstruieren kann, die Φ definiert. (Das „S“ steht für die Standardbasis; es wird gleich verallgemeinert.) Und zwar gilt:

$$\begin{aligned} \Phi(x) &= \Phi(b_1 \cdot a_1 + \dots + b_n \cdot a_n) = \Phi(b_1) \cdot a_1 + \dots + \Phi(b_n) \cdot a_n = \\ &= (\Phi(b_1) \ \dots \ \Phi(b_n)) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \underbrace{(\Phi(b_1) \ \dots \ \Phi(b_n))}_{=: D_{SB}(\Phi)} \cdot D_B(x) \end{aligned}$$

Um dies im Klartext auszudrücken: Man nimmt die Bilder $\Phi(b_i)$ der Basisvektoren und schreibt sie als Spalten in die Matrix $D_{SB}(\Phi)$. Dies sollte man sich auf jeden Fall merken, weil das Aufstellen einer solchen Matrix zu den Standardaufgaben in der Linearen Algebra gehört.

Oft möchte man aber das Ergebnis als Koordinatenvektor bezüglich einer geeigneten Basis C von W darstellen, das heißt man sucht $D_{CB}(\Phi)$ so, dass $D_C(\Phi(x)) = D_{CB}(\Phi) \cdot D_B(x)$ ist. Das ist analog genau dann der Fall, wenn $D_{CB}(\Phi) = (D_C(\Phi(b_1)) \ \dots \ D_C(\Phi(b_n)))$ ist. (Es muss ja auch $C \cdot D_{CB}(\Phi) = D_{SB}(\Phi)$ sein.) Die Spalten der Matrix $D_{CB}(\Phi)$ sind also nicht mehr die Bilder $\Phi(b_i)$ als Vektoren bezüglich der Standardbasis, sondern als Koordinatenvektoren bezüglich der Basis C .

Überhaupt muss W ja kein Vektorraum K^n sein, der eine Standardbasis besitzt. (Dann ist $D_{SB}(\Phi)$ eigentlich gar keine richtige Matrix.) $D_{CB}(\Phi)$ kann man jedoch immer bilden. Damit sind alle linearen Abbildungen, die es gibt, charakterisiert.

4.2.4 Basiswechsel

Für $\Phi : V \rightarrow W$ mit der Abbildungsmatrix $D_{C_1 B_1}(\Phi)$ bezüglich zwei Basen B_1 und C_1 von V bzw. W muss man oft eine andere Abbildungsmatrix $D_{C_2 B_2}(\Phi)$ bezüglich zwei anderen Basen konstruieren. Ich möchte hier kurz darstellen, wie man diese Matrix am schnellsten findet, wenn alle Basen als Matrizen geschrieben werden können. Diese sind dann sogar invertierbar, weil Basen linear unabhängig sind.

Ich gehe erst einmal davon aus, dass V und W Standardbasen S besitzen. Dann ist $D_{SS}(\Phi) \cdot B_1 = D_{SB_1}(\Phi) = C_1 \cdot D_{C_1 B_1}(\Phi)$, also $D_{SS}(\Phi) = C_1 \cdot D_{C_1 B_1}(\Phi) \cdot B_1^{-1}$. Analog ist aber auch $D_{SS}(\Phi) = C_2 \cdot D_{C_2 B_2}(\Phi) \cdot B_2^{-1}$. Also ist insgesamt $D_{C_2 B_2}(\Phi) = C_2^{-1} \cdot C_1 \cdot D_{C_1 B_1}(\Phi) \cdot B_1^{-1} \cdot B_2$. Dies gilt sogar ohne die Existenz der Standardbasis.

Diese Formel kann man sich wahrscheinlich schlecht merken, vor allem die Reihenfolge der einzelnen Faktoren. Allerdings kann man sich die einzelnen Schritte schnell herleiten, wenn man einmal verstanden hat, wie die Koordinatenvektoren funktionieren. Und zwar liefert $D_{CB}(\Phi)$ immer einen Koordinatenvektor bezüglich C , wenn man einen Vektor einsetzt, denn so war $D_{CB}(\Phi)$ ja gerade definiert. Um dies zu kompensieren, muss man mit C von links multiplizieren. Aber $D_{CB}(\Phi)$ nimmt einen Koordinatenvektor bezüglich B , den man erhält, indem man einen Vektor bezüglich der Standardbasis mit B^{-1} multipliziert. Analysiert man die obige Gleichung danach, an welcher Stelle was für ein Typ von Vektor vorliegt, dann sollte die Reihenfolge klar sein.

Besonders wichtig ist noch der Spezialfall $V = W, B_1 = C_1, B_2 = C_2$. Dann wird die obige Formel nämlich zu $D_{B_2 B_2}(\Phi) = B_2^{-1} \cdot B_1 \cdot D_{B_1 B_1}(\Phi) \cdot B_1^{-1} \cdot B_2$, und man sieht, dass die beiden Matrizen invers zueinander sind. (Es gilt ja $(B_1^{-1} \cdot B_2)^{-1} = B_2^{-1} \cdot B_1$). Eine Aufgabe wie „Finden Sie eine invertierbare Matrix S , so dass $D_{B_2 B_2}(\Phi) = S^{-1} \cdot D_{B_1 B_1}(\Phi) \cdot S$ ist“, ist also einfach durch $S = B_1^{-1} \cdot B_2$ zu lösen. Auch hier muss man natürlich aufpassen, dass man die Reihenfolge einhält.

Ich möchte hier noch auf einen kleinen Rechenrick aufmerksam machen, mit dem man eine Matrix gleichzeitig invertieren und von rechts mit einer anderen multiplizieren kann, hier z.B. $B_1^{-1} \cdot B_2$: Normalerweise wendet man beim Invertieren den Gaußalgorithmus auf die Matrix $(B_1|E)$ an; dabei werden B_1 und E von links mit einer invertierbaren Matrix C multipliziert, so dass $C \cdot B_1 = E$ ist, also $C = B_1^{-1}$ (siehe 3.3.3). Rechts bekommt man $C \cdot E = B_1^{-1} \cdot E = B_1^{-1}$. Aber in dieser Formel kann man schon erkennen, dass man statt E auch direkt B_2 benutzen kann (d.h. man fängt mit $(B_1|B_2)$ an), um statt $B_1^{-1} \cdot E$ eben $B_1^{-1} \cdot B_2$ zu bekommen. Übrigens muss B_2 zu diesem Zweck nicht notwendigerweise quadratisch sein.

Die Matrizen, die Zeilenumformungen durchführen, wenn man von links mit ihnen multipliziert, bewirken die gleichen Umformungen auf den Spalten, wenn man von rechts damit multipliziert. Auch damit kann man eine Matrix invertieren, denn es gilt $(A^{-1})^T = (A^T)^{-1}$ für alle invertierbaren Matrizen A . D.h. man kann zum Invertieren von B_1 auch Spaltenoperationen auf

$$\begin{pmatrix} B_1 \\ E \end{pmatrix}$$

durchführen, und das Produkt $B_2 \cdot B_1^{-1}$ analog zu $B_1^{-1} \cdot B_2$ berechnen, indem man E durch B_2 ersetzt. Oder man nutzt direkt aus, dass $B_2 \cdot B_1^{-1} = ((B_1^T)^{-1} \cdot B_2^T)^T$ ist.

4.2.5 Dimensionsformel

Für eine lineare Abbildung $\Phi : V \rightarrow W$ gilt die folgende Dimensionsformel, die man sich unbedingt merken sollte:

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

Dass Kern und Bild sich gegenseitig sozusagen ergänzen, kann man sich noch ganz gut vorstellen. Aber wie soll man sich merken, was auf der rechten Seite steht, $\dim V$ oder $\dim W$? Ganz einfach: Es kann nicht W sein, denn den Vektorraum W kann man problemlos vergrößern, ohne dass sich an Φ etwas ändern muss (d.h. sowohl Kern als auch Bild können gleich bleiben). Also muss V da stehen.

Diese Dimensionsformel kann man an vielen Stellen anwenden, wo es um Dimensionen und lineare Abbildungen geht. Zusammen mit dem nun folgenden Dimensionsformeln für Summen und Faktorräume kann man damit viele Dimensionsaufgaben schnell lösen.

4.3 Summen und Faktorräume

4.3.1 Summen von Vektorräumen

Die Summe von zwei Vektorräumen ist das, was entsteht, wenn man Vektoren aus beiden Vektorräumen addiert. Da die Addition kommutativ ist, ist dies das Erzeugnis der beiden Vektorräume, genauer gesagt das Erzeugnis der Vereinigung.

Ähnlich wie beim Erzeugnis einzelner Vektoren kann man dabei betrachten, wie viel beim Bilden der Summe überflüssig ist. Das heißt, dass einer der beiden Vektorräume kleiner sein könnte, und es ergibt sich trotzdem das selbe Ergebnis. Es ist genau dann der Fall, wenn der Schnitt ein nichttrivialer Vektorraum ist (also nicht nur aus dem Nullvektor besteht). Besteht der Schnitt nur aus dem Nullvektor, könnte man sagen, dass die beiden Vektorräume nichts miteinander zu tun haben, genau wie linear unabhängige Vektoren in gewisser Weise nichts miteinander zu tun haben. Man nennt dann die Summe „direkt“.

Für die Dimension der Summe gilt also genau das, was man erwarten würde: Ist die Summe direkt, dann haben wir zwei Vektorräume, die nichts miteinander zu tun haben, also addieren sich die Dimensionen. Z.B. ist die Summe von zwei verschiedenen Geraden (1-dimensional) die Ebene (2-dimensional), die diese Geraden enthält. Im allgemeinen Fall ist der Schnitt der beiden Vektorräume ein Indikator dafür, wie viel überflüssig ist, also muss man die Dimension des Schnitts noch subtrahieren.

Wenn zu jedem der beiden Vektorräume ein Erzeugendensystem gegeben ist, dann ist die Vereinigung ein Erzeugendensystem der Summe. Sind beides disjunkte Basen, dann bildet die Vereinigung also genau dann eine Basis, wenn die Summe direkt ist. Um die Direktheit festzustellen, kann man also prüfen, ob die Basisvektoren der beiden Vektorräume insgesamt linear unabhängig sind.

4.3.2 Faktorräume

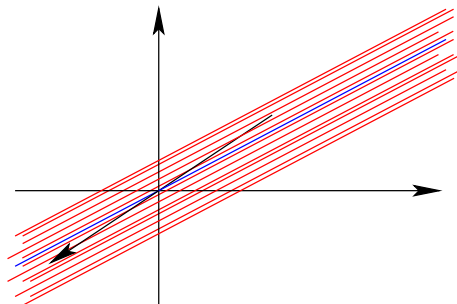
Um zu erklären, was bei der Faktorisierung von Vektorräumen passiert, möchte ich kurz einschieben, wie die Faktorisierung bei Gruppen und Ringen funktioniert hat. Ich habe das in diesem Skript nicht beschrieben, weil nur das Rechnen in den Restklassenringen \mathbb{Z}_m in der Vorlesung wirklich von Bedeutung war. Deshalb werde ich anhand dieses Beispiels die allgemeine Faktorisierung von Gruppen erläutern.

Zur Faktorisierung einer Menge gehören immer eine Klasseneinteilung (Partition) und eine Äquivalenzrelation, die sich gegenseitig bedingen (siehe 2.4). Normalerweise wird die Äquivalenzrelation angegeben, aber man kann auch erst die Klasseneinteilung vornehmen und sich dann eine möglichst einfache Relation dazu überlegen. Im Falle \mathbb{Z}_m hat man die Klassen $[x]$ mit $x \in \mathbb{Z}$, wobei $[0] =$

$[m] = \dots, [1] = [m+1] = \dots$ ist usw. Die Klassen sind aber Mengen, und zwar die Mengen aller Elemente mit gleicher Klasse. Z.B. ist $[0] = \{\dots, -m, 0, m, 2 \cdot m, \dots\} = \{m \cdot z : z \in \mathbb{Z}\} =: m \cdot \mathbb{Z}$. Oder $[1] = \{\dots, -m+1, 1, m+1, 2 \cdot m+1, \dots\} = \{1 + m \cdot z : z \in \mathbb{Z}\} =: 1 + m \cdot \mathbb{Z}$. Allgemein ist $[x] = \{x + m \cdot z : z \in \mathbb{Z}\} = x + m \cdot \mathbb{Z}$. Das heißt: In der Klasse von x liegen x selbst und alle Zahlen, die sich um ein Vielfaches von m davon unterscheiden. \mathbb{Z}_m , die Menge dieser Klassen, ist also: $\mathbb{Z}_m = \{[x] : x \in \mathbb{Z}\} = \{x + m \cdot \mathbb{Z} : x \in \mathbb{Z}\} =: \mathbb{Z}/m \cdot \mathbb{Z}$.

Genau so bildet man auch Faktorräume. Bei der Faktorisierung eines Vektorraums V nach einem Untervektorraum U sind die entstehenden Klassen $[x] = x + U = \{x + u : u \in U\}$ für $x \in V$. Es ist also $V/U = \{x + U : x \in V\}$. Wie diese Mengen im \mathbb{R}^3 aussehen, kann man sich anhand des „Spaghetti- und Lasagne-Modells“ klarmachen:

Es gibt im \mathbb{R}^3 genau zwei nichttriviale Arten von Unterräumen, nämlich Geraden und Ebenen. Was passiert also, wenn U eine Gerade ist (durch den Ursprung, sonst wäre es kein Vektorraum)? Dann sind auch alle Menge $x + U$ Geraden, allerdings nicht durch den Ursprung. Es sind alle Geraden, die parallel zu U verlaufen, wie Spaghetti. Und die Faktormenge ist die Menge aller dieser Geraden, also eine Menge von Spaghetti, in der jedes Element eine Spaghetti ist. Mit Ebenen verhält es sich gleich; hier hat man es mit Lasagne statt Spaghetti zu tun.



Nur diese Mengen zu betrachten, wäre uninteressant. Faktorisierung enthält immer auch Verknüpfungen, die man auf der Faktormenge definiert. Und zwar überträgt man die Verknüpfungen von den Elementen auf die Klassen, z.B. $[x] + [y] := [x + y]$. Bei Gruppen ist dies nur unter einer speziellen Voraussetzung möglich; bei Vektorräumen ist diese Voraussetzung immer erfüllt. Also bildet die Faktormenge wieder einen Vektorraum.

Wieder ist das „Spaghetti-/Lasagne-Modell“ gefragt: Jede Spaghetti bzw. jedes Lasagneblatt bildet einen Vektor. Um Vektoren zu addieren und zu multiplizieren, könnte man z.B. ein Brett durch die Spaghetti legen, oder vielleicht sollten es dann besser Nägel sein. Jetzt hat man eine Fläche, die isomorph zu \mathbb{R}^2 ist. Aber man kommt auch ohne das Brett aus, wenn man die Spaghetti aus einem Blickwinkel betrachtet, aus dem sie zu Punkten werden.

Hier sieht man auch schon, wie sich die Dimensionen verhalten: Es ist $\dim V/U = \dim V - \dim U$.

4.4 Dualräume

4.4.1 Definition

Ist V ein K -Vektorraum, dann ist der Dualraum V^* kurz gesagt die Menge der linearen Abbildungen von V nach K , als Vektorraum über K . Diese nennt man auch „Linearformen“. Im Gegensatz zu den anderen Objekten der Linearen Algebra gibt es dazu weder eine anschauliche Erklärung noch einen Spezialfall, den man schon kennt. Also muss man sich beim Arbeiten damit immer wieder die Definition ins Gedächtnis rufen und alles darauf zurückführen.

4.4.2 Duale Basis

Ist V n -dimensional, dann ist V^* isomorph zu V . Ein entsprechender Isomorphismus ist zwar meistens künstlich und hat keine theoretische Bedeutung; wichtig ist aber die Folgerung, dass jede Basis von V^* n Elemente hat. Ist $B = (b_1 \ b_2 \ \dots \ b_n)$ eine (geordnete) Basis von V , dann kann man nach einer Basis $B^* = (b_1^* \ b_2^* \ \dots \ b_n^*)$ von V^* suchen, die mit B ein besonderes Verhältnis hat. B^* nennt man die „duale Basis“ zu B .

Die Elemente b_i^* von B^* sind natürlich Linearformen, d.h. lineare Abbildungen von V nach K . Worauf könnten sie ein $x \in V$ abbilden? D.h. wie steht x mit b_i in Verbindung? Es gibt eigentlich nur eine Verbindung zwischen x und b_i , nämlich die i -te Koordinate von x . Am besten schreibt man den Koordinatenvektor von x aus:

$$D_B(x) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Dann ist a_i die Koordinate zu b_i . Die duale Basis B^* ist gerade so definiert, dass $b_i^*(x) = a_i$ ist.

Anstatt für ein beliebiges $x \in V$ anzugeben, worauf b_i^* es abbilden soll, reicht es bekanntlich, die Bilder der Basisvektoren b_j von V festzulegen (siehe 4.2.3). Was ist aber der Koordinatenvektor von b_j ? Er hat an der j -ten Stelle eine 1, und alle anderen Stellen sind 0. Also ist eine einfachere Definition der dualen Basis:

$$b_i^*(b_j) = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{sonst} \end{cases}$$

Oder als Abbildungsmatrix:

$$D_{EB}(b_i^*) = (0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0)$$

(wobei die 1 in der i -ten Spalte steht und $E = (1)$ die Standardbasis von K ist)

Soll man die duale Basis zu B bestimmen, dann ist es aber eher angebracht, konkret für jedes $x \in V$ anzugeben, worauf es von b_i^* abgebildet wird. Ist $V = K^n$ mit der Standardbasis S , dann ist $D_{ES}(b_i^*)$ die i -te Zeile von B^{-1} . Das ergibt sich beim Basiswechsel (siehe 4.2.4), oder man überlegt sich, wie man $D_B(x)$ aus x berechnet: Da $x = B \cdot D_B(x)$ ist, gilt $D_B(x) = B^{-1} \cdot x$, und damit ist die i -te Komponente von $D_B(x)$ gerade die i -te Zeile von B^{-1} mal x .

Ist umgekehrt eine duale Basis $B^* = (b_1^* \ b_2^* \ \dots \ b_n^*)$ gegeben und B zu berechnen (im K^n), dann bestimmt man zunächst einmal Abbildungsmatrizen der b_i^* bezüglich der Standardbasis, schreibt diese als Zeilen in eine einzige $n \times n$ -Matrix und invertiert diese. Das Ergebnis ist dann B .

Wenn V nicht der Standardraum K^n ist, dann sucht man am besten zunächst eine möglichst einfache Basis C von V und untersucht $D_{EC}(b_i^*)$ statt $D_{ES}(b_i^*)$. In diesem Fall sollte man wirklich den Basiswechsel durchführen, oder man stellt sich die Elemente aus V einfach als Koordinatenvektoren bezüglich C vor und rechnet wie in K^n .

4.4.3 Duale Abbildung

Ist eine lineare Abbildung Φ zwischen zwei K -Vektorräumen V und W gegeben ($\Phi : V \rightarrow W$), dann kann man sich überlegen, wie man eine analoge Abbildung zwischen den Dualräumen V^* und W^* definiert. Es ist nicht kompliziert, aus Φ eine Linearform zu machen, d.h. eine lineare Abbildung

nach K . Dazu muss man Φ nur mit einer linearen Abbildung von W nach K verketteten; dann erhält man eine lineare Abbildung von V nach K . Aber eine lineare Abbildung von W nach K ist ein Element aus W^* , und eine lineare Abbildung von V nach K ist ein Element aus V^* . Damit erhält man eine Abbildung:

$$\Phi^* : W^* \rightarrow V^*, y \mapsto y \circ \Phi$$

Diese Abbildung ist selbst linear, man nennt sie die „duale Abbildung“ zu Φ . Man beachte, dass sie in der umgekehrten Richtung definiert ist. Aber: Auch wenn im endlichdimensionalen Fall V zu V^* und W zu W^* isomorph ist, liefert das keine lineare Abbildung von W nach V , die allein durch Φ definiert ist. Denn die Isomorphismen sind immer abhängig von willkürlich gewählten Basen von V und W .

Im diesem Fall kann man jedoch für V und W Basen B und C festlegen und Φ mit Hilfe von $D_{CB}(\Phi)$ angeben. Wie sieht die Abbildungsmatrix von Φ^* bezüglich C^* und B^* aus? Es ist $D_{B^*C^*}(\Phi^*) = (D_{CB}(\Phi))^T$, weswegen man Φ^* auch als „transponierte Abbildung“ Φ^T bezeichnet.

Da Quelle und Ziel von Φ^* gegenüber Φ gerade vertauscht sind, ist es wahrscheinlich nicht verwunderlich, dass sich auch die Richtung der Verkettung umdreht, d.h. $(\Phi \circ \Psi)^* = \Psi^* \circ \Phi^*$. Außerdem sind Kern und Bild vertauscht in dem Sinne, dass man Kern Φ^* aus Bild Φ berechnen kann und umgekehrt:

$$\begin{aligned} \text{Kern } \Phi^* &= \{y \in W^* : \Phi^*(y) = y \circ \Phi = 0\} = \\ &= \{y \in W^* : y(\Phi(x)) = 0 \forall x \in V\} = \\ &= \{y \in W^* : y(z) = 0 \forall z \in \text{Bild } \Phi\} = \\ &= \{y \in W^* : y(\text{Bild } \Phi) = \{0\}\} \\ \text{Bild } \Phi^* &= \{z \in V^* : \exists y \in W^* : z = \Phi^*(y) = y \circ \Phi\} = \\ &= \{z \in V^* : \exists y \in W^* : z(x) = y(\Phi(x)) \forall x \in V\} = \\ &= \{z \in V^* : z(\text{Kern } \Phi) = \{0\}\} \text{ (Homomorphiesatz, siehe Beispiel in 3.1.4)} \\ \text{Kern } \Phi &= \{x \in V : \Phi(x) = 0\} = \\ &= \{x \in V : y(\Phi(x)) = (\Phi^*(y))(x) = 0 \forall y \in W^*\} = \\ &= \{x \in V : z(x) = 0 \forall z \in \text{Bild } \Phi^*\} \\ \text{Bild } \Phi &= \{z \in W : \exists x \in V : z = \Phi(x)\} = \\ &= \{z \in W : \exists x \in V : y(z) = y(\Phi(x)) = (\Phi^*(y))(x) \forall y \in W^*\} = \\ &= \{z \in W : y(z) = 0 \forall y \in \text{Kern } \Phi^*\} \end{aligned}$$

Insbesondere ist Φ^* genau dann injektiv, wenn Φ surjektiv ist, und umgekehrt.

Außerdem gilt für eine weitere lineare Abbildung $\Psi : W \rightarrow X$ genau dann $\text{Kern } \Psi = \text{Bild } \Phi$, wenn $\text{Kern } \Phi^* = \text{Bild } \Psi^*$ ist:

$$\begin{aligned} \text{Kern } \Psi = \text{Bild } \Phi &\Rightarrow \text{Kern } \Phi^* = \{y \in W^* : y(\text{Bild } \Phi) = y(\text{Kern } \Psi) = \{0\}\} = \text{Bild } \Psi^* \\ \text{Kern } \Phi^* = \text{Bild } \Psi^* &\Rightarrow \text{Kern } \Psi = \{z \in W : y(z) = 0 \forall y \in \text{Bild } \Psi^* = \text{Kern } \Phi^*\} = \text{Bild } \Phi \end{aligned}$$

Ein kleiner Anwendungsfall ist die Darstellung eines gegebenen Untervektorraums U von V (endlichdimensional) als Lösungsmenge eines homogenen LGS, also als Kern einer linearen Abbildung $\Phi : V \rightarrow W$, wobei W ein endlichdimensionaler Vektorraum über dem selben Grundkörper ist.

Sei B eine geordnete Basis von U , dann hat die Abbildung $\Psi : K^{\dim U} \rightarrow V, x \mapsto B \cdot x$ als Bild gerade U . Kern $\Psi^ = \{y \in V^* : y(\text{Bild } \Psi) = y(U) = \{0\}\}$ (s.o.). Man bestimme eine geordnete Basis C davon (die Abbildungsmatrix von Ψ^* bezüglich entsprechender Basen ist B^T) und erhält*

damit $\Phi^* : W^* \rightarrow V^*$, $z \mapsto C \cdot D_{S^*}(z)$ (mit entsprechend gewähltem W mit Standardbasis S), so dass $\text{Bild } \Phi^* = \text{Kern } \Psi^*$ ist und deshalb $\text{Kern } \Phi = \text{Bild } \Psi = U$.

Dieses Beispiel soll verdeutlichen, dass Fälle, in denen mit transponierten Matrizen gearbeitet wird, oft allgemeiner mit Hilfe von Dualräumen und dualen Abbildungen beschrieben werden können.

4.4.4 Alternative Definition für Abbildungsmatrizen

Da duale Basen eine Darstellung für die Koordinaten eines Vektors bezüglich der ursprünglichen Basis liefern, liegt es nahe, dass Abbildungsmatrizen mit Hilfe von dualen Basen einfacher definiert werden können. Sind V, W, B, C, Φ wie immer und $D_{CB}(\Phi) = ((a_{ij}))$, wie kann man dann ein a_{ij} genau berechnen? Bei der Definition (4.2.3) wurden Abbildungsmatrizen immer spaltenweise angegeben:

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = D_C(\Phi(b_j))$$

Aber jetzt gibt es die Möglichkeit, eine einzelne Komponente von $D_C(\Phi(b_j))$ und damit ein einzelnes a_{ij} zu ermitteln. Die duale Basis C^* war gerade so definiert, dass $c_i^*(x)$ die i -te Komponente von $D_C(x)$ für ein $x \in W$ ist, also:

$$a_{ij} = c_i^*(\Phi(b_j))$$

4.5 Determinanten

4.5.1 Definition

Die übliche Definition der Determinante ist schwer zu verstehen und anzuwenden. Also merkt man sich lieber nur, wie man sie berechnet. Das taugt natürlich auch als Definition; allerdings gibt es beim Berechnen sehr viele Wahlmöglichkeiten, und es leuchtet nicht sofort ein, dass das Ergebnis davon unabhängig ist. Benutzt man die mathematische Definition, ist dies wiederum klar.

Definiert ist die Determinante für eine quadratischen Matrix $R^{n \times n}$ über einem kommutativen Ring R . Bei endlichdimensionalen Vektorräumen kann man auch von der Determinante einer linearen Abbildung $\Phi : V \rightarrow V$ sprechen, denn die Determinante der Abbildungsmatrix $D_{BB}(\Phi)$ ist unabhängig von der Basis B . Die Determinante drückt gewissermaßen die gesamte Matrix in einer Zahl aus. Dabei gehen natürlich Informationen verloren, aber es bleiben einige Eigenschaften erhalten, wie z.B. die Invertierbarkeit und das Verhalten bei der Multiplikation.

Um die Determinante auszurechnen, merkt man sich am besten die folgenden Regeln (in dieser Reihenfolge). Ich habe jeweils typische Beispiele dazugeschrieben, bei denen man sich leicht überlegen kann, wie es weitergeht.

1. Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente:

$$\det \begin{pmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix} = \det \begin{pmatrix} a & 0 & 0 \\ * & b & 0 \\ * & * & c \end{pmatrix} = a \cdot b \cdot c$$

Insbesondere ist die Determinante der Einheitsmatrix immer 1.

2. Addiert man ein Vielfaches einer Zeile oder Spalte zu einer anderen, ändert sich die Determinante nicht:

$$\det \begin{pmatrix} a & b & c \\ a & d & e \\ a & f & g \end{pmatrix} = \det \begin{pmatrix} a & b & c \\ 0 & d-b & e-c \\ 0 & f-b & e-c \end{pmatrix}$$

3. Multipliziert man *eine* Zeile oder Spalte mit einem Faktor k , vervielfacht sich die Determinante um k . Multipliziert man die *ganze* Matrix mit k , vervielfacht sich die Determinante also um k^n :

$$\det \begin{pmatrix} a & ab & ac \\ 1 & d & e \\ 1 & f & g \end{pmatrix} = a \cdot \det \begin{pmatrix} 1 & b & c \\ 1 & d & e \\ 1 & f & g \end{pmatrix}$$

(Vorsicht: Hier ist $k = \frac{1}{a}$. Da der Wert der Determinante durch a geteilt wird, muss man mit a multiplizieren, um den ursprünglichen Wert wiederzubekommen.)

4. Vertauscht man zwei Zeilen oder Spalten, ändert sich das Vorzeichen der Determinante:

$$\det \begin{pmatrix} a & * & * \\ 0 & 0 & c \\ 0 & b & * \end{pmatrix} = -\det \begin{pmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{pmatrix}$$

Da in einem Körper die Schritte des Gaußalgorithmus offenbar den Wert der Determinante nur um einen Faktor $k \neq 0$ verändern und das Ergebnis des Gaußalgorithmus immer eine Dreiecksmatrix ist, ist die Determinante ein Kriterium dafür, ob die Matrix invertierbar ist. Ist dies nämlich der Fall, dann stehen hinterher auf der Diagonalen nur Einsen, und damit ist die Determinante von 0 verschieden. Ist es nicht der Fall, dann steht auf der Diagonalen mindestens eine Null, so dass die Determinante gleich 0 ist. Die Determinante ist also immer dann 0, wenn der Rang der Matrix kleiner als n ist; z.B. schon dann, wenn die Matrix eine Nullzeile oder -spalte enthält.

Besteht eine Matrix nur aus Zahlen, kann man mit diesen Regeln immer recht schnell die Determinante berechnen. Problematischer wird es, wenn Variablen oder Polynome in der Matrix stehen; dann sollte man unbedingt auch die folgenden Verfahren anwenden können:

4.5.2 Formeln

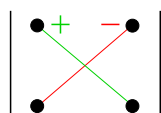
Für $n \leq 3$ existieren einfache Formeln, die man sich merken muss. Die Definition der Determinante ergibt sofort, dass die Determinante für $n = 0$ immer 1 ist, und dass für $n = 1$ gilt:

$$\det(a) = a$$

Für $n = 2$ hat man:

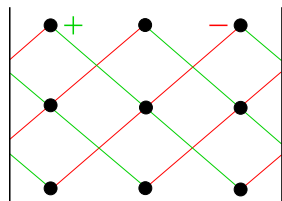
$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \cdot d - b \cdot c$$

D.h. die Diagonale von links oben nach rechts unten geht positiv in die Determinante ein, die andere negativ:



Wie merkt man sich am besten, welche Diagonale positiv und welche negativ zählt? Man kann sich z.B. den Spezialfall der Einheitsmatrix anschauen; die Determinante davon muss 1 sein, damit muss $a \cdot d$ positiv eingehen.

Für $n = 3$ wird die Formel schon recht lang, deshalb sollte man sich direkt das Schaubild einprägen:



Es sieht erst einmal sehr ähnlich aus, aber es gibt einen wesentlichen Unterschied: Hier zählen auch die Diagonalen, die über den Rand hinausgehen. Während man bei $n = 2$ nur 2 Diagonalen betrachten musste, sind es bei $n = 3$ schon 6. Für $n = 4$ lohnt es sich schon nicht mehr, die Formel explizit anzugeben; es kommen auch nicht mehr nur Diagonalen darin vor.

4.5.3 Aufteilung der Matrix

Für Determinanten von Matrizen gilt (wie an vielen anderen Stellen), dass man eine Matrix aufteilen kann in einzelne Teilmatrizen, um dann mit den Determinanten der einzelnen Matrizen die Determinante der großen Matrizen auszurechnen. Leider ist aber der Matrizenring nicht kommutativ, so dass man sich überlegen muss, welche der Formeln noch gelten, wenn die Elemente von Matrizen selbst wieder Matrizen sind. Ein Variante funktioniert immer:

Teilt man die Matrix so in Blöcke auf, dass die Diagonale nur quadratische Blöcke enthält und alle Blöcke unter- oder oberhalb der Diagonalen 0 sind, dann ist die Determinante der Matrix das Produkt der Determinanten der Blöcke auf der Diagonalen:

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

(wobei A , B und C quadratische Matrizen sind)

Beispiel:

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \cdot \det \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \cdot \det (5) = (1 \cdot 3 - 2 \cdot 2) \cdot (2 \cdot 2 - 3 \cdot 1) \cdot 5 = -5$$

4.5.4 Laplace-Entwicklung

Mit der Laplace-Entwicklung kann man eine Matrix schrittweise verkleinern, bis man die Determinante leicht ausrechnen kann. Man sucht sich eine Zeile oder Spalte mit vielen Nullen (im obigen Beispiel z.B. die letzte Zeile). Dann streicht man die Zeile bzw. Spalte, und außerdem nacheinander

die Spalten bzw. Zeilen, in denen der weggestrichene Eintrag nicht 0 war (oben wäre das die 5, also nur die letzte Spalte).

Die Determinanten der so gebildeten Matrizen muss man mit den jeweiligen Einträgen multiplizieren und entweder addieren oder subtrahieren, je nachdem, wo der Eintrag stand. Das geht nach folgendem schachbrettartigen Schema:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Entwickelt man im Beispiel oben nach der letzten Zeile, erhält man:

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} = 5 \cdot \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Wenn man will, kann man aber auch z.B. nach der ersten Spalte entwickeln:

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} = 1 \cdot \det \begin{pmatrix} 3 & 4 & 5 & 6 \\ 0 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 5 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

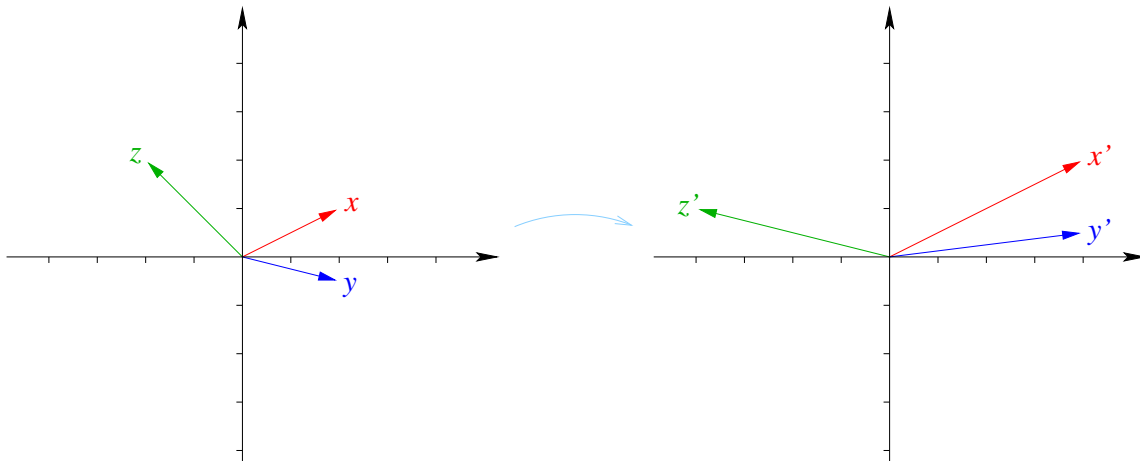
Im ersten Summanden wurde die erste Zeile weggestrichen, im zweiten Summanden die zweite, entsprechend dem jeweiligen Eintrag. Außerdem war der zweite Eintrag an einer Stelle, wo im Schema ein Minus steht, daher das Minus vor der 2.

4.6 Eigenwerte

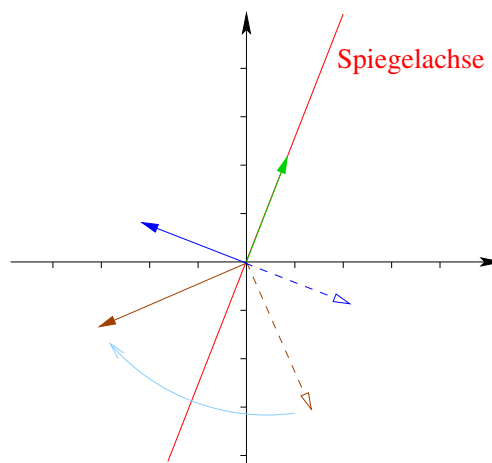
4.6.1 Definition

Wenn $\Phi : V \rightarrow V$ eine lineare Abbildung ist, ist es interessant zu wissen, ob ein bestimmter Vektor $x \in V$ von Φ z.B. wieder auf sich selbst abgebildet wird, d.h. $\Phi(x) = x$. Allgemeiner kann man auch fragen, ob der Vektor zwar nicht auf sich selbst abgebildet wird, aber auf ein Vielfaches davon, d.h. $\Phi(x) = c \cdot x$, $x \in K$. In dem Fall nennt man x einen „Eigenvektor“ zum „Eigenwert“ c . Der Spezialfall $\Phi(x) = x$ besagt damit, dass x Eigenvektor zum Eigenwert 1 ist.

In dieser Illustration ist eine lineare Abbildung $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dadurch dargestellt, was sie mit verschiedenen Vektoren x, y, z macht ($x' = \Phi(x)$, $y' = \Phi(y)$, $z' = \Phi(z)$). x ist ein Eigenvektor zum Eigenwert 2, d.h. er wird durch Φ um den Faktor 2 gestreckt. Auch y und z werden durch den Eigenwert 2 beeinflusst: Schaut man sich die beiden Grafiken genau an, sieht man, dass Alles in Richtung von x gestreckt wurde. x und y sind aber selbst keine Eigenvektoren zu irgendeinem Eigenwert. Irgendwo ist noch der Eigenwert 1 versteckt; es dürfte eine gute Übungsaufgabe sein, einen Eigenvektor dazu zu finden.



Um den Sinn von Eigenwerten besser zu verstehen, kann man sich z.B. die Eigenwerte bei einer Spiegelung im \mathbb{R}^2 anschauen. Die Vektoren, die direkt auf der Spiegelachse liegen, werden nicht verändert, d.h. sie sind Eigenvektoren zum Eigenwert 1. Die Vektoren, die genau senkrecht dazu stehen, werden negiert, sind also Eigenvektoren zum Eigenwert -1 . Die übrigen Vektoren sind keine Eigenvektoren:



Ist x ein Eigenvektor zum Eigenwert c , dann ist klar, dass auch $k \cdot x$ für $k \in K$ ein Eigenvektor zum Eigenwert c ist. Das Gleiche gilt für beliebige Linearkombinationen von Vektoren zum gleichen Eigenwert. D.h. die Menge $\text{Eig}_c(\Phi)$ der Eigenvektoren zum Eigenwert c ist ein Unterraum (wird auch mit $E_c(\Phi)$ bezeichnet, aber man darf es nicht mit der Bezeichnung für die Einheitsmatrix verwechseln). Speziell ist $\text{Eig}_0(\Phi) = \text{Kern } \Phi$. Man nennt c eigentlich nur dann „Eigenwert von Φ “, wenn dies nicht der Nullraum ist, d.h. wenn es einen Eigenvektor gibt, der nicht der Nullvektor ist. Man spricht aber trotzdem immer vom „Eigenraum zum Eigenwert c “.

Der Schnitt von zwei Eigenräumen ist immer der Nullraum, denn ein Vektor kann nicht gleichzeitig Eigenvektor zu zwei verschiedenen Eigenwerten sein. Führt man diesen Gedanken fort, gelangt man zu dem Schluss, dass auch mehr Eigenräume immer eine direkte Summe bilden. Allerdings ist die Summe aller Eigenräume *nicht* immer der ganze Vektorraum. Z.B. hat im \mathbb{R}^2 eine Drehung um den Ursprung gar keine Eigenwerte.

Noch allgemeiner kann man sich auch beliebige Unterräume anschauen und prüfen, ob jeder Vektor aus dem Unterraum wieder darin landet, d.h. $\Phi(U) \subset U$. Man nennt U dann „ Φ -invariant“. Die Fragestellung hängt damit zusammen, weil man für eindimensionales U damit $\Phi(x) = c \cdot x$ ($c \in K$) für alle $x \in U$ erreicht. Auch andere Φ -invariante Unterräume lassen sich über Eigenwerte charakterisieren (Stichwort „Haupträume“), aber das darf man nicht mit den Eigenräumen verwechseln.

4.6.2 Einfache Fälle

Ist $V = K^n$ und $\Phi : V \rightarrow V, x \mapsto A \cdot x$ die Multiplikation mit einer Diagonalmatrix

$$A = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_n \end{pmatrix},$$

dann ist es recht einfach, n Eigenvektoren anzugeben: Es ist

$$\Phi \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \Phi \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ c_2 \\ 0 \\ \vdots \end{pmatrix}, \quad \dots, \quad \Phi \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ c_n \end{pmatrix},$$

also sind diese Vektoren jeweils Eigenvektoren zu den Eigenwerten c_1 bis c_n . Damit ist ganz V die direkte Summe der Eigenräume, die von den Standardbasisvektoren aufgespannt werden. Die c_i müssen natürlich nicht alle verschieden sein; bei gleichen c_i bekommt man Eigenräume höherer Dimension.

Man kann diesen Spezialfall etwas verallgemeinern, indem man nur voraussetzt, dass die Abbildungsmatrix bezüglich einer beliebigen Basis eine Diagonalmatrix ist. Die Basisvektoren sind dann die Eigenvektoren. Hat man umgekehrt eine Basis aus Eigenvektoren (indem man vorher die Eigenräume bestimmt hat und die Summe davon tatsächlich ganz V ist), dann ist die Abbildungsmatrix bezüglich dieser Basis immer eine Diagonalmatrix.

Ebenfalls ein wichtiger Spezialfall sind Projektionen. Man nennt Φ eine „Projektion“, wenn $\Phi^2 = \Phi$ gilt, was man z.B. anhand einer Abbildungsmatrix recht leicht überprüfen kann. Da alle Vektoren aus dem Bild wieder auf sich selbst abgebildet werden müssen, hat Φ bezüglich einer Basis, die nur aus Vektoren aus Kern und Bild bestehen, eine Diagonalmatrix als Abbildungsmatrix. Auf der Diagonalen stehen nur Nullen und Einsen, deshalb sind das die beiden einzigen Eigenwerte, die vorkommen. Sind umgekehrt 0 und 1 die einzigen Eigenwerte einer Abbildung, und ist die Summe der zugehörigen Eigenräume ganz V , dann ist die Abbildung eine Projektion.

4.6.3 Allgemeine Formeln

Ist V endlichdimensional und B eine Basis, dann ist $x \in V$ genau dann Eigenvektor zum Eigenwert c von Φ , wenn $D_{BB}(\Phi) \cdot D_B(x) = c \cdot D_B(x)$ ist. Man kann also statt V auch gleich K^n betrachten und statt Φ die Multiplikation mit der Matrix $D_{BB}(\Phi)$. Die Eigenwerte sind die gleichen, und die Eigenvektoren sind gerade die Koordinatenvektoren der Eigenvektoren von Φ . Daher kann man sich von jetzt an auch einschränken auf $V = K^n$ und $\Phi : V \rightarrow V, x \mapsto A \cdot x$ mit einer $n \times n$ -Matrix A . Man spricht dann auch von den „Eigenwerten von A “.

In $A \cdot x = c \cdot x$ kann man aber das Distributivgesetz anwenden; es ist äquivalent zu $A \cdot x - c \cdot x = (A - c \cdot E_n) \cdot x = 0$. Das liefert schon einmal eine Möglichkeit, den Eigenraum zum Eigenwert c auszurechnen; es ist der Kern von $A - c \cdot E_n$. Jetzt muss man nur noch alle Eigenwerte ermitteln, damit man weiß, welche c man einsetzen muss.

Um ein Gefühl dafür zu bekommen, welche Rolle hier Polynome und Determinanten spielen, kann man sich erst einmal überlegen, dass der Term $A - c \cdot E_n$ sehr danach aussieht, als wäre hier die

Matrix A in das Polynom $X - c$ eingesetzt worden. Zwar ist der Wert davon nicht selbst 0, aber zumindest werden alle $x \in \text{Eig}_c(\Phi)$ auf 0 abgebildet. Das wird gleich noch wichtig.

Zunächst wird aber der Term auf eine *andere* Art als Polynom betrachtet: Man sucht ja alle c , so dass $(A - c \cdot E_n) \cdot x = 0$ auch für ein $x \neq 0$ gilt, d.h. dass das Gleichungssystem nichttrivial lösbar ist. Das ist genau dann der Fall, wenn die Determinante von $A - c \cdot E_n$ gleich 0 ist. Ersetzt man c (nicht A) durch X , dann ist die Determinante ein Polynom. Die Eigenwerte von A sind die Nullstellen. Man nennt dieses Polynom das „charakteristische Polynom“ von A oder Φ .

Zur Bestimmung des charakteristischen Polynoms muss man also in A auf der Diagonalen überall X subtrahieren und dann die Determinante bilden. Damit das nicht zu kompliziert wird, sollte man geschickt die verschiedenen Methoden (wie Laplace-Entwicklung, siehe 4.5.4) ausnutzen. Das Resultat ist auf jeden Fall ein Polynom, aber es hat eine recht komplizierte Form; je nachdem, wie geschickt man sich anstellt.

Um die Nullstellen zu bestimmen, rät man jeweils eine Nullstelle und benutzt dann Polynomdivision (siehe 3.2.5), oder man verwendet direkte Formeln. Auf jeden Fall sollte man beim Ausrechnen der Determinante schon darauf achten, gemeinsame Terme auszuklammern, wo es möglich ist. Teilt man die Matrix in Blöcke auf (siehe 4.5.3), dann ist dies automatisch gegeben. Bei der Laplace-Entwicklung sollte am besten nur ein einziger Term übrig bleiben, sonst muss man selbst nach Möglichkeiten zum Ausklammern suchen.

Das charakteristische Polynom hat immer den Grad n , und der erste Koeffizient ist $(-1)^n$. Wenn die Summe der Eigenräume nicht ganz V ist, dann kann sich das auf zwei verschiedene Arten im Polynom widerspiegeln: Entweder das Polynom lässt sich nicht vollständig als Produkt von Faktoren der Form $(X - c_i)$ schreiben (man sagt, es „zerfällt“ nicht in „Linearfaktoren“), oder ein Faktor $(X - c_i)$ kommt mehrmals vor (d.h. als $(X - c_i)^k$). Das heißt nämlich noch *nicht*, dass die Dimension des zugehörigen Eigenraums k ist; sie kann auch kleiner sein.

Ist p das charakteristische Polynom von A bzw. Φ , dann gilt nach dem Satz von Cayley-Hamilton $p(A) = 0$ bzw. $p(\Phi) = 0$. Für den Fall, dass p in Linearfaktoren zerfällt, kann man sich nach der Bemerkung oben (darüber, was passiert, wenn man A in das Polynom $X - c$ einsetzt) vielleicht ungefähr vorstellen, warum das so ist. Es ist natürlich kein Beweis.

Dies kann übrigens ganz nützlich sein, um zu überprüfen, ob man das charakteristische Polynom richtig ausgerechnet hat. Es ist aber recht mühsam. Zerfällt das Polynom in Linearfaktoren, dann rechnet man besser zu den gefundenen Eigenwerten die Eigenräume aus; oft ist es ohnehin Teil einer Aufgabe.

4.7 Jordan-Normalform

4.7.1 Beschreibung

Jede lineare Selbstabbildung eines n -dimensionalen K -Vektorraums V , zu der eine Abbildungsmatrix $A \in K^{n \times n}$ bezüglich einer Basis gegeben ist, lässt sich durch Basiswechsel (siehe 4.2.4) in die sogenannte „Jordan-Normalform“ bringen. Diese ist bis auf die Reihenfolge bestimmter Teile (entspricht der Reihenfolge der Basisvektoren) eindeutig und hat eine sehr einfache Gestalt. Das hat viele Vorteile, z.B.:

- Ist eine Abbildungsmatrix in Jordan-Normalform gegeben, kann man Vieles direkt ablesen, denn es ist eine Dreiecksmatrix: Rang, Determinante, Eigenwerte, charakteristisches Polynom, Verhalten beim Potenzieren, invariante Unterräume, usw. Diese Eigenschaften ändern sich beim Basiswechsel nicht.

- Da jede Matrix eine Jordan-Normalform besitzt, kann man sich bei Beweisen oft auf Matrizen in Jordan-Normalform beschränken.
- Sie liefert ein Entscheidungskriterium, ob zwei Matrizen durch Basiswechsel ineinander überführt werden können (d.h. ob sie „ähnlich“ sind). Denn dann haben sie die gleiche Jordan-Normalform, weil diese eindeutig bestimmt ist.
- Manchmal ist es hilfreich, eine Matrix erst in Jordan-Normalform zu überführen und dann damit zu rechnen.

Gesucht ist also die Jordan-Normalform J von A und eine Basiswechselmatrix B , so dass $B^{-1} \cdot A \cdot B = J$ ist (siehe 4.2.4). Ist $V = K^n$ und A eine Abbildungsmatrix bezüglich der Standardbasis, dann ist J die Abbildungsmatrix bezüglich der Basis B ; deshalb spricht man oft von der „Jordan-Basis“ und nicht von der „Jordan-Basiswechselmatrix“. Der Einfachheit halber sei deshalb von jetzt an $\Phi : K^n \rightarrow K^n, x \mapsto A \cdot x$ definiert.

Ist A diagonalisierbar, dann ist J die entsprechende Diagonalmatrix und B die Basis aus Eigenvektoren (siehe 4.6.2). Das ist z.B. der Fall, wenn das charakteristische Polynom p n verschiedene Nullstellen hat.

Falls p zwar in Linearfaktoren zerfällt, aber einige Faktoren gleich sind (also Terme der Form $(X - c)^k$ vorkommen), beruht die Existenz der Jordan-Normalform darauf, dass V die direkte Summe der „Haupträume“ $\text{Kern}(A - c \cdot E_n)^k$ ist. Das Produkt der Terme der Form $(A - c \cdot E_n)^k$ ist $p(A)$, also nach dem Satz von Cayley-Hamilton 0 (siehe 4.6.3). Jeder Vektor wird also 0 , wenn er nacheinander mit diesen Faktoren multipliziert wird. Es dürfte daher einleuchten, dass die direkte Summe der Kerne ganz V ist (auch wenn dies wieder kein Beweis ist). Außerdem ist die Dimension des Hauptraums k , also der Exponent im charakteristischen Polynom. Auch das ist nicht schwer einzusehen, denn die Summe der Dimensionen ist ja n .

Diese Haupträume sind Φ -invariant, d.h. für $x \in \text{Kern}(A - c \cdot E_n)^k$ gilt auch $\Phi(x) \in \text{Kern}(A - c \cdot E_n)^k$. Im Fall $k = 1$ ist es klar, denn dies ist dann der Eigenraum zum Eigenwert c . Aber auch sonst kann man es leicht nachrechnen: Sei $x \in \text{Kern}(A - c \cdot E_n)^k$, d.h. $(A - c \cdot E_n)^k \cdot x = 0$. Zu zeigen ist, dass $(A - c \cdot E_n)^k \cdot \Phi(x) = (A - c \cdot E_n)^k \cdot A \cdot x = 0$ ist. Aber $(A - c \cdot E_n)^k \cdot A$ ist die Matrix A eingesetzt in das Polynom $(X - c)^k \cdot X = X \cdot (X - c)^k$, also gleich $A \cdot (A - c \cdot E_n)^k$. Also gilt $(A - c \cdot E_n)^k \cdot A \cdot x = A \cdot (A - c \cdot E_n)^k \cdot x = A \cdot 0 = 0$. Man sieht hier sehr deutlich, dass es sich lohnt, den Zusammenhang von Matrizen und Polynomen genauer zu beleuchten. Theoretisch kann man es sich natürlich auch anders erklären: Multipliziert man nämlich $(A - c \cdot E_n)^k \cdot A$ aus, dann erhält man eine Linearkombination von Potenzen von A , bei denen man auf der linken Seite wieder A ausklammern kann.

Ist B nun eine Basis aus Vektoren der Haupträume, dann hat die Abbildungsmatrix von Φ bezüglich B eine Blockgestalt, in der zu jedem Hauptraum ein quadratischer Block („Jordan-Block“) auf der Diagonalen gehört. Denn ein Vektor, der in einem Hauptraum liegt, wird wieder in den Hauptraum abgebildet; Entsprechendes gilt für den Koordinatenvektor bezüglich B .

4.7.2 Berechnung

Zunächst zerfalle das charakteristische Polynom in Linearfaktoren, $p = (c_1 - X) \cdot (c_2 - X) \cdot \dots \cdot (c_n - X)$, die aber nicht unbedingt verschieden sein müssen. Man kann sich auf diesen Fall beschränken, weil z.B. in \mathbb{C} jedes Polynom in Linearfaktoren zerfällt. Es gibt aber auch noch eine verallgemeinerte Normalform, die ohne diese Beschränkung auskommt.

Dann kann man J und B wie folgt berechnen:

1. Auf der Diagonalen von J stehen die Eigenwerte c_1 bis c_n aus dem charakteristischen Polynom. Dabei müssen gleiche Werte nebeneinander stehen, um die im vorherigen Abschnitt angesprochene Blockgestalt zu erreichen.
2. Die Blöcke sollte man markieren. Die Matrix sieht dann ungefähr so aus:

$$\left(\begin{array}{ccc|ccc} c_1 & & & & & \\ & c_2 & & & & 0 \\ \hline & & c_3 & & & \\ & & & c_4 & & \\ & 0 & & & \ddots & \\ & & & & & c_n \end{array} \right)$$

(Dabei ist z.B. $c_1 = c_2$.) Die folgenden Schritte beziehen sich auf jeden einzelnen Block, und zwar zum Eigenwert c :

3. Unterhalb der Diagonalen steht jeweils entweder 0 oder 1, der Rest ist 0. Die Stellen, an denen eine Null steht, teilen den Jordan-Block in „Jordan-Kästchen“ auf:

$$\left(\begin{array}{ccc|ccc} c & & 0 & & & \\ \color{red}{1} & c & & & & 0 \\ 0 & \color{red}{1} & c & & & \\ \hline & & \color{red}{0} & c & & 0 \\ & 0 & & \color{red}{1} & \ddots & \\ & & & 0 & \color{red}{1} & c \end{array} \right)$$

4. Ganz rechts und überall dort, wo eine 0 steht, gibt es nur den Eigenwert auf der Diagonalen. Ein Koordinatenvektor, der nur an der entsprechenden Stelle den Wert 1 und sonst 0 hat, wird von J auf c mal sich selbst abgebildet. Also muss der entsprechende Basisvektor in B ein Eigenvektor zum Eigenwert c sein. Fazit: Die Anzahl der Kästchen ist die Dimension des Eigenraums zum Eigenwert c , also die Dimension des Kerns von $A - c \cdot E_n$:

$$\left(\begin{array}{ccc|ccc} c & & \color{red}{\downarrow} & & & \color{red}{\downarrow} \\ 1 & c & & & & \\ & \color{red}{1} & c & & & \\ \hline & & \color{red}{0} & c & & \\ & & & \color{red}{1} & \ddots & \\ & & & & \color{red}{1} & c \end{array} \right)$$

(Die Pfeile markieren die Spalten, zu denen in B ein Eigenvektor gehört.)

5. Nun sollte man sich überlegen, ob dies schon ausreicht, damit der Jordanblock eindeutig bestimmt ist. Hat der Eigenraum die Dimension 1, gibt es z.B. nur ein Kästchen. Ist die Dimension die Größe des Jordanblocks, gibt es nur Nullen außerhalb der Diagonalen. Oder wenn die Größe des Blocks z.B. 3 ist und die Dimension des Eigenraums 2, dann gibt es zwei Kästchen, also eines der Größe 2 und eines der Größe 1. Die Reihenfolge der Kästchen ist aber nicht eindeutig; sie kann höchstens per Konvention festgelegt werden (oft der Größe nach absteigend sortiert). Also ist man auch in diesem Fall fertig, wenn man B nicht bestimmen muss.

6. Ansonsten muss man wissen, dass in jedem Kästchen der Basisvektor ganz rechts ein Vektor aus $\text{Kern}(A - c \cdot E_n)$ ist (der Eigenvektor eben), der links daneben aus $\text{Kern}(A - c \cdot E_n)^2$, der nächste aus $\text{Kern}(A - c \cdot E_n)^3$, usw. (Das ergibt sich daraus, wie man im nächsten Schritt die Basisvektoren sucht.) D.h. man rechnet nacheinander diese Kerne aus und schaut, wie viele Vektoren man darin findet, die mit den vorherigen linear unabhängig sind. Die Anzahl ist die Differenz der Dimensionen der beiden Kerne. Die Vektoren verteilt man gedanklich auf die einzelnen Kästchen (am besten beginnend mit dem ersten, so dass die Kästchen der Größe nach absteigend sortiert sind), und damit weiß man (oft schon beim zweiten Exponenten), wie groß die Kästchen jeweils sein müssen.

Vorsicht: Dies verrät nur die *Anzahl* der Vektoren und damit die *Größe* der Kästchen. Möchte man B bestimmen, dann muss man noch Folgendes beachten:

7. Unterhalb der Diagonalen steht im Kästchen überall eine 1, d.h. für zwei benachbarte Basisvektoren b_1 und b_2 gilt $A \cdot b_1 = c \cdot b_1 + 1 \cdot b_2$, also $b_2 = A \cdot b_1 - c \cdot b_1$. Deswegen fängt man beim größten Jordankästchen an; es habe die Größe g . Man sucht einen beliebigen Vektor b_1 aus $\text{Kern}(A - c \cdot E_n)^g$, der nicht schon in $\text{Kern}(A - c \cdot E_n)^{g-1}$ liegt (am besten durch Raten und Überprüfen). Dies ist der Basisvektor, der zur linken Spalte des Jordankästchens gehört. Der nächste Basisvektor b_2 berechnet sich als $b_2 = A \cdot b_1 - c \cdot b_1 = (A - c \cdot E_n) \cdot b_1 \in \text{Kern}(A - c \cdot E_n)^{g-1}$, dann $b_3 = A \cdot b_2 - c \cdot b_2 \in \text{Kern}(A - c \cdot E_n)^{g-2}$, usw. bis b_g .

Das war das größte Jordankästchen; jetzt schaut man sich das nächstkleinere an. Hat es die Größe h , dann braucht man wieder einen Vektor, der in $\text{Kern}(A - c \cdot E_n)^h$, aber nicht in $\text{Kern}(A - c \cdot E_n)^{h-1}$ liegt. Allerdings darf er auch nicht linear abhängig mit den bereits bestimmten Vektoren sein, denn es soll ja eine Basis werden. (Am besten wieder raten und überprüfen. Alle Vektoren muss man nicht unbedingt überprüfen, sondern nur die, die selbst auch in $\text{Kern}(A - c \cdot E_n)^h$ liegen. Jedoch reicht es nicht, die Vektoren einzeln zu prüfen, sondern man muss wie immer ein LGS mit *allen* relevanten Vektoren aufstellen.)

Die Vektoren für die weiteren Kästchen findet man analog.

Anhand dieser Schritte kann man sowohl J als auch B bestimmen. Jedoch kommt man mit weniger Rechnung aus, wenn noch mehr Informationen gegeben sind:

Ist g die Größe des größten Jordan-Kästchens zum Eigenwert c , dann liegen alle Basisvektoren zum Jordan-Block in $\text{Kern}(A - c \cdot E_n)^g$. D.h. $\text{Kern}(A - c \cdot E_n)^g$ ist bereits der Hauptraum zum Eigenwert c . Man nennt g den „Index“ des Hauptraums. Wenn es mehrere Kästchen gibt (d.h. die Dimension des Eigenraums ist größer als 1), dann ist g offensichtlich kleiner als die Größe des Jordan-Blocks, also als der Exponent im charakteristischen Polynom.

Ersetzt man im charakteristischen Polynom p den Faktor $(X - c)^k$ durch $(X - c)^g$, dann ist immer noch $p(A) = 0$, denn $(A - c \cdot E_n)^g$ schickt genau die gleichen Vektoren auf 0 wie $(A - c \cdot E_n)^k$. Tut man dies bei allen Faktoren, dann erhält man das kleinste Polynom $p \neq 0$, so dass $p(A) = 0$ ist; dies nennt man das „Minimalpolynom“ von A . Das Minimalpolynom hat also die gleichen Faktoren wie das charakteristische Polynom, aber die Exponenten können kleiner sein (jedoch nie 0).

Ist das Minimalpolynom von A bereits bekannt, dann ist der Exponent dort also die Größe des größten Jordankästchens. Damit kann man oft die Jordan-Normalform ohne Rechnung schon genau angeben (wenn auch das charakteristische Polynom und eventuell die Dimensionen der Eigenräume gegeben sind).