Mathe-II-Skript

Markus Junker

9. Juni 2016

Wichtiger Hinweis

Dieses Skript ist kein Lehrbuch; es enthält mit Sicherheit Lücken und Fehler, die hoffentlich nach und nach korrigiert und ergänzt werden. Bitte teilen Sie mir Korrekturen oder Verständnisschwierigkeiten mit!

Genese

Das Skript basiert auf der in den Sommersemestern 2012 und 2013 gehaltenen Vorlesung „Mathematik II für Studierende der Informatik“. Im Sommersemester 2012 hat Lisa Schüttler eine Mitschrift angefertigt; auf der Grundlage dieser Mitschrift und meiner eigenen Notizen ist im Sommersemester 2013 ein unvollständiges Skript entstanden, das von David Zschocke ergänzt und in schöne Form gebracht wurde. Es wurde in den folgenden Semestern überarbeitet, zuletzt im Sommersemester 2016.
Beiden – Lisa Schüttler und David Zschocke – gilt mein herzlicher Dank!

„Plagiats-Disclaimer“

Das Skript ist nach in der Mathematik gängiger Vorgehensweise angefertigt. Dies bedeutet, dass es keinen Anspruch auf eine eigene wissenschaftliche Leistung erhebt und keine eigenen Ergebnisse wiedergibt, sondern die Ergebnisse anderer darstellt. Diese Ergebnisse sind über Jahrhunderte gewachsen; da Mathematik weitgehend ahistorisch betrieben wird, lässt sich in der Regel nicht mehr zurückverfolgen, von wem welche Fragestellungen, Begriffe, Sätze, Beweise oder Beweistechniken stammen. Vereinzelt gibt es überlieferte Zuweisungen von Sätzen oder von Beweisen zu Mathematikern (die aber nicht immer historisch exakt sein müssen).
Die Darstellung des Stoffes orientiert sich an den von mir selbst gehörten Vorlesungen, an Skripten von Kollegen und an Büchern. Diese verschiedenen Einflüsse sind nicht zu trennen und können daher nicht einzeln dargelegt werden. Fehler dagegen sind von mir zu verantworten. Insbesondere bei Formeln empfiehlt sich eine kritische Lektüre, da kleine Tippfehler aufgrund mangelnder Redundanz gleich massive Fehler bewirken.

Inhaltsverzeichnis

Teil I.

Lineare Algebra

1. Grundlegende algebraische Strukturen

1.1. Strukturen

Informelle Definition

Eine algebraische Struktur besteht aus einer nicht-leeren Grundmenge M mit einer oder mehreren *Operationen* (oder *Verknüpfungen*), die gewisse „schöne" Eigenschaften haben. Die Operationen können *innere Operationen* sein, das sind Funktionen/Abbildungen[1] $M^n \to M$, oder *äußere Operationen*, dies sind z. B. Abbildungen $R \times M \to M$ für eine feste Struktur R, etwa den Körper $\mathbb{R}$ der reellen Zahlen. Außerdem kann eine Struktur ausgezeichnete Elemente („*Konstanten*") besitzen.

Bei inneren Operation $\alpha : M^n \to M$ heißt n die *Stelligkeit* der Operation. Es ist also $\alpha : M \to M$ eine *einstellige* oder *unäre* Operation, $\alpha : M^2 \to M$ eine *zweistellige* oder *binäre* Operation; $\alpha : M^3 \to M$ eine *dreistellige* oder *ternäre* Operation, usw. Der mathematische Formalismus erlaubt es auch, *nullstellige* Operationen $\alpha : M^0 \to M$ zu betrachten, Da $M^0 = \{\emptyset\}$ eine einelementige Menge ist, kann man eine nullstellige Operation mit dem Bild dieses Elementes, also mit einer Konstanten identifizieren.

In den wichtigen mathematischen Strukturen werden in der Regel ein- und zweistellige Operationen sowie Konstanten betrachtet. Drei- und höherstellige Operationen, die nicht aus einfacheren Operationen zusammengesetzt sind, kommen selten vor.

Beispiele

1. Die Struktur $(\mathbb{Z}, +)$: Hier bilden die ganzen Zahlen $M = \mathbb{Z}$ die Grundmenge; die Addition „$+$" $: M \times M \to M$ ist darauf eine zweistellige innere Operation.
2. Die Struktur $(\mathbb{Z}, \cdot, 1)$: die ganzen Zahlen $M = \mathbb{Z}$ mit der Multiplikation „$\cdot$" $: M \times M \to M$ und der Konstanten 1 als ausgezeichnetem Element.
3. Die Struktur $(\mathbb{Z}, +, \cdot)$: Hier betrachtet man die ganzen Zahlen $\mathbb{Z}$ mit zwei zweistelligen Operationen (Addition und Multiplikation) gleichzeitig.
4. Die Menge der Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ als Grundmenge M mit der zweistelligen Operation „$\circ$", d. h. der Hintereinanderausführung von Funktionen, als zweistelliger innerer Operation.
5. Die Menge $M = A^*$ aller Wörter über einem Alphabet A. Wörter sind endliche Folgen von Symbolen. Eine zweistellige Verknüpfung auf A^* ist die *Konkatenation*, das Hintereinanderschreiben zweier Wörter.

[1] beide Begriffe benutzte ich synonym

Definition: Wichtige Eigenschaften von Operationen
Folgende wichtige Eigenschaften von zweistelligen Operationen $* : M^2 \to M$ und $\circ : M^2 \to M$ werden wir betrachten:

- $*$ heißt *kommutativ*, wenn für alle $m_1, m_2 \in M$ gilt: $m_1 * m_2 = m_2 * m_1$.
- $*$ heißt *assoziativ*, wenn für alle $m_1, m_2, m_3 \in M$ gilt: $m_1*(m_2*m_3) = (m_1*m_2)*m_3$.
- $*$ heißt *distributiv über* $\circ$, wenn für alle $m_1, m_2, m_3 \in M$ gilt: $m_1 * (m_2 \circ m_3) = (m_1 * m_2) \circ (m_1 * m_3)$ und $(m_2 \circ m_3) * m_1 = (m_2 * m_1) \circ (m_3 * m_1)$.
- $*$ besitzt ein *neutrales Element*, falls es ein $m_0 \in M$ gibt, so dass für alle $m \in M$ gilt: $m_0 * m = m * m_0 = m$.
- Falls $*$ ein neutrales Element $m_0 \in M$ besitzt, so heißt $m_2 \in M$ *inverses Element* von $m_1 \in M$ (bezüglich $*$), falls $m_1 * m_2 = m_2 * m_1 = m_0$.

Beispiele
Die zweistelligen Operationen in den Beispielen 1, 2 sind kommutativ, in 4 und 5 nicht; alle vier sind assoziativ. Im Beispiel 3 ist $\cdot$ distributiv über $+$; die Zahl 0 ist neutrales Element bezüglich der Addition und die Zahl 1 neutrales Element bezüglich der Multiplikation. Im Beispiel 4 ist die identische Abbildung $\mathrm{id}_{\mathbb{R}}$ neutrales Element bezüglich der Komposition; die Funktion $x \mapsto \frac{1}{2}x$ ist inverses Element der Funktion $x \mapsto 2x$.

Bemerkung:
Wichtige Strukturen sind Vektorräume (engl. *vector spaces*), Gruppen (*groups*), Ringe (*rings*) und Körper (*fields*). Diese werden nun in den weiteren Kapiteln Thema sein: Vektorräume vor allem in Teil I, die anderen Strukturen in Teil II der Vorlesung.

1.2. Monoide

Definition: Monoid
Ein *Monoid* besteht aus einer nicht-leeren Grundmenge M und einer assoziativen, zweistelligen Verknüpfung $\circ$ mit einem neutralen Element $e \in M$. Es gibt also eine Abbildung $\circ : M \times M \to M$, die

- *assoziativ* ist, d. h. $(m_1 \circ m_2) \circ m_3 = m_1 \circ (m_2 \circ m_3)$ für alle $m_1, m_2, m_3 \in M$ erfüllt,
- und ein *neutrales Element* $e \in M$ besitzt, d. h. es gilt $e \circ m = m \circ e = m$ für alle $m \in M$.

Ein Monoid $(M, \circ)$ heißt *kommutatives Monoid*, wenn die Verknüpfung $\circ$ zusätzlich

- *kommutativ* ist, d. h. $m_1 \circ m_2 = m_2 \circ m_1$ für alle $m_1, m_2 \in M$ gilt.

Erläuterung
„Monoid“ ist sächlich („das Monoid“) und wird „Mono-id“ mit Betonung auf der letzten Silbe ausgesprochen. Das Zeichen $\circ$ ist ein Platzhalter für die Verknüpfung; in einem kon-

kreten Monoid kann dafür auch ein anderes Zeichen stehen, etwa $+$ im Monoid $(\mathbb{N}, +, 0)$ der natürlichen Zahlen bezüglich der Addition.

Bemerkung:

Das neutrale Element e ist eindeutig bestimmt, d. h. es können nicht zwei oder mehrere neutrale Elemente für die gleiche Operation existieren. Denn falls e und e' neutrale Elemente sind, so gilt per Definition $e = e \circ e' = e'$.

Verschiedene Operationen haben dagegen in der Regel auch unterschiedliche neutrale Elemente. So sind die natürlichen Zahlen $\mathbb{N}$ sowohl bezüglich der Addition ein Monoid – mit neutralem Element 0 – als auch bezüglich der Multiplikation – mit neutralem Element 1.

Notation: Weglassen von Klammern

Wegen der Assoziativität kann man bei iterierten Verknüpfungen Klammern weglassen. „Iterierte Verknüpfung“ bedeutet, dass ein durch eine Verknüpfung gegebenes Element erneut verknüpft wird.

Im einfachsten Fall steht also $m_1 \circ m_2 \circ m_3$ für einen der beiden Ausdrücke $(m_1 \circ m_2) \circ m_3$ oder $m_1 \circ (m_2 \circ m_3)$, falls es nur auf das Ergebnis der Verknüpfung ankommt, da dann beide Ausdrücke das gleiche Ergebnis liefern.

Beispiele

- Die natürlichen Zahlen $\mathbb{N}$ bilden mit der Addition $+$ ein kommutatives Monoid mit neutralem Element 0.
- Die natürlichen Zahlen $\mathbb{N}$ bilden mit der Multiplikation $\cdot$ ein kommutatives Monoid mit neutralem Element 1.
- Die echt positiven natürlichen Zahlen $\mathbb{N} \setminus \{0\}$ bilden mit der Multiplikation $\cdot$ ein kommutatives Monoid mit neutralem Element 1.
- Die Abbildungen $\mathrm{Abb}(A, A)$ einer Menge A in sich selbst bilden unter der Komposition $\circ$, d. h. der Hintereinanderausführung von Abbildungen, ein Monoid, dessen neutrales Element die identische Abbildung id_A ist. Wenn A mindestens zwei Elemente $a \neq b$ besitzt, ist diese Monoid nicht kommutativ, wie man an den konstanten Abbildungen $x \mapsto a$ und $x \mapsto b$ sieht, die nicht miteinander vertauschen.
- Wenn A eine Menge ist (in diesem Kontext auch *Alphabet* genannt), bildet die Menge A^* der endlichen Folgen von Elementen aus A (die „*Wörter über A*“) mit der *Konkatenation* (d. h. dem Hintereinandersetzen) $^\frown$ ein Monoid. Mit $A = \{a, b, c\}$ ist also z. B. $abaac^\frown ccb = abaacccb$. Das neutrale Element ist das *leere Wort*, d. h. die Folge der Länge 0, das oft mit λ oder ε bezeichnet wird. Wenn A mindestens zwei Elemente enthält, ist A^* nicht kommutativ.

Gegenbeispiele

- Die echt positiven natürlichen Zahlen $\mathbb{N} \setminus \{0\}$ bilden mit der Addition $+$ kein Monoid, da es kein neutrales Element gibt.
- Die natürlichen Zahlen $\mathbb{N}$ bilden mit der Exponentiation kein Monoid, da die Exponentiation nicht assoziativ ist, denn z. B. ist $2^{(3^2)} = 2^9 = 512$, aber $(2^3)^2 = 8^2 = 64$.

Zudem gibt es zwar ein „rechtsneutrales Element“ (da $n^1 = n$ für alle $n \in \mathbb{N}$), aber kein „linksneutrales Element“.

Notation: Weglassen von Teilen der Definition

Wenn die Menge M mit der Verknüpfung $\circ$ und dem neutralem Element e ein Monoid bildet, schreibt man dafür üblicherweise $(M, \circ, e)$ oder $(M, \circ)$, da e durch $\circ$ festgelegt ist. Wenn man sauber arbeitet, unterscheidet man notationell zwischen der Struktur und der zugrundeliegenden Menge und schreibt dann gerne für die Struktur den entsprechenden Buchstaben in einem anderen Schriftart, also z. B. $\mathcal{M}$ oder $\mathfrak{M}$ für ein Monoid mit Grundmenge M. Oft erlaubt man sich aber die notationelle Unsauberkeit, für die Struktur und die Grundmenge das gleiche Symbol (hier z. B. M) zu verwenden.

Bei der Angabe eines Monoids entfällt bisweilen die Angabe der Verknüpfung, wenn aus dem Kontext heraus offensichtlich ist, welche gemeint ist, oder wenn es eine besonders natürliche Verknüpfung gibt. Wenn man z. B. vom Monoid der Wörter über einem Alphabet spricht oder dem Monoid der Abbildungen einer Menge in sich selbst, meint man die oben angegebenen Standardbeispiele. Das doppelte Beispiel der natürlichen Zahlen – einmal mit Addition und einmal mit Multiplikation – zeigt aber, dass man i. a. auf die Angabe der Verknüpfung nicht verzichten kann und selbst eine natürlich wirkende Operation nicht unbedingt einen Alleinstellungsanspruch hat.

Wenn mehrere (abstrakte) Monoide gleichzeitig betrachtet werden, werden oft die gleichen Notationen für die Verknüpfungen und neutralen Elemente gebraucht. Es kann also vorkommen, dass man Monoide $(M, \circ, e)$ und $(N, \circ, e)$ betrachtet. Zur Verdeutlichung schreibt man dann manchmal $\circ_M$ für Verknüpfung und e_M für das neutrale Element von M und analog $\circ_N$ und e_N für die Verknüpfung und das neutrale Element von N.

Analoge Bemerkungen zur Notation gelten für alle weiteren betrachteten algebraischen Strukturen!

1.3. Gruppen

Definition: Gruppe

Ein *Gruppe* besteht aus einer nicht-leeren Grundmenge G und einer zweistelligen Verknüpfung $\circ$ auf G (der „*Gruppenoperation*“), die

- *assoziativ* ist,
- ein *neutrales Element* $e \in G$ besitzt
- und bezüglich der es *inverse Elemente* gibt, d. h. zu jedem $g \in G$ gibt es ein Element $h \in G$ mit $h \circ g = g \circ h = e$.

Eine Gruppe $(G, \circ)$ heißt *kommutative Gruppe*[2], wenn die Verknüpfung $\circ$ zusätzlich *kommutativ* ist.

[2] oder auch *Abelsche Gruppe*, nach dem norwegischen Mathematiker Niels Henrik Abel (1802–1829)

Bemerkung:
Jede (kommutative) Gruppe ist also insbesondere ein (kommutatives) Monoid.

Bemerkung:
In einer Gruppe hat jedes Element g ein eindeutig bestimmtes inverses Element, denn sind h_1, h_2 invers zu g, so gilt

$$h_1 = h_1 \circ e = h_1 \circ (g \circ h_2) = (h_1 \circ g) \circ h_2 = e \circ h_2 = h_2.$$

Notation: inverses Element
Das bezüglich der Gruppenoperation zu $g \in G$ inverse Element wird mit g^{-1} bezeichnet.

Notation: gebräuchliche Notationen für Gruppen
Es gibt drei gebräuchliche Notationen für Gruppen:

	Verknüpfung	neutrales Element	inverses Element
allgemein:	$\circ$	e	g^{-1}
multiplikativ:	$\cdot$	1	g^{-1}
additiv:	$+$	0	$-g$

Die additive Schreibweise ist im allgemeinen kommutativen Gruppen vorbehalten. Bei der multiplikativen Schreibweise lässt man den Multiplikationspunkt auch gerne weg.

Beispiele

- $(\mathbb{Z}, +, 0)$ ist kommutative Gruppe.
- $(\mathbb{Q}, +, 0)$, $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$ und $(\mathbb{Q}^{>0}, \cdot, 1)$ mit $\mathbb{Q}^{>0} = \{q \in \mathbb{Q} \mid q > 0\}$ sind kommutative Gruppen.
- $(\mathbb{R}, +, 0)$, $(\mathbb{R} \setminus \{0\}, \cdot, 1)$ und $(\mathbb{R}^{>0}, \cdot, 1)$ mit $\mathbb{R}^{>0} = \{r \in \mathbb{R} \mid r > 0\}$ sind kommutative Gruppen.
- $(\mathbb{C}, +, 0)$ und $(\mathbb{C} \setminus \{0\}, \cdot, 1)$ sind kommutative Gruppen.
- Ein wichtiges Beispiel einer Gruppe ist die „verallgemeinerte Uhren-Arithmetik", d. i. die kommutative Gruppe $\mathbb{Z}_m = \big(\{0, \dots, m-1\}, +_m, 0\big)$, wobei

  $$x +_m y := \text{„Rest von } x + y \text{ bei Division durch } m\text{“} = \begin{cases} x + y & \text{falls } x + y < m \\ x + y - m & \text{falls } x + y \geq m \end{cases}$$

  Für $n = 12$ ist dies die Art, wie man mit Uhrzeiten rechnet („*8 Uhr + 5 Stunden = 1 Uhr*“).
- $(\mathrm{Sym}(A), \circ, \mathrm{id})$ ist eine Gruppe, die *symmetrische Gruppe über A*. Hierbei bezeichnet $\mathrm{Sym}(A)$ die Menge der *Permutationen* von A, d. h. der Bijektionen von einer Menge A in sich selbst, und $\circ$ ist die Komposition von Abbildungen. Wenn A mindestens drei Elemente enthält, ist die symmetrische Gruppe über A nicht kommutativ.

- Die *triviale Gruppe* besteht nur aus einem Element, ihrem neutralen Element. Genau genommen gibt es viele verschiedene Realisierungen der trivialen Gruppe: Zum Beispiel besteht $(\mathbb{Z}_1, +_1)$ nur aus einem Element und auch $\mathrm{Sym}(A)$ für eine einelementige Menge A. Alle diese Realisierungen sind aber untereinander *isomorph*, d. h. (informell) nur verschiedene Bezeichnungen für dieselbe Gruppe. Die mathematische Präzisierung der „Isomorphie" folgt in Teil II.
- Zu jeder Struktur M gibt es die *Automorphismengruppe* $\mathrm{Aut}(M)$, welche aus den „strukturerhaltenden" Permutationen von M besteht mit der Komposition von Funktionen als Gruppenoperation. Was genau „strukturerhaltend" bedeutet, wird noch an Beispielen klar werden.

Gegenbeispiele

Die folgenden Strukturen sind keine Gruppen:

- $(\mathbb{Z} \setminus \{0\}, \cdot, 1)$: Alle Elemente bis auf 1 und -1 haben keine Inverse.
- $(\mathbb{Q}, \cdot, 1)$: 0 hat kein Inverses.

Definition: Gruppentafel

Die Gruppentafel ist eine Tabelle, in der alle möglichen Verknüpfungen zweier Elemente der Gruppe aufgeführt sind. Eine Gruppe ist kommutativ, wenn die Gruppentafel mit der Diagonale von links oben nach rechts unten eine Symmetrieachse besitzt. Bei nichtkommutativen Gruppen muss man klarstellen, in welcher Reihenfolge die Verknüpfung in der Tabelle naufzufassen ist.

Beispiele

- Zu $\mathbb{Z}_4$ ist die Gruppentafel:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

- Allgemeiner kann man natürlich für jede zweistellige Verküpfung solch eine Verknüpfungstafel aufstellen. Wenn man etwas das Monoid $\mathrm{Abb}(A, A)$ für die zweielementige Menge $A = \{a, b\}$ betrachtet, so besteht $\mathrm{Abb}(A, A)$ aus den folgenden vier Abbildungen: $\mathrm{id}_A : x \mapsto x$, $c_a : x \mapsto a$, $c_b : x \mapsto b$ und $\tau : a \mapsto b, b \mapsto a$. Hierfür ist die Verküpfungstafel:

$\circ$	id_A	c_a	c_b	τ
id_A	id_A	c_a	c_b	τ
c_a	c_a	c_a	c_a	c_a
c_b	c_b	c_b	c_b	c_b
τ	τ	c_b	c_a	id_A

mit der Konvention, dass in der Tafel $f \circ g$ dargestellt ist, wobei f in der ersten Spalte und g in der obersten Zeile angegeben ist. Dieses Monoid ist nicht kommutativ, was man an der fehlenden Symmetrie der Verküpfungstafel sieht. Daher ist es wichtig anzugeben, in welcher Reihenfolge die Verküpfung aufzufassen ist.

- Die kleinste nicht-kommutative Gruppe ist $\mathrm{Sym}(B)$ für eine drei-elementige Menge B. diese Gruppe hat sechs Elemente. Als Übung kann man die Gruppentafel aufstellen.

Bemerkung:
Man sieht bei genauerem Hinschauen, dass manche der in den Beispielen angegebenen Gruppen von einfachen Beispielen für Monoide herstammen. Diese Monoide wurden so verändert, dass sie auch die Anforderungen an Gruppen erfüllen. Man kann zum einen versuchen, fehlende inverse Elemente hinzuzunehmen (Beispiel: Konstruktion von $(\mathbb{Z}, +)$ aus $(\mathbb{N}, +)$). Dies ist aber nicht immer möglich. Manchmal genügt es dann, wenige störende Elemente wegzulassen (Beispiel: Konstruktion von $(\mathbb{Q}^{>0}, \cdot)$ aus $(\mathbb{N}, \cdot)$ unter Weglassen der Null). Zum andern erhält man manchmal aus Monoiden interessante Gruppen, indem man die Elemente herausgreift, die bereits Inverse haben (Beispiel: $\mathrm{Sym}(A)$ in $\mathrm{Abb}(A, A)$).

Zur Zahl 0 in $(\mathbb{N}, \cdot)$ kann man kein inverses Element hinzunehmen, ohne die Assoziativität aufzugeben. Denn gäbe es in einer Erweiterung ein Element 0^{-1}, müsste z. B.

$$1 = 0 \cdot 0^{-1} = (2 \cdot 0) \cdot 0^{-1} = 2 \cdot (0 \cdot 0^{-1}) = 2 \cdot 1 = 2$$

gelten.

Ähnlich sieht man bei Abbildungen, dass es kein (Links-)Inverses für h geben kann, wenn $h \circ g_1 = h \circ g_2$ für $g_1 \neq g_2$ gilt, und kein (Rechts-)Inverses, wenn $g_1 \circ h = g_2 \circ h$ gilt.

1.4. Ringe

Definition: Ring
Ein *Ring* besteht aus einer nicht-leeren Menge R, zwei zweistelligen Verknüpfungen $+$ und $\cdot$ auf R (in der Regel Addition und Multiplikation genannt) und Elementen 0 und 1 (in der Regel Null und Eins genannt), für die gilt:

- $(R, +, 0)$ ist eine kommutative Gruppe;
- $(R, \cdot, 1)$ ist ein Monoid;
- $\cdot$ ist distributiv über $+$, d. h. es gelten die *Distributivgesetze*:

$$\begin{aligned} (r_1 + r_2) \cdot s &= (r_1 \cdot s) + (r_2 \cdot s) \\ s \cdot (r_1 + r_2) &= (s \cdot r_1) + (s \cdot r_2) \end{aligned}$$

  für alle $r_1, r_2, s \in R$.

Ein Ring $(R, +, \cdot)$ heißt *kommutativer Ring*, wenn die Multiplikation zusätzlich *kommutativ* ist.

Erläuterung

Genauer handelt es sich hier um *Ringe mit Eins* oder *unitäre Ringe*. Es gibt ein allgemeineres Konzept von „Ring“, bei dem es kein neutrales Element der Multiplikation zu geben braucht. Bei der Lektüre anderer Skripte oder Bücher muss man daher vorsichtig sein, da eine andere Definition benutzt sein könnte.

In einem kommutativen Ring folgt natürlich jedes der beiden Distributivgesetze aus dem anderen.

Notation: Weglassen von Klammern

Zur Ersparnis von Klammern führt man die üblichen „Vorfahrtsregeln“ ein, also „Punkt vor Strich“. Den Multiplikationspunkt lässt man gerne weg. Das erste Distributivgesetz kann man also kurz als $(r_1 + r_2)s = r_1 s + r_2 s$ schreiben.

Bemerkung: Vertraute Rechenregeln

Aus den Axiomen für Ringe ergibt sich, dass $r \cdot 0 = 0 \cdot r = 0$ für alle $r \in R$ ist. Denn es gilt $r \cdot 0 = r \cdot (0 + 0) = r \cdot 0 + r \cdot 0$. Also ist

$$0 = r \cdot 0 + (-(r \cdot 0)) = r \cdot 0 + r \cdot 0 + (-(r \cdot 0)) = r \cdot 0 + 0 = r \cdot 0,$$

und analog für die vertauschte Reihenfolge.

Ähnlich sieht man, dass $(-r) \cdot s = r \cdot (-s) = -(r \cdot s)$ für alle $r, s \in R$ gilt. Auch hier kann man daher Klammern einsparen.

Vorsicht: Nicht alle aus dem Ring der ganzen Zahlen vertrauten Rechenregeln gelten in beliebigen Ringen. Zum Beispiel gilt im Ring $\mathbb{Z}_6$ (siehe in den folgenden Beispielen) $2 \cdot_6 3 = 0$, ohne dass $2 = 0$ oder $3 = 0$ gelten würde.

Beispiele

- Die Definition verbietet nicht, dass $0 = 1$ ist. In diesem Fall folgt aber $r = r \cdot 1 = r \cdot 0 = 0$ für alle $r \in R$, und es liegt der sogenannte *triviale Ring* vor, der nur aus einem einzigen Element besteht.
- $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ – jeweils mit der üblichen Addition und Multiplikation – sind kommutative Ringe.
- Die Gruppe $\mathbb{Z}_m$ (siehe Beispiele zu 1.3) kann durch eine analog definierte Multiplikation $\cdot_m$ zu einem kommutativen Ring gemacht werden: $x \cdot_m y$ rechnet man dadurch aus, dass man von dem normalen Produkt in $\mathbb{Z}$ den Rest bei der Division durch m nimmt, also solange m abzieht, bis man im Bereich $\{0, \ldots, m-1\}$ landet.
- Die Polynome mit Koeffizienten in einem Ring R und der Unbekannten X bilden mit der bekannten Polynomaddition und -multiplikation den *Polynomring* $R[X]$, also z. B. $\mathbb{R}[X]$: Polynome mit einer Unbekannten X und Koeffizienten in $\mathbb{R}$, oder $\mathbb{Z}[X]$: Polynome mit einer Unbekannten X und Koeffizienten in $\mathbb{Z}$.

  Nimmt man mit einer neuen Unbekannten Y z. B. den Polynomring $\mathbb{R}[X]$ als Koeffizientenbereich, erhält man den Polynomring mit zwei Unbekannten X und Y mit Koeffizienten in $\mathbb{R}$, also $\mathbb{R}[X][Y] = \mathbb{R}[X, Y]$.

1.5. Körper

Definition: Körper

Ein *Körper* besteht aus einer nicht-leeren Menge K, zwei zweistelligen Verknüpfungen $+$ und $\cdot$ auf K (Addition und Multiplikation) und Elementen 0 und 1 (Null und Eins), für die gilt:

- $0 \neq 1$;
- $(K, +, 0)$ und $(K \setminus \{0\}, \cdot, 1)$ sind kommutative Gruppen[3];
- es gelten die Distributivgesetze wie bei Ringen.

Erläuterung

Mit der gleichen Rechnung wie bei Ringen zeigt man, dass $0 \cdot k = 0$ für alle $k \in K$ ist. Damit sieht man, dass die Multiplikation auf ganz K assoziativ ist und 1 als neutrales Element hat, d. h. dass $(K, \cdot, 1)$ ein kommutatives Monoid ist. Jeder Körper ist also insbesondere ein kommutativer, nicht-trivialer Ring:

Beispiele

- $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ mit der üblichen Addition und Multiplikation sind Körper.
- Für Primzahlen p ist $\mathbb{Z}_p$ mit den definierten Operationen $+_m$ und $\cdot_m$ ein Körper und wird dann oft mit $\mathbb{F}_p$ bezeichnet.
- $\mathbb{R}(x)$ ist der Körper der rationalen Funktionen über $\mathbb{R}$,

$$\mathbb{R}(x) = \left\{ \frac{P(x)}{Q(x)} \,\middle|\, P, Q \in \mathbb{R}[x], Q \neq 0 \right\}.$$

Definition: $\mathbb{F}_2$

Besonders interessant für die Informatik ist der Körper $\mathbb{F}_2$, der aus den beiden Elementen 0 und 1 besteht mit folgenden Verknüpfungen:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

[3] Es gibt auch das allgemeineres Konzept eines *Schiefkörper*, bei dem die Multiplikation nicht kommutativ zu sein braucht.

1.6. Exkurs: Äquivalenzrelation

Definition: binäre Relationen

Sei M eine Menge. Eine *zweistellige Relation* (oder *binäre Relation*) R auf M ist eine Eigenschaft von Paaren von Elementen von M. Sie kann mit der Teilmenge der Paare von $M \times M$ identifiziert werden, auf die die Eigenschaft zutrifft.

Für $a, b \in M$ schreibt man aRb (oder auch Rab), wenn R auf (a, b) zutrifft.

Beispiele

- Auf $M = \mathbb{N}$ sind die Ordnungsrelationen $<$, $\leq$, $>$ und $\geq$ vier Beispiele binärer Relationen. Zum Beispiel gilt $2 < 3$, d. h. die durch $<$ ausgedrückte Eigenschaft „kleiner als“ trifft auf das Paar $(2, 3)$ zu, während $2 < 2$ nicht gilt, d. h. die Kleiner-Eigenschaft, trifft auf das Paar $(2, 2)$ nicht zu. Man kann die Kleiner-Relation durch die (manchmal *Graph der Relation* genannte) Menge $\{(a, b) \in \mathbb{N} \times \mathbb{N} \mid a < b\}$ beschreiben.
- Ein weiteres Beispiel einer binären Relation auf $\mathbb{N}$ ist die *Teilbarkeitsrelation*, die mit einem senktrechten Strich $\mid$ bezeichnet wird: $a \mid b$ ist genau dann wahr, wenn die Zahl a die Zahl b ohne Rest teilt. Es gilt also zum Beispiel $3 \mid 15$, aber nicht $3 \mid 14$. dafür schreibt man $3 \nmid 14$.
- Eine besondere Relation ist die *Gleichheitsrelation* $=$, die genau auf die Paare zutrifft, deren beiden Komponenten gleich sind. Zu beachten ist hier, dass links und rechts des Gleichheitszeichens in der Regel nur Namen für Elemente stehen (z. B. Rechenausdrücke) und nicht die Elemente selbst. So gilt z. B. in den natürlichen Zahlen $3 + 5 = 8$, weil darin sowohl „$3 + 5$“ als auch „8“ Bezeichnungen desselben Elements sind. Ist man dagegen in $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +\}^*$, so sind „$3 + 5$“ und „8“ verschiedene Wörter über der gegebenen Symbolmenge.

Definition: Eigenschaften binärer Relationen

Sei R eine binäre Relation.

- R heißt „reflexiv“, falls Rmm für alle $m \in M$ gilt.
- R heißt „symmetrisch“, falls für alle $m_1, m_2 \in M$ gilt: $Rm_1m_2 \Leftrightarrow Rm_2m_1$.
- R heißt „transitiv“, falls für alle $m_1, m_2, m_3 \in M$ gilt: wenn Rm_1m_2 und Rm_2m_3, dann auch Rm_1m_3.

Beispiele

Von den oben betrachteten Relationen auf $\mathbb{N}$ sind $=, \leq, \geq$ und $\mid$ reflexiv, $<$ und $>$ sind nicht reflexiv. Abgesehen von $=$ ist keine der Relationen symmetrisch. Alle betrachteten Relationen sind transitiv.

Definition: Äquivalenzrelation und Äquivalenzklassen

Eine Äquivalenzrelation $\sim$ auf M ist eine reflexive, symmetrische und transitive binäre Relationen auf M. Die Äquivalenzklasse von $m \in M$ bzgl. $\sim$ ist $m/_\sim := \{n \in M \mid m \sim n\}$.[4]

Erläuterung

Für Äquivalenzklassen gibt es keine Standardnotation. Andere verbreitete Schreibweisen sind $[m]_\sim$, $[\![m]\!]_\sim$ oder auch kurz $[m]$, $[\![m]\!]$ oder $\overline{m}$, falls aus dem Kontext klar ist, um welche Relation es sich handelt.

Bemerkung:

Die Äquivalenzklassen bilden eine Partition von M, d. h.

- $\bigcup_{m\in M} m/_\sim = M$;
- zwei verschiedene Äquivalenzklassen sind disjunkt.

Die Äquivalenzklassen von Elementen m_1, m_2 sind also entweder gleich (nämlich genau dann, wenn $m_1 \sim m_2$) oder disjunkt (wenn $m_1 \not\sim m_2$).

Umgekehrt liefert jede Partition von M eine Äquivalenzrelation, deren Äquivalenzklassen gerade die Teilmengen der Partition sind: Zwei Elemente sind genau dann äquivalent, wenn sie in derselben Teilmenge der Partition liegen.

Definition: Repräsentant, Repräsentantensystem

Falls $K \subseteq M$ eine Äquivalenzklasse ist und $m \in K$, dann heißt m *Vertreter* (oder *Repräsentant*) der Klasse. Ein *Vertreter*- oder *Repräsentantensystem* von $\sim$ ist eine Teilmenge von M, die aus jeder Äquivalenzklasse genau einen Vertreter enthält.

Erläuterung

Ein in der Mathematik sehr häufiges Verfahren besteht darin, Äquivalenzklassen als neue mathematische Objekte einzuführen. Darin kann man einen Abstraktionsprozess sehen: Die Äquivalenzrelation drückt eine gemeinsame Eigenschaft aus; die Äquivalenzklasse steht für das jeweils Gemeinsame. Als nicht-mathematisches Beispiel könnte man sich eine Menge von Gegenständen vorstellen, auf denen man die Äquivalenzrelationen „gleiche Form" oder „gleiche Farbe" betrachtet. Die Äquivalenzklassen entsprechen dann den Formen bzw. Farben, für die man u. U. (noch) keine Namen hat. Mathematisch gesprochen könnte man dann die Äquivalenzklassen als die Formen bzw. Farben definieren.

Im mathematischen Kontext kommt es häufig vor, dass man die Menge der Äquivalenzklassen selbst wieder als eine Struktur auffassen möchte und darauf Operationen definieren will. Dies geschieht in der Regel dadurch, dass man die Operationen auf Vertretern

[4] Achtung: Für Äquivalenzklassen gibt es keine Standardnotation. Andere Schreibweisen sind $[m]_\sim$, $[\![m]\!]_\sim$, $[m]$, $[\![m]\!]$, $\bar{m}$, wobei die letzten Notationen voraussetzen, dass die Äquivalenzrelation aus dem zusammenhang bekannt ist.

der Äquivalenzklassen definiert, und zwar entweder auf einem ausgewählten Vertretersystem oder auf beliebigen Vertretern. In letzterem Fall muss man zeigen, dass die Definition *vertreterunabhängig* („*wohldefiniert*“) ist, d. h. nicht von der Wahl der Vertreter abhängt. Ein bekanntes Beispiel soll dies verdeutlichen:

Beispiele

Brüche, d. h. die rationalen Zahlen $\mathbb{Q}$, werden als Äquivalenzklassen von Paaren ganzer Zahlen eingeführt. Genauer betrachtet man auf der Menge $M = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ die Äquivalenzrelation

$$(m_1, n_1) \sim (m_2, n_2) :\Longleftrightarrow\ m_1 \cdot n_2 = m_2 \cdot n_1.$$

Die Äquivalenzklasse von (m, n) entspricht dabei dem Bruch $\frac{m}{n}$. Ein Beispiel für ein Vertretersystem ist $\{(m, n) \mid n > 0, m$ und n teilerfremd$\}$, was der gekürzten Darstellung von Brüchen mit positivem Nenner entspricht.

Wenn man nun die Addition von Brüchen definieren will, kann man das auf diesem Vertretersystem tun durch

$$(m, n) + (m', n') := \left(\frac{mn' + m'n}{\mathrm{ggT}(mn' + m'n, nn')}, \frac{nn'}{\mathrm{ggT}(mn' + m'n, nn')} \right)$$

(wobei „ggT“ für den positiven größten gemeinsamen Teiler steht) oder auf beliebigen Repräsentanten durch

$$(m, n) + (m', n') := (mn' + m'n, nn').$$

Letzteres ist als Definition viel einfacher, aber überhaupt nur sinnvoll, wenn das Ergebnis nicht von der Wahl der Repräsentanten abhängt. Dies bedeutet: Falls $(m_1, n_1) \sim (m_2, n_2)$ und $(m'_1, n'_1) \sim (m'_2, n'_2)$, dann muss $(m_1, n_1) + (m'_1, n'_1) \sim (m_2, n_2) + (m'_2, n'_2)$ gelten.

Man kann nun nachrechnen, dass dies stimmt! Denn nach Voraussetzung ist $m_1 n_2 = m_2 n_1$ und $m'_1 n'_2 = m'_2 n'_1$. Also ist

$$\begin{aligned}(m_1 n'_1 + m'_1 n_1) \cdot n_2 n'_2 &= m_1 n'_1 n_2 n'_2 + m'_1 n_1 n_2 n'_2 \\ &= m_2 n'_2 n_1 n'_1 + m'_2 n_2 n_1 n'_1 = (m_2 n'_2 + m'_2 n_2) \cdot n_1 n'_1\end{aligned}$$

2. Vektorräume

2.1. Vektorräume

Sei K ein Körper, also z. B. $K = \mathbb{R}$ oder $K = \mathbb{F}_2$ (dies werden die hauptsächlichen Beispiele in dieser Vorlesung sein). Zur Verdeutlichung sind die Körperelemente und -operationen vorübergehend mit einem Index K gekennzeichnet, also $+_K, -_K, \cdot_K, 0_K, 1_K$.

Definition: Vektorraum

Ein K-*Vektorraum* V besteht aus einer nicht-leeren Menge V zusammen mit einer zweistelligen inneren Verknüpfung $+ : V \times V \to V$ (der Addition) und einer äußeren Verknüpfung $\cdot : K \times V \to V$ (der *Skalarmultiplikation*), für die gilt:

- $(V, +)$ ist eine kommutative Gruppe mit neutralem Element 0_V;
- es gelten folgende Regeln für die Skalarmultiplikation:

$$\begin{aligned} k \cdot (v_1 + v_2) &= (k \cdot v_1) + (k \cdot v_2) \\ (k_1 +_K k_2) \cdot v &= (k_1 \cdot v) + (k_2 \cdot v) \\ (k_1 \cdot_K k_2) \cdot v &= k_1 \cdot (k_2 \cdot v) \\ 1_K \cdot v &= v \end{aligned}$$

  für alle $k, k_1, k_2 \in K$ und $v, v_1, v_2 \in V$.

Falls aus dem Kontext klar ist, um welchen Körper K es geht, spricht man auch kurz von „Vektorraum“ statt von „K-Vektorraum“. Elemente von V heißen *Vektoren*, Elemente von K *Skalare*.

Bemerkung:

Im Unterschied zu einem Ring kann man Vektoren in einem allgemeinen Vektorraum nicht miteinander multiplizieren.[1]. Manche Rechenregeln gelten aber wie in Ringen und lassen sich analog beweisen, so gilt für alle $k \in K$ und $v \in V$:

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

Hier steht der Klarheit halber $-_K k$ für das additive Inverse von k im Körper K und $-_V v$ für das additive Inverse von v im Vektorraum V.

[1] In speziellen Fällen gibt es allerdings Vektorprodukte

Notation:
In Vektorräumen benutzt man die gleichen notationellen Kurzformen wie bei Ringen (Klammersparregeln und Weglassen des Multiplikationspunktes). Auch werde ich von nun an die Indizes K und V in der Regel weglassen. Dadurch bekommen 0, $+$, $-$ und $\cdot$ zwar eine doppelte Bedeutung; es sollte aber aus der Situation immer klar werden, welche Null gemeint ist bzw. in welcher Struktur gerade gerechnet wird. Eine Skalarmultiplikation liegt immer dann vor, wenn links ein Körperelement und rechts ein Vektor steht. Wenn auf beiden Seiten ein Körperelement steht, handelt es sich um die Multiplikation im Körper. Die Addition kann nur zwischen zwei Vektoren oder zwischen zwei Körperelementen stehen.

Beispiele

- $\mathbb{R}^n$, also die Menge der n-Tupel reeller Zahlen, ist ein $\mathbb{R}$-Vektorraum mit komponentenweiser Addition und Skalarmultiplikation. Die Tupel können als z. B. als *Zeilenvektoren* $(r_1, \dots, r_n)$ geschrieben werden. Dann ist also

$$\begin{aligned}(r_1, \dots, r_n) + (s_1, \dots, s_n) &= (r_1 + s_1, \dots, r_n + s_n)\\ r \cdot (r_1, \dots, r_n) &= (r \cdot r_1, \dots, r \cdot r_n)\end{aligned}$$

- Spezialfälle hiervon:
  Für $n = 2$ erhält man die *koordinatisierte reelle Ebene*: Wenn man zwei verschiedene Koordinatenachsen in der Ebene wählt, kann man jeden Punkt der Ebene mit dem Paar (x, y) seiner Koordinaten identifizieren.
  Für $n = 3$ erhält man analog den *koordinatisierten reellen Raum*: Die Wahl dreier nicht in einer Ebene liegender Koordinatenachsen erlaubt es, jeden Punkt des Raumes mit dem Tripel (x, y, z) seiner Koordinaten identifizieren.
  Für $n = 1$ erhält man die *koordinatisierte reelle Gerade*: Die Wahl des Koordinatensystems reduziert sich in diesem Fall auf die Wahl des Ursprungs und des Maßstabes.
  Ein Element von $\mathbb{R}^1$, also ein 1-Tupel (r) mit $r \in \mathbb{R}$, kann man mit der reellen Zahl r identifizieren.[2] in diesem Fall sind also Vektorraum und Skalarenkörper gleich.
  Für $n = 0$ erhält man den einelementigen Vektorraum $\mathbb{R}^0 = \{0\}$.
- Allgemeiner kann man Folgen reeller Zahlen betrachten, also den $\mathbb{R}$-Vektorraum $\mathbb{R}^\infty := \big\{(r_0, r_1, r_2, \dots) \;\big|\; r_i \in \mathbb{R}\big\}$, ebenfalls mit komponentenweisen Operationen, also

$$\begin{aligned}(r_0, r_1, r_2, \dots) + (s_0, s_1, s_2, \dots) &= (r_0 + s_0, r_1 + s_1, r_2 + s_2, \dots)\\ r \cdot (r_0, r_1, r_2, \dots) &= (r \cdot r_0, r \cdot r_1, r \cdot r_2, \dots)\end{aligned}$$

[2] Man kann n-Tupel auf verschiedene Weise definieren, z. B. n-Tupel über $\mathbb{R}$ als Funktionen $\{1, \dots, n\} \to \mathbb{R}$. In diesem Fall haben Elemente von $\mathbb{R}^1$ formal einen anderen Typ als Elemente von $\mathbb{R}$ und das Weglassen der Klammer von (r) nach r steht tatsächlich für eine Identifikation. Bei anderen Definitionen ist u. U. $\mathbb{R}^1$ tatsächlich gleich $\mathbb{R}$; dann sind (r) und r nur zwei Notationen für dasselbe Element.

- Die Polynome mit Koeffizienten aus $\mathbb{R}$ bilden ebenfalls einen $\mathbb{R}$-Vektorraum mit der üblichen Addition und der Skalarmultiplikation $r \cdot \sum_{i=1}^{n} r_i X^i = \sum_{i=1}^{n} (r \cdot r_i) X^i$. Wenn man Skalare mit konstanten Polynomen identifiziert, ist dies gewissermaßen ein Teil der Ringstruktur auf $\mathbb{R}[X]$.
- All die bisherigen Beispiele funktionieren für beliebige Körper, d. h. für jeden Körper K erhält man K-Vektorräume K^n, K^∞, $K[X]$.
- Da $\mathbb{R}$ ein Teilkörper von $\mathbb{C}$ ist, kann man jeden $\mathbb{C}$-Vektorraum auch als $\mathbb{R}$-Vektorraum betrachten, indem man die Skalarmultiplikation auf reelle Skalare einschränkt. Insbesondere ist $\mathbb{C}$ selbst sowohl $\mathbb{C}$-Vektorraum als auch $\mathbb{R}$-Vektorraum. Als $\mathbb{R}$-Vektorraum kann man ihn mit $\mathbb{R}^2$ identifizieren („Gaußsche Zahlenebene“).
- $\mathbb{R}$ ist dagegen *kein* $\mathbb{F}_2$-Vektorraum. $\mathbb{R}$ enthält zwar ebenfalls Elemente 0 und 1 wie $\mathbb{F}_2$; diese verhalten sich aber in $\mathbb{F}_2$ anders als in $\mathbb{R}$ (d. h. $\mathbb{F}_2$ ist kein Teil- oder Unterkörper von $\mathbb{R}$), denn $1_{\mathbb{F}_2} +_{\mathbb{F}_2} 1_{\mathbb{F}_2} = 0_{\mathbb{F}_2}$, aber $1_{\mathbb{R}} +_{\mathbb{R}} 1_{\mathbb{R}} \neq 0_{\mathbb{R}}$.
  So gilt z. B. $2\sqrt{2} = (1 \cdot \sqrt{2}) +_{\mathbb{R}} (1 \cdot \sqrt{2}) \neq (1 +_{\mathbb{F}_2} 1) \cdot \sqrt{2} = 0 \cdot \sqrt{2} = 0$.
- Die aus der Schule als „Pfeile in der Ebene“ (oder analog im Raum) betrachteten Vektoren kann man auf mehrere Weisen in den Begriff des Vektorraums einsortieren.
  1. Man betrachtet Pfeile als orientierte Geradenstücke in der Ebene und definiert darauf die Äquivalenzrelation der „Parallelität“: Zwei Pfeile sind parallel, falls sie gleiche Länge und Richtung (inklusive Orientierung) haben, also durch eine Parallelverschiebung der Ebene ineinander übergehen. Vektoren sind nun Parallelitätklassen von Pfeilen: Die Skalarmultiplikation eines Pfeiles mit einer reellen Zahl r besteht dann aus der Streckung um das r-fache (de facto eine Stauchung, falls $|r| < 1$, und orientierungsumkehrend, falls $r < 0$); die Addition durch „Dreiecksbildung“: man wählt einen Repräsentanten v_0 aus der Klasse von v, den Repräsentanten w_0 aus der Klasse von w, dessen Anfangspunkt der Endpunkt von v_0 ist, und setzt für $v + w$ die Äquivalenzklasse des Pfeils vom Anfangspunkt von v_0 zum Endpunkt von w_0. Natürlich muss man dann zeigen, dass diese Operationen repräsentantenunabhängig sind.
  2. Man wählt ein Repräsentantensystem der Äquivalenzklasse der Pfeile, nämlich diejenigen, welche von einem festgewählten Ursprung ausgehen. Die Streckung bei der Skalarmultiplikation geht dann immer vom Ursprung aus; bei der Addition muss man beide Pfeilen zu einem Parallelogramm ergänzen und die vom Ursprung ausgehende Diagonale wählen (man muss dies passend interpretieren, falls beide Vektoren in die gleiche Richtung gehen).
  3. Man kann durch ein fest gewähltes Koordinatensystem jeden Punkt (x, y) von $\mathbb{R}^2$ mit dem Pfeil von $(0, 0)$ nach (x, y) identifizieren.

  All dies sind verschiedene Betrachtungsweisen der gleichen Struktur. Die vielleicht am umständlichsten erscheinende erste Version hat den Vorteil, unabhängig von der Wahl eines Koordinatensystems oder Ursprungs zu sein.

Notation: Zeilen- und Spaltenvektoren

Für Elemente v aus dem K-Vektorraum K^n gibt es zwei Standardschreibweisen:

- als Zeilenvektor $(k_1, k_2, \ldots, k_n)$ oder $(k_1\ k_2\ \ldots\ k_n)$ (die Kommata dienen nur der Lesbarkeit und haben keine Bedeutung)
- als Spaltenvektor $\begin{pmatrix} k_1 \\ k_2 \\ \ldots \\ k_n \end{pmatrix}$

Beides sind nur verschiedene Schreibweisen desselben Objekts. In den kommenden Abschnitten wird es aber, abhängig von der Situation, günstiger sein, die eine oder die andere Variante zu wählen.

2.2. Untervektorräume und Erzeugende

In diesem Abschnitt sei V stets ein K-Vektorraum.

Definition: Untervektorraum

$U \subseteq V$ heißt K-*Untervektorraum* von V, falls U unter den eingeschränkten Operationen selbst ein K-Vektorraum ist, d. h. falls $0 \in U$ und für alle $u, u_1, u_2 \in U$ und $k \in K$ die Elemente $u_1 + u_2$, $-u$ und $k \cdot u$ in U liegen. Man schreibt dafür $U \leqslant V$.

Wenn der Körper K durch den Kontext bekannt ist, sagt man auch kurz „Untervektorraum“ statt „K-Untervektorraum“. Außerdem verkürzt man bisweilen „Untervektorraum“ zu „Unterraum“.

Bemerkung:

Man kann sich leicht davon überzeugen, dass sich Regeln wie Assoziativität, Kommutativität und Distributivität oder die Neutralität von 0 automatisch auf Teilmengen übertragen.

Die Abgeschlossenheit bezüglich Negation folgt aus den anderen Regeln, da $-u = (-1) \cdot u$. Wenn $U \neq \emptyset$, etwa $u \in U$, folgt $0 = u + (-u) \in U$. Untervektorräume sind also genau die nicht-leeren, bezüglich Addition und Skalarmultiplikation abgeschlossenen Teilmengen.

Beispiele

Sei $K = \mathbb{R}$ und $V = \mathbb{R}^2$. Die $\mathbb{R}$-Untervektorräume von V sind dann:

- der *triviale Untervektorraum* $\{0_V\}$;
- alle Teilmengen der Form $\{(x, y) \in \mathbb{R}^2 \mid ax + by = 0\}$ für feste $a, b \in \mathbb{R}$ – dies sind die Geraden durch den Ursprung $(0, 0)$;
- der ganze Vektorraum $\mathbb{R}^2$.

Gegenbeispiele

Keine Untervektorräume sind:

- Die Punkte eines Kreises bilden keinen Untervektorraum des $\mathbb{R}^2$ (weder abgeschlossen unter Addition, noch unter Skalarmultiplikation).

- Die Fläche zwischen zwei sich schneidenden Geraden ist kein Untervektorraum des $\mathbb{R}^2$ (abgeschlossen unter Skalarmultiplikation, aber nicht unter Addition).
- Die Punkte mit ganzzahligen Koordinaten, also das „Gitter“ $\mathbb{Z}^2$ (abgeschlossen unter Addition, aber nicht unter Skalarmultiplikation).

Satz 1 Der Schnitt von beliebig vielen K-Untervektorräumen von V ist wieder ein K-Untervektorraum von V.

Beweis zu 1:

Man prüft leicht anhand der Definition nach, dass dies gilt. Falls zum Beispiel $u, v \in \bigcap_{i\in I} U_i$ für Untervektorräume U_i, so sind $u, v \in U_i$ für alle $i \in I$, also ist auch $u+v \in U_i$ für alle $i \in I$ und mithin $u + v \in \bigcap_{i\in I} U_i$. Analog für die anderen Eigenschaften.

Definition: erzeugter Untervektorraum

Sei $X \subseteq V$. Der von X in V *erzeugte Untervektorraum* $\langle X\rangle$ ist der Schnitt aller Untervektorräume von V, die X enthalten. Wegen dem vorangehenden Satz ist dies der bezüglich Inklusion kleinste Untervektorraum von V, der X enthält.

Sprech- und Schreibweisen

Für $\langle\{v_i \mid i \in I\}\rangle$ schreibt man auch kurz $\langle v_i \mid i \in I\rangle$ und für $\langle\{v_1, \ldots, v_n\}\rangle$ kurz $\langle v_1, \ldots, v_n\rangle$.

Der von X erzeugte Untervektorraum heißt auch das *Erzeugnis* von X.

Ist $V = \langle v_i \mid i \in I\rangle$, so sagt man

- die v_i $(i \in I)$ „erzeugen V“ oder
- die v_i $(i \in I)$ „sind Erzeuger (oder Erzeugende) von V“ oder
- $\{v_i \mid i \in I\}$ „ist ein *Erzeugendensystem* von V“

oder Varianten hiervon.

V heißt *endlich erzeugt*, falls es ein endliches Erzeugendensystem gibt.

Definition: Linearkombination

Sei $X \subseteq V$. Eine *Linearkombination* von X ist ein Ausdruck der Form $k_1x_1 + \cdots + k_nx_n$ mit $n \in \mathbb{N}$, $k_i \in K$ und $x_i \in X$. Die Linearkombination heißt *nicht trivial*, wenn mindestens ein k_i nicht null ist.

Notation:

Falls X unendlich ist, soll für Ausdrücke $\sum_{x\in X} k_x x$ gelten, dass alle k_x bis auf endlich viele null sind und die Summe nur über die endlich vielen $k_x x$ gebildet wird, für die $k_x \neq 0$ ist. Damit bezeichent $\sum_{x\in X} k_x x$ also eine Linearkombination von X.

Satz 2 Der von $X \subseteq V$ erzeugte Untervektorraum besteht aus allen durch Linearkombinationen von X beschriebenen Elemente von V. Insbesondere ist $\langle v_1, \dots, v_n \rangle = \{k_1 v_1 + \dots + k_n v_n \mid k_1, \dots, k_n \in K\}$.

Beweis zu 2:

Da jeder Untervektorraum unter Summen und Skalarmultiplikation abgeschlossen ist, enthält er mit $v_1, \dots, v_n$ auch jedes durch eine Linearkombination von $v_1, \dots, v_n$ gegebene Element. Dies gilt also insbesondere für das Erzeugnis einer $v_1, \dots, v_n$ enthaltenden Menge. Also gilt die Inklusion „$\supseteq$“ im Satz.

Für die umgekehrte Inklusion „$\subseteq$“ reicht es zu sehen, dass die Menge der durch Linearkombinationen von X beschriebenen Elemente unter Addition und Skalarmultiplikation abgeschlossen ist und alle Elemente von X enthält: Dies gilt, da $\sum_{x \in X} k_x x + \sum_{x \in X} k'_x x = \sum_{x \in X} (k_x + k'_x) x$, $k \cdot \sum_{x \in X} k_x x = \sum_{x \in X} (k \cdot k_x) x$ und $x = 1 \cdot x$.

Erläuterung

Falls $X = \emptyset$ ist nach Definition $\langle \emptyset \rangle = \{0\}$. Satz 2 stimmt auch in diesem Fall, da der Wert der „leeren Summe“ $\sum_{x \in \emptyset} k_x x$ als 0 definiert wird.

Beispiele

- $(1,0,0), (0,1,0), (0,0,1)$ erzeugen $\mathbb{R}^3$, da sich jedes Element $(x,y,z) \in \mathbb{R}^3$ schreiben lässt als $x \cdot (1,0,0) + y \cdot (0,1,0) + z \cdot (0,0,1)$.
- Ebenso ist $(-1,0,0), (0,2,0), (0,0,1), (1,1,1)$ ein Erzeugendensystem von $\mathbb{R}^3$.
- $(0,1,0), (0,0,2), (0,3,-2)$ dagegen erzeugen einen echten Untervektorraum von $\mathbb{R}^3$, nämlich $\{(0,r,s) \mid r, s \in \mathbb{R}\}$.
- Die Folgen $(1,0,0,0,\dots), (0,1,0,0,\dots), (0,0,1,0,\dots), \dots$ erzeugen einen echten Untervektorraum von $\mathbb{R}^\infty$, nämlich den Untervektorraum der Folgen von endlichem Träger, das sind Folgen $(r_0, r_1, r_2, \dots)$, bei denen alle r_i b is auf endlich viele null sind.

2.3. Lineare Unabhängigkeit, Basis, Dimension

Sei wieder stets V ein K-Vektorraum, und sei $X \subseteq V$ eine Menge von Vektoren.

Definition: Lineare Abhängigkeit

Ein Vektor $v \in V$ ist *linear abhängig* von X, falls $v \in \langle X \rangle$, d. h. falls es $x_1, \dots, x_n \in X$ und $k_1, \dots, k_n \in K$ gibt mit $v = k_1 x_1 + \dots + k_n x_n$.

X ist *linear unabhängig*, falls kein $x \in X$ linear abhängig von $X \setminus \{x\}$ ist.

Satz 3 Eine Menge von unendlich vielen Vektoren ist genau dann linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist.

Beweis zu 3:

Folgt unmittelbar aus der Definition.

Vorsicht

vor den Tücken der Mengenschreibweise bei Doppelnennungen: Angenommen die Menge $\{v_1, v_2\}$ ist linear unabhängig und $v_2 = v_3$. Dann ist $\{v_1, v_2, v_3\} = \{v_1, v_2\}$ linear unabhängig, aber v_3 ist linear abhängig von $\{v_1, v_2\}$. Dies liegt daran, dass hier $\{v_1\} = \{v_1, v_2, v_3\} \setminus \{v_3\} \neq \{v_1, v_2\}$. Diese Schwierigkeit wird mit der folgenden Definition umgangen.

Definition: Menge ohne Doppelnennungen

$\{v_i \mid i \in I\}$ heißt *Beschreibung einer Menge ohne Doppelnennungen*, falls $v_i \neq v_j$ für $i \neq j$, also falls die Elemente v_i für $i \in I$ paarweise verschieden sind.

Anders ausgedrückt: die Abbildung $I \to V$, $i \mapsto v_i$ ist injektiv, oder, noch einmal anders ausgedrückt, $v_j \notin \big\{v_i \mid i \in I \setminus \{j\}\big\}$ für alle $j \in I$.

Der Kürze halber spreche ich von „Menge ohne Doppelnennungen“, obwohl es sich nicht um eine Eigenschaft der Menge, sondern ihrer Beschreibung handelt.

Satz 4 $\{v_1, \ldots, v_n\}$ ist genau dann linear unabhängig und ohne Doppelnennungen, wenn nur die triviale Linearkombination Null ergibt, d. h. wenn $k_1v_1 + \cdots + k_nv_n = 0$ nur für $k_1 = 0, \ldots, k_n = 0$ gilt.

Beweis zu 4:

Wenn die Menge linear abhängig ist oder Doppelnennungen vorliegen, gilt etwa $v_1 \in \langle v_2, \ldots, v_n \langle$ (sonst Umindizieren!), also $v(-1) \cdot v_1 + k_2v_2 + \cdots + k_nv_n = 0$.

Wenn es umgekehrt eine Darstellung $k_1v_1 + \cdots + k_nv_n = 0$ gibt, bei der etwa $k_1 \neq 0$, so folgt $v_1 = -\frac{k_2}{k_1}v_2 + \cdots + (-\frac{k_n}{k_1})v_n$, also ist entweder $v_1 \in \{v_2, \ldots, v_n\}$ und es gibt Doppelnennungen oder die Menge $\{v_1, \ldots, v_n\}$ ist linear abhängig.

Aus Satz 4 folgt unmittelbar eine allgemeine Version auch für unendliche Mengen:

Satz 5 Eine Menge $\{v_i \mid i \in I\}$ ist genau dann linear unabhängig und ohne Doppelnennungen, wenn keine nicht-triviale Linearkombination der Menge 0 ergibt.

Definition: Basis
Eine *Basis* eines Vektorraums V ist ein linear unabhängiges Erzeugendensystem.

Satz 6 $\{v_i \mid i \in I\}$ ist eine Basis von V
$\iff$ $\{v_i \mid i \in I\}$ ist eine maximale linear unabhängige Teilmenge von V
$\iff$ $\{v_i \mid i \in I\}$ ist ein minimales Erzeugendensystem von V
(„maximal“ und „minimal“ sind bezüglich der Teilmengenbeziehung)

Beweis zu 6:

Sei zunächst $B = \{v_i \mid i \in I\}$ eine Basis. Da B linear unabhängig ist, gilt für jedes $b \in B$, dass $b \notin B \setminus \{b\}$, also ist keine echte Teilmenge von B ein Erzeugendensystem von V. Da umgekehrt B Erzeugendensystem von V ist, gilt für beliebiges $v \in V \setminus B$, dass $v \in \langle B \rangle = \langle (B \cup \{v\}) \setminus \{v\} \rangle$, also ist keine echte Obermenge $B \cup \{v\}$ von B linear unabhängig.

Sei nun B maximal linear unabhängig und $v \in V \setminus B$. Dann ist $B \cup \{v\}$ linear abhängig, also existiert eine nicht-triviale Linearkombination $k_1 v_1 + \cdots + k_n v_n + kv = 0$ mit paarweise verschiedenen $v_i \in B$. Es kann nicht $k = 0$ sein, da sonst eine nicht-triviale Linearkombination von Elementen von B null wäre, im Widerspruch zur linearen Unabhängigkeit von B, also ist $v = -\frac{k_1}{k} v_1 + \cdots + -\frac{k_n}{k} v_n \in \langle B \rangle$ und B ist Erzeugendensystem.

Sei nun B minimales Erzeugendensystem und $b \in B$. Dann it $b \notin \langle B \setminus \{b\} \rangle$, mithin ist B linear unabhängig.

Satz 7 Jeder endlich erzeugte Vektorraum besitzt Basen; jedes endliche Erzeugendensystem enthält eine Basis und jede linear unabhängige Teilmenge lässt sich zu einer Basis vergrößern.

Beweis zu 7:

Die erste und die zweite Aussage folgen unmittelbar aus dem vorigen Satz, da sich ein endliches Erzeugendensystem zu einem minimalen Erzeugendensystem verkleinern lässt. Ist eine linear unabhängige Teilmenge X gegeben und ein endliches Erzeugendensystem E, so ist auch $X \cup E$ ein Erzeugendensystem. Nun kann keine echte Teilmenge X' von X ein Erzeugendensystem sein, weil X' sonst als linear unabhängiges Erzeugendensystem zwar eine Basis wäre, aber nicht maximal linear unabhängig. Also muss es unter den Teilmengen Y mit $X \subseteq Y \subseteq X \cup E$ ein minimales Erzeugendensystem geben, das also eine Erweiterung von X zu einer Basis darstellt.

Erläuterung

Dieser Satz gilt auch für unendlich dimensionale Vektorräume, ist aber langwieriger zu beweisen und beruht auf einem etwas komplizierteren mengentheoretischen Axiom.

Satz 8 Je zwei Basen eines Vektorraums haben die gleiche Anzahl von Elementen (im unendlichen Fall: die gleiche Mächtigkeit, d. h. es gibt eine Bijektion zwischen zwei Basen).

Definition: Dimension
Die Anzahl der Elemente der Basen eines K-Vektorraums V heißt *Dimension* von V (über K). Man schreibt dafür $\dim_K V$ oder kurz $\dim V$, wenn K im Kontext festgeschrieben ist.

Beweis zu 8:
Dieser Satz bleibt vorerst ohne Beweis. Für endlich erzeugte Vektorräume folgt der Beweis später aus dem Gauß-Verfahren (man muss sich aber davon überzeugen, dass der Satz für das Gauß-Verfahren nicht gebraucht wird). Für Vektorräume mit unendlichen Basen wird der Satz nicht bewiesen.

Beispiele

- $\mathbb{R}^n$ hat eine Basis $\{e_1, \ldots, e_n\}$ mit $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, etc. Diese Basis heißt *Standardbasis* des $\mathbb{R}^n$. Man sieht, dass $\dim_{\mathbb{R}} \mathbb{R}^n = n$.
- Im Fall $n = 1$ besteht die Standardbasis also aus 1; im Fall $n = 0$ ist die Standardbasis (wie jede andere Basis) die leere Menge.
- $\{(1, 2, 3), (4, 5, 6), (7, 8, 0)\}$ ist eine Basis des $\mathbb{R}^3$. Ein Verfahren zum Überprüfen, ob gegebene Elemente des $\mathbb{R}^n$ eine Basis bilden, wird das Gauß-verfahren liefern.
- $\mathbb{R}[X]$ besitzt (gewissermaßen per Definition) die Basis $\{1, X, X^2, X^3, \ldots\} = \{X^i \mid i \in \mathbb{N}\}$. Auch diese Basis heißt Standardbasis von $\mathbb{R}[X]$. Man sieht, dass $\mathbb{R}[X]$ unendliche Dimension hat.
- $\mathbb{R}^\infty$ hat ebenfall unendliche Dimension; es ist aber keine explizite Basis des Vektorraums bekannt. Die Folgen $(1, 0, 0, 0, \ldots), (0, 1, 0, 0, \ldots), (0, 0, 1, 0, \ldots), \ldots$ sind zwar linear unabhängig, bilden aber kein Erzeugendensystem.
- Alle voranstehenden Beispiele gelten entsprechend für andere Körper wie $\mathbb{F}_2$ oder $\mathbb{C}$, insbesondere hat $\mathbb{F}_2^n$ die Dimension n.
- $\mathbb{C}$ hat als $\mathbb{C}-$Vektorraum die Dimension 1 (mit Standardbasis 1), als $\mathbb{R}-$Vektorraum die Dimension 2, z. B. mit der Basis $\{1, \mathrm{i}\}$. Allgemeiner ist $\dim_{\mathbb{C}} \mathbb{C}^n = n$ und $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$. Eine $\mathbb{R}$-Basis von $\mathbb{C}^n$ ist $\{(1, 0, 0, \ldots, 0), (\mathrm{i}, 0, 0, \ldots, 0), (0, 1, 0, \ldots, 0), (0, \mathrm{i}, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, 1), (0, 0, \ldots, 0, \mathrm{i})\}$.

Satz 9 Seien $v_1, \ldots, v_n$ paarweise verschiedene Elemente. Dann ist $\{v_1, \ldots, v_n\}$ genau dann eine Basis von V, wenn es für jedes $v \in V$ eine eindeutige Darstellung $v = k_1 v_1 + \cdots + k_n v_n$ gibt.

Definition: Koordinaten
Die eindeutig bestimmten Skalare $k_1, \ldots, k_n$ aus Satz 9 werden die *Koordinaten* von v bezüglich der Basis genannt.

Beweis zu 9:
Zunächst ist klar, dass genau dann für jedes $v \in V$ solch eine Darstellung existiert, wenn $\{v_1, \ldots, v_n\}$ ein Erzeugendensystem ist. Angenommen nun $v = k_1 v_1 + \cdots + k_n v_n = k'_1 v_1 + \cdots + k'_n v_n$. Dann gilt $0 = (k_1 - k'_1)v_1 + \cdots + (k_n - k'_n)v_n$, d. h. es gibt genau dann zwei verschiedene Darstellungen für einen Vektor, falls es eine nicht-triviale Linearkombination der Null gibt, was nach Satz 4 genau dann der Fall ist, wenn $\{v_1, \ldots, v_n\}$ nicht linear unabhängig ist.

Auch für diesen Satz kann man eine „unendliche Version“ angeben, die unmittelbar aus Satz 9 folgt:

Satz 10 Eine Teilmenge $\{v_i \mid i \in I\}$ von V ohne Doppelnennungen ist genau dann eine Basis von V, wenn es für jedes $v \in V$ eine eindeutige Darstellung $v = \sum_{i \in I} k_i v_i$ mit $k_i \in K$ gibt.

2.4. Lineare Abbildungen

Seien V und W K-Vektorräume.

Definition: Lineare Abbildung/Vektorraumhomomorphismus
Eine Abbildung $\phi : V \to W$ ist eine *K-lineare Abbildung* oder ein *K-Vektorraumhomomorphismus*, falls ϕ mit der Gruppenstruktur und der Skalarmultiplikation verträglich ist, d. h. falls für alle $v, v_1, v_2 \in V$ und $k \in K$ gilt [3]:

- $\phi(v_1 +_V v_2) = \phi(v_1) +_W \phi(v_2)$, $\phi(0_V) = 0_W$ und $\phi(-_V v) = -_W \phi(v)$
- $\phi(k \cdot_V v) = k \cdot_W \phi(v)$.

Falls aus dem Kontext klar ist, um welchen Körper K es sich handelt, spricht man auch kurz von „linearen Abbildungen“ bzw. „Vektorraumhomomorphismen“.

Bemerkung:
Man kann zeigen, dass die beiden Bedingungen $\phi(0) = 0$ und $\phi(-v) = -\phi(v)$ aus der Additivität $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$ folgt, da $(V, +)$ eine Gruppe ist.

Definition: Isomorphismus und Isomorphie
Eine Abbildung $\phi : V \to W$ ist ein K-*Vektorraumisomorphismus*, falls ϕ eine bijektive Abbildung ist und sowohl ϕ als auch die Umkehrabbildung ϕ^{-1} K-linear sind.

V und W heißen *isomorph* (als K-Vektorräume), falls ein K-Vektorraumisomorphismus $\phi : V \to W$ existiert. Man schreibt dafür $V \cong W$.

Bemerkung:

Man kann zeigen, dass die Umkehrabbildung einer bijektiven K-linearen Abbildung automatisch K-linear ist.

Erläuterung

Der Begriff „isomorph“ und die Notation $V \cong W$ werden auch bei anderen Strukturen eingesetzt (z. B. Gruppen, Ringe). Wenn sie ohne nähere Spezifikation verwendet werden, setzen sie voraus, dass aus dem Kontext klar ist, welche Art von Strukturen betrachtet werden, hier also K-Vektorräume. Ebenso verkürzt man dann auch „Vektorraumisomorphismus“ und „Vektorraumhomomorphismus“ zu „Isomorphismus“ bzw. „Homomorphismus“.

Satz 11 Sei $\{v_i \mid i \in I\}$ eine Basis von V ohne Doppelnennungen, und seien w_i beliebige Elemente von W. Dann gibt es genau eine lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = w_i$ für alle $i \in I$. Außerdem ist ϕ ist genau dann ein Isomorphismus, wenn $\{w_i \mid i \in I\}$ eine Basis von W ohne Doppelnennungen ist.

Beweis zu 11:

Wenn es überhaupt solch eine lineare Abbildung gibt, muss $\phi(k_1 v_{i_1} + \cdots + k_n v_{i_n}) = k_1 w_{i_1} + \cdots + k_n w_{i_n}$ gelten. Da nach Satz 9 jedes v eine eindeutige Darstellung $v = \sum_{j=1}^n k_j v_{i_j}$ besitzt mit $n \in \mathbb{N}$, paarweise verschiedenen $i_j \in I$ und $k_j \in K$, kann man durch $\phi(v) := \sum_{j=1}^n k_j w_{i_j}$ auch tatsächlich eine Abbildung $V \to W$ definieren. Man sieht dann auch leicht ein, dass diese Abbildung tatsächlich linear ist.

Das Bild von ϕ besteht dann aus den Vektoren $\sum_{j=1}^n k_j w_{i_j}$, also ist ϕ genau dann surjektiv, wenn $\{w_i \mid i \in I\}$ ein Erzeugendensystem ist. Wenn ϕ nicht injektiv ist, gibt es zwei verschiedene Vektoren $k_1 v_{i_1} + \cdots + k_n v_{i_n}$ und $k'_1 v_{j_1} + \cdots + k'_m v_{j_m}$ mit gleichem Bild $k_1 w_{i_1} + \cdots + k_n w_{i_n} = k'_1 w_{j_1} + \cdots + k'_m w_{j_m}$. Dann ist $\{w_i \mid i \in I\}$ keine Basis ohne Doppelnennungen, da die Eindeutigkeit der Darstellung aus Satz 9 verletzt ist.

Wenn umgekehrt ϕ bijektiv ist, also ein Isomorphismus ist, gilt $w = \sum_{j1}^n k_j w_{i_j}$ genau dann, wenn $\phi^{-1}(w) = \sum_{j1}^n k_j \phi^{-1}(w_{i_j}) = \sum_{j1}^n k_j v_{i_j}$. Aus der Eindeutigkeit der Darstellung bezüglich der Basis $\{v_i \mid i \in I\}$ folgt damit die Eindeutigkeit der Darstellung bezüglich $\{w_i \mid i \in I\}$. Mit Satz 9 folgt dann, dass $\{w_i \mid i \in I\}$ eine Basis ohne Doppelnennungen ist.

Erläuterung

Ein Isomorphismus ist soviel wie eine Umbenennung der Elemente des Vektorraums und überträgt alle aus der Vektorraumsstruktur definierbaren Eigenschaften. Insbesondere bildet er ein Erzeugendensystem auf ein Erzeugendensystem, eine linear unabhängige Menge auf eine linear unabhängige Menge und eine Basis auf eine Basis ab, und kann also nur zwischen Vektorräumen gleicher Dimension bestehen!

Folgerung 12 Eine lineare Abbildung $\phi : V \to W$ ist durch die Bilder einer Basis festgelegt.

Folgerung 13 Genau dann gibt es einen K-Vektorraumisomorphismus $\phi : V \to W$, wenn $\dim_K V = \dim_K W$.

Beweis zu 13:

Wenn $\phi : V \to W$ ein Isomorphismus ist und B eine Basis von V, dann ist $\{\phi(b) \mid b \in B\}$ eine Basis von W der gleichen Mächtigkeit.

Wenn B und B' Basen gleicher Mächtigkeit von V bzw. W sind, angezeigt durch eine Bijektion $\beta : B \to B'$, dann setzt sich β zu einer bijektiven linearen Abbildung $V \to W$, also einem Isomorphismus, fort.

Definition: angeordnete Basis

Eine *angeordnete Basis* $(v_1, \ldots, v_n)$ ist eine Basis $\{v_1, \ldots, v_n\}$ ohne Doppelnennungen zusammen mit einer festen Reihenfolge der Elemente (nämlich der Anordnung, in der die Elemente als Komponenten des n-Tupels auftreten).[4]

Satz 14 Sei V ein n-dimensionaler K-Vektorraum. Dann wird durch jede angeordnete Basis $B = (v_1, \ldots, v_n)$ ein Vektorraumisomorphismus $i_B : V \to K^n$, $v_i \mapsto e_i$ festgelegt. Dabei wird $v = k_1 v_1 + \cdots + k_n v_n$ auf seine Koordinaten $(k_1, \ldots, k_n)$ bezüglich der Basis B abgebildet.

Umgekehrt bestimmt jeder Vektorraumisomorphismus $i : V \to K^n$ eine angeordnete Basis B von V, nämlich $(i^{-1}(e_1), \ldots, i^{-1}(e_n))$, und es ist $i = i_B$.

Beispiele

Sei nun stets $K = \mathbb{R}$ (wobei die Überlegungen, abgesehen von der geometrischen Anschauung, ebenso für jeden anderen Körper K gelten) und $\phi : V \to W$ eine $\mathbb{R}$-lineare Abbildung zwischen endlich-dimensionalen $\mathbb{R}$-Vektorräumen $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$. Dann ist ϕ festgelegt durch die Bilder der Standardbasis $\{e_1, \ldots, e_n\}$. Es ist nun üblich und günstig, die Elemente von V und W als Spaltenvektoren zu schreiben. Wir betrachten zunächst drei Spezialfälle:

- Sei zunächst $n = m = 1$. Dann ist $e_1 = 1$. Mit $\lambda := \phi(e_1) = \phi(1) \in \mathbb{R}$ gilt dann:

$$\phi(r) = \phi(r \cdot 1) = r \cdot \phi(1) = \lambda \cdot r.$$

  Die linearen Abbildungen $\mathbb{R} \to \mathbb{R}$ sind also genau die Multiplikationen mit einer festen reellen Zahl.

- Sei nun n beliebig und $m = 1$. Mit $\lambda_1 := \phi(e_1), \ldots, \lambda_n := \phi(e_n)$ gilt dann:

$$\phi\Big(\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}\Big) = \phi\Big(\sum_{i=1}^{n} r_i \cdot e_i\Big) = \sum_{i=1}^{n} r_i \cdot \phi(e_i) = \lambda_1 \cdot r_1 + \cdots + \lambda_n \cdot r_n$$

  Die Urbilder der Elemente der Bildraums $\mathbb{R}$ bilden parallele, zu $(\lambda_1, \ldots, \lambda_n)$ senkrechte Hyperebenen im $\mathbb{R}^n$. Man kann die Abbildung geometrisch verstehen als die Projektion auf die Gerade durch den Ursprung in Richtung $(\lambda_1, \ldots, \lambda_n)$, die noch um die Länge von $(\lambda_1, \ldots, \lambda_n)$, also um den Faktor $\sqrt{\lambda_1^2 + \cdots + \lambda_n^2}$, skaliert (d. h. gestreckt oder gestaucht) wird.

- Sei nun $n = 1$ und m beliebig. Mit $\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} := \phi(e_1) = \phi(1)$ gilt dann:

$$\phi(r) = \phi(r \cdot 1) = r \cdot \phi(1) = r \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = \begin{pmatrix} \mu_1 \cdot r \\ \vdots \\ \mu_m \cdot r \end{pmatrix}$$

  Das Bild von ϕ ist also die Gerade durch den Punkt $\phi(1)$; die Abbildung ϕ bildet $\mathbb{R}$ unter Streckung bzw. Stauchung (Skalierung um die Länge von $\phi(1)$) auf diese Gerade ab.

- Seien schließlich im allgemeinen Fall n und m beliebig. Mit

$$\begin{pmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{m1} \end{pmatrix} := \phi(e_1)\ , \begin{pmatrix} \mu_{12} \\ \mu_{22} \\ \vdots \\ \mu_{m2} \end{pmatrix} := \phi(e_2)\ , \ldots, \begin{pmatrix} \mu_{1n} \\ \mu_{2n} \\ \vdots \\ \mu_{mn} \end{pmatrix} := \phi(e_n)$$

  gilt dann:

$$\phi\Big(\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}\Big) = \phi\Big(\sum_{i=1}^{n} r_i \cdot e_i\Big) = \sum_{i=1}^{n} r_i \cdot \phi(e_i) =$$

$$= r_1 \cdot \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{m1} \end{pmatrix} + \cdots + r_n \cdot \begin{pmatrix} \mu_{1n} \\ \vdots \\ \mu_{mn} \end{pmatrix} = \begin{pmatrix} \mu_{11} \cdot r_1 + \cdots + \mu_{1n} \cdot r_n \\ \vdots \\ \mu_{m1} \cdot r_1 + \cdots + \mu_{mn} \cdot r_n \end{pmatrix}$$

Um diese Abbildungen besser beschreiben zu können, führt man Matrizen ein.

Definition: Matrix
Eine $(m \times n)$-*Matrix* über eine Körper K ist eine rechteckige Anordnung von mn Körperelementen a_{ij} für $i = 1, \dots, m$ („Zeilenindex") und $j = 1, \dots, n$ („Spaltenindex") in m Zeilen und n Spalten:

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

Die Menge aller $(m \times n)$-Matrizen mit Einträgen aus K wird mit $\mathrm{Mat}_{m\times n}(K)$ bezeichnet.

Notation:
Wenn nicht explizit anders angegeben, werden die Einträge einer mit einem Großbuchstaben bezeichneten Matrix durch die entsprechenden Kleinbuchstaben beschrieben. Es hat also z. B. die Matrix C in der Regel Einträge c_{ij}, d. h.. $C = (c_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$.

Eine $(m \times n)$-Matrix A besteht aus

- m Zeilenvektoren $z_1 = (a_{11}, a_{12}, \dots, a_{1n})$, …, $z_m = (a_{m1}, a_{m2}, \dots, a_{mn})$
- und aus n Spaktenvektoren $s_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, s_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$.

Dies deute ich bei Bedarf durch die Schreibweisen $A = \left(\begin{array}{c} z_1 \\ \hline \vdots \\ \hline z_m \end{array}\right)$ bzw. $A = (s_1 | \dots | s_n)$ an.

Definition: Multiplikation einer Matrix mit einem Vektor
Man definiert die *Multiplikation einer* $(m \times n)$*-Matrix mit einem Spaltenvektor* aus K^n durch die Formel:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} a_{11}r_1 + a_{12}r_2 + \dots + a_{1n}r_n \\ a_{21}r_1 + a_{22}r_2 + \dots + a_{2n}r_n \\ \vdots \\ a_{m1}r_1 + a_{m2}r_2 + \dots + a_{mn}r_n \end{pmatrix}$$

Satz 15 Durch diese Definition ergibt sich, dass die linearen Abbildungen $K^n \to K^m$ genau die Multiplikationen (von links) mit $(m \times n)$-Matrizen sind. Zur linearen Abbildung $\phi : K^n \to K^m$ gehört dabei die $(m \times n)$-Matrix

$$\big(\, \phi(e_1) \,\big|\, \dots \,\big|\, \phi(e_m) \,\big).$$

Man sagt dafür auch, dass die lineare Abbildung durch die Matrix *dargestellt* wird.

Erläuterung

In Zukunft werde ich oft die $(m \times n)$-Matrix A mit der linearen Abbildung $K^n \to K^m$, $v \mapsto A \cdot v$ identifizieren und zum Beispiel von der „Abbildung A“ sprechen.

2.5. Matrixmultiplikation

Satz 16 Seien $\phi : K^n \to K^m$ und $\psi : K^m \to K^l$ beides K-lineare Abbildungen. Dann ist $\psi \circ \phi : K^n \to K^l$ ebenfalls K-linear.

Beweis zu 16:

Man rechnet nach, dass $(\psi \circ \phi)(v_1 + v_2) = \psi(\phi(v_1 + v_2)) = \psi(\phi(v_1) + \phi(v_2)) = \psi(\phi(v_1)) + \psi(\phi(v_2)) = (\psi \circ \phi)(v_1) + (\psi \circ \phi)(v_2)$ und $(\psi \circ \phi)(k \cdot v) = \psi(\phi(k \cdot v)) = \psi(k \cdot \phi(v)) = k \cdot \psi(\phi(v)) = k \cdot (\psi \circ \phi)(v)$.

Frage

Die Abbildungen ϕ, ψ und $\psi \circ \phi$ aus Satz 16 werden durch eine $(m \times n)$-Matrix A, eine $(l \times m)$-Matrix B und eine $(l \times n)$-Matrix C dargestellt. Wie hängt nun C mit A und B zusammen? Wie kann man C aus A und B ausrechnen?

Dazu rechnet man $C \cdot v = (B \cdot A) \cdot v$ aus (siehe Formelkasten in Abbildung 2.1) und stellt fest, dass der (i,k)-Eintrag der Matrix C sich berechnet als

$$c_{ik} = \sum_{j=1}^{m} b_{ij} a_{jk} = \begin{pmatrix} b_{i1} & \dots & b_{im} \end{pmatrix} \cdot \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} = \begin{array}{c} \\ i\text{-te Zeile} \\ \\ \end{array} \begin{pmatrix} \dots & \dots & \dots \\ b_{i1} & \dots & b_{im} \\ \dots & \dots & \dots \end{pmatrix} \cdot \begin{array}{c} j\text{-te Spalte} \\ \begin{pmatrix} \vdots & a_{1k} & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & a_{mk} & \vdots \end{pmatrix} \end{array},$$

wobei hierfür die i-te Zeile von B mit der k-ten Spalte von A so multipliziert wird, wie im letzten Abschnitt definiert (dies heißt auch *Skalarprodukt* des i-ten Zeilenvektors von B mit dem k-ten Spaltenvektor von A, siehe Defintion 2.9).

$$C \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \psi(\phi(\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix})) = B \cdot \big(A \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}\big) =$$

$$= B \cdot \begin{pmatrix} \sum\limits_{i=1}^{n} a_{1i} v_i \\ \vdots \\ \sum\limits_{i=1}^{n} a_{mi} v_i \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{m} b_{1i} \sum\limits_{i=1}^{n} a_{ji} v_i \\ \vdots \\ \sum\limits_{j=1}^{m} b_{li} \sum\limits_{i=1}^{n} a_{ji} v_i \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{m} \sum\limits_{i=1}^{n} b_{1i} a_{ji} v_i \\ \vdots \\ \sum\limits_{j=1}^{m} \sum\limits_{i=1}^{n} b_{li} a_{ji} v_i \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{n} (\sum\limits_{j=1}^{m} b_{1j} a_{ji}) v_i \\ \vdots \\ \sum\limits_{i=1}^{n} (\sum\limits_{j=1}^{m} b_{lj} a_{ji}) v_i \end{pmatrix}$$

$$= \begin{pmatrix} \sum\limits_{j=1}^{m} b_{1j} a_{j1} & \dots & \sum\limits_{j=1}^{m} b_{1j} a_{jn} \\ \vdots & & \vdots \\ \sum\limits_{j=1}^{m} b_{lj} a_{j1} & \dots & \sum\limits_{j=1}^{m} b_{lj} a_{jn} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Abbildung 2.1.: Matrizenmultiplikation

Definition: Matrixprodukt
Das *Matrixprodukt* $B \cdot A$ einer $(l \times m)$-Matrix B mit einer $(m \times n)$-Matrix A ist die $(l \times n)$-Matrix C mit Einträgen $c_{ik} = \sum\limits_{j=1}^{m} b_{ij} a_{jk}$.

Erläuterung
Das Matrixprodukt $B \cdot A$ ist also dann und nur dann definiert, wenn die Anzahl der Spalten von B gleich der Anzahl der Zeilen von A ist. Als Merkregel für die Dimensionen der Matrizen kann man sich „$(l \times m) \cdot (m \times n) = (l \times n)$“ einprägen; der gemeinsame mittlere Term verschwindet also.

Das Matrixprodukt wurde genau so definert, dass folgendes gilt:

Satz 17 Wenn A eine $(m \times n)$-Matrix über K ist und B eine $(l \times m)$-Matrix über K und $v \in K^n$, so gilt

$$(B \cdot A) \cdot v = B \cdot (A \cdot v).$$

Erläuterung
Im letzten Abschnitt wurde das Produkt $B \cdot v$ einer $(l \times m)$-Matrix B mit einem Spal-

tenvektor $v \in K^m$ definiert. Nun ist solch ein Spaltenvektor v nichts anderes als eine $(m \times 1)$-Matrix A. Somit ist also das Produkt $B \cdot v$ eigentlich doppelt definiert, aber man kann sich leicht anhand der Formeln davon überzeugen, dass beide Definitionen übereinstimmen.

Dass dies kein Zufall ist, sieht man folgendermaßen ein: Man kann einen Vektor $v \in K^m$ mit der linearen Abbildung $K^1 \to K^m$, $1 \mapsto v$ identifizieren, deren Matrix gerade der Spaltenvektor v ist. (Die Abbildung ist also die Multiplikation eines Skalars mit v.) Die Verküpfung dieser Abbildung mit der durch B beschriebenen linearen Abbildung ist dann die lineare Abbildung $K^1 \to K^l$, welche 1 auf $B \cdot v$ abbildet. Die Matrix dieser Abbildung berechnet sich als das Matrixprodukt von B und v, ist aber andererseits der Spaltenvektor $B \cdot v$.

Man hätte sich aber auch umgekehrt die Matrixmultiplikation aus der Multiplikation einer Matrix mit einem Vektor herleiten können. Wenn A die lineare Abbildung $\phi : K^n \to K^m$ darstellt und B die Abbildung $\psi : K^m \to K^l$, so gilt $(\psi \circ \phi)(e_i) = \psi(\phi(e_i)) = \psi(A \cdot e_i) = B \cdot (A \cdot e_i)$, d. h. der i-te Spaltenvektor der Matrix zu $\psi \circ \phi$ ist $B \cdot s_i$, wobei s_i der i-te Spaltenvektor von A ist. Wenn A nur aus einer Spalte besteht, ist dies also die schon bekannte Multiplikation der Matrix B mit dem Spaltenvektor.

Man sieht also, dass das Matrixprodukt $B \cdot A$ „spaltenweise in A“ funktioniert, d. h. wenn $A = (s_1 | \dots | s_n)$, so ist $B \cdot A = (B \cdot s_1 | \dots | B \cdot s_n)$. Umgekehrt funktioniert es „zeilenweise in B“, d. h. wenn $B = \left(\begin{array}{c} z_1 \\ \hline \vdots \\ \hline z_m \end{array}\right)$, so ist $B \cdot A = \left(\begin{array}{c} z_1 \cdot A \\ \hline \vdots \\ \hline z_m \cdot A \end{array}\right)$, woebi hier in den Zeilen also das Matrixprodukt der Zeilenvektoren von B, aufgefasst als $(1 \times m)$-Matrixzen, mit der $(m \times n)$-Matrix A steht.

Beispiele

- Ein Beispiel für eine (willkürlich gewählte) Matrixmultiplikation:
  $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-1) + 2 \cdot 0 + 3 \cdot 1 & 1 \cdot 0 + 2 \cdot 2 + 3 \cdot 3 \\ 4 \cdot (-1) + 5 \cdot 0 + 6 \cdot 1 & 4 \cdot 0 + 5 \cdot 2 + 6 \cdot 3 \end{pmatrix} = \begin{pmatrix} 2 & 13 \\ 2 & 28 \end{pmatrix}$
- Die Verküpfung „Spiegelung an der y-Achse $\circ$ Spiegelung an der x-Achse“ wird beschrieben durch
  $$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$
  ergibt also die Matrix der Punktspiegelung am Ursprung.
- Eine Drehung um den Winkel α mit anschließender Drehung um den Winkel β ergibt insgesamt eine Drehung um $\alpha + \beta$. Aus der Berechnung des Matrixprodukts ergeben sich dadurch die Additionstheoreme für Sinus und Cosinus:

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

Definition: Einheitsmatrix

Die zur Identitätsabbildung id : $K^n \to K^n$ gehörige Matrix ist die *Einheitsmatriz* genannte $(n \times n)$-Matrix I_n, deren Spalten (bzw. Zeilen) gerade die Standardbasisvektoren sind.

Die zur konstanten Nullabbildungen $K^n \to K^m$, $v \mapsto 0$, gehörige Matrix ist die *Nullmatrix*, deren Einträge alle 0 sind. Sie wird meist ebenfalls mit 0 bezeichnet.

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

Exkurs zur Komplexität der Matrizenmultiplikation

Matrizenmultiplikationen spielen in vielen algorithmischen Anwendungen eine große Rolle; es ist daher interessant und nützlich, möglichst schnelle Verfahren zu finden. Das Verfahren, das der Definition folgt, läuft für zwei $(n \times n)$-Matrizen in $O(n^3)$: pro Eintrag n Multiplikationen und $n-1$ Additionen. Für große Matrizen gibt es aber schnellere Verfahren: Das erste solche wurd 1969 von Volker Strassen[5] entwickelt und läuft in $O(n^{2{,}807})$. Er wurde nach und nach verbessert; den letzten großen Schritt lieferte 1990 der Coppersmith-Winograd-Algorithmus[6] mit $O(n^{2{,}3737})$. Etwas überraschend kam 2010 nochmals eine Verbesserung durch Andrew Stothers; der derzeit letzte Stand ist ein Algorithmus von Virginia Vassilevska Williams aus dem Jahre 2011 mit einer Laufzeit von $O(n^{2{,}3727})$. Als untere Schranke hat man sicher $O(n^2)$, da n^2 Einträge auszurechnen sind; einige Forscher vermuten, dass diese untere Schranke optimal ist, also dass es Algorithmen in $O(n^2)$ gibt.

(Zu bedenken ist dabei, dass kleinere Exponenten wegen der in der O-Notation versteckten Konstanten evtl. nur für sehr große Matrizen Verbesserungen bringen; außerdem sagt die Laufzeit nichst über die Güte des Algorithmus hinsichtlich Stabilität (Fehleranfälligkeit) aus. Die Verbesserung des Exponenten in der dritten Nachkommastelle scheint

[5] Volker Strassen ($*$ 1936), ehemaligere Student der Universität Freiburg, zuletzt Professor in Konstanz.

[6] nach Don Coppersmith ($*$ ca. 1950) und Shmuel Winograd ($*$ 1936), damals IBM.

zunächst vernachlässigbar, es ist aber bereits $1000^{2,3737} - 1000^{2,3727} \approx 10^5$; bei vielen Multiplikationen großer Matrizen kann sich also ein spürbarer Effekt ergeben.)

Satz 18 Die Matrizenmulitplikation ist assoziativ, aber i. a. nicht kommutativ, auch bei $(n \times n)$-Matrizen untereinander. Die Einheitsmatrizen sind neutrale Elemente in dem Sinn, dass $I_m \cdot A = A$ und $A \cdot I_n = A$ für jede $(m \times n)$-Matrix A gelten. Nullmatrizen sind *absorbierende Elemente*, d. h. es gilt $0 \cdot A = 0$ und $A \cdot 0 = 0$ (für die Nullmatrix passender Größe, so dass also die Multiplikationen definiert sind).

Beweis zu 18:

Alle Eigenschaften folgen daraus, dass sie auf Seite der zugehörigen Abbildungen gelten. Die nicht vorhandene Kommutativität sieht man z. B. an

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}.$$

Bemerkung:

Eine (1×1)-Matriz (a_{11}) kann man mit der Zahl a_{11} identifizieren. Die Multiplikation von (1×1)-Matrizen ist also kommutativ.

Abgesehen von der fehlenden Kommutativität gibt es noch andere Eigenschaften, welche die Matrizenmultiplikation von der Multiplikation z. B. reeller Zahlen unterscheidet. So gibt es sogenannte „nilpotente“ Elemente, das sind Matrizen $A \neq 0$ mit $A^n = 0$ für ein $n > 0$. Zum Beispiel gilt:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Insbesondere folgt für Matrizen aus $A \cdot B = 0$ nicht $A = 0$ oder $B = 0$!

Definition: Vektorräume $\mathrm{Abb}(K^n, K^m)$ **und** $\mathrm{Lin}(K^n, K^m)$

Abbildungen $\phi, \psi : K^n \to K^m$ kann man addieren durch $(\phi + \psi)(v) := \phi(v) + \psi(v)$ und skalar multiplizieren durch $(k \cdot \phi)(v) := k \cdot \phi(v)$. Die Menge der Abbildungen wird dadurch zu einem K-Vektorraum $\mathrm{Abb}(K^n, K^m)$. Die Teilmenge der linearen Abbildungen $K^n \to K^m$ bildet darin einen Untervektorraum $\mathrm{Lin}(K^n, K^m)$.

Man kann nun die Addition und Skalarmultiplikation mittels der Identifikation von linearen Abbildungen und Matrizen in Satz 15 auf Matrizen ausdehnen, so dass die Menge $\mathrm{Mat}_{m \times n}(K)$ zu einem zu $\mathrm{Lin}(K^n, K^m)$ isomorphen K-Vektorraum wird. Man kann nun leicht nachrechnen, dass die folgende Definition die *Matrizenaddition* und die *Skalarmultiplikation von Matrizen* beschreibt:

Definition: Vektorraumstruktur auf $\mathrm{Mat}_{m\times n}(K)$
Seien A und B $(m \times n)$-Matrizen über K und $k \in K$. Dann ist

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

$$k\cdot A=k\cdot\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn}\end{pmatrix}:=\begin{pmatrix} ka_{11} & \dots & ka_{1n} \\ \vdots & & \vdots \\ ka_{m1} & \dots & ka_{mn}\end{pmatrix}$$

Satz 19
(a) Die $(m \times n)$-Matrizen über K bilden einen mn-dimensionalen, zu $\mathrm{Lin}(K^n, K^m)$ isomorphen K-Vektorraum $\mathrm{Mat}_{m\times n}(K)$. Das neutrale Element der Addition ist die $(m\times n)$-Nullmatrix.
Die Matrizen E_{ij}, deren (i, j)-Eintrag jeweils 1 ist und alle anderen Einträge 0, bilden eine Basis, die *Standardbasis* von $\mathrm{Mat}_{m\times n}(K)$ genannt wird. Jede Aufzählung der Standardbasis liefert einen Vektorraum-Isomorphismus $\mathrm{Mat}_{m\times n}(K) \to K^{mn}$, der die Standardbasis von $\mathrm{Mat}_{m\times n}(K)$ in der gewählten Reihenfolge auf die Standardbasis von K^{mn} in der natürlichen Reihenfolge abbildet.

(b) Es gelten die Distributivgesetze, d. h. immer dann, wenn die Operationen definiert sind, gelten $A\cdot(B_1+B_2) = (A\cdot B_1)+(A\cdot B_2)$ und $(B_1+B_2)\cdot A = (B_1\cdot A)+(B_2\cdot A)$.

(c) Die quadratischen $(n \times n)$-Matrizen $\mathrm{Mat}_{n\times n}(K)$ bilden mit Matrizenaddition und -multiplikation einen (für $n \geq 2$ nicht-kommutativen) Ring mit Eins I_n.

(d) $k \cdot I_n$ ist die „$(n \times n)$-Diagonalmatrix“ mit Einträgen k auf der Hauptdiagonale von links oben nach rechts unten und Einträgen 0 an allen anderen Stellen. Es gilt dann $k\cdot A = (k\cdot I_n)\cdot A = A\cdot(k\cdot I_n)$. Es folgt daraus, dass die Skalarmultiplikation mit der Matrizenmultiplikation vertauscht, d. h. es gilt $k\cdot(A\cdot B) = (k\cdot A)\cdot B = A\cdot(k\cdot B)$, sofern das Produkt $A \cdot B$ definiert ist.

Beweis zu 19:
Die Matrizen E_{ij} bilden eine Basis, da sich jede Matrix eindeutig schreiben lässt als $A = \sum_{i,j} a_{ij}E_{ij}$. Die Distributivgesetze und Teil (d) gelten, weil es auf der Seite der linearen Abbildungen gilt. Alles andere folgt aus der bisher entwickelten Theorie.

2.6. Basiswechsel

Die in diesem Abschnitt betrachteten Vektorräume seien alle endlich-dimensional.

Definition: invertierbare Matrizen
Eine $(n \times n)$-Matrix A über K heißt *invertierbar*, wenn die zugehörige lineare Abbildung $K^n \to K^n$ invertierbar ist, d. h. wenn eine $(n \times n)$-Matrix A^{-1} existiert (nämlich die Matrix zur Umkehrabbildung) mit

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

Bemerkung:
Wegen der Eindeutigkeit der Umkehrabbildung (alternativ durch die gleiche Überlegung wie in Gruppen) sieht man, dass die Matrix A^{-1} durch die Eigenschaft $A \cdot A^{-1} = I_n$ oder $A^{-1} \cdot A = I_n$ bereits eindeutig bestimmt ist.

Satz 20 A ist genau dann invertierbar, wenn die Spaltenvektoren $A \cdot e_1, \ldots, A \cdot e_n$ von A eine Basis von K^n bilden. Die Umkehrabbildung ist dann durch die Zuordnung $A \cdot e_i \mapsto e_i$ festgelegt.
Offensichtlich ist A^{-1} selbst wieder invertierbar und es gilt $(A^{-1})^{-1} = A$.
Falls A und B invertierbare $(n \times n)$-Matrizen sind, so ist auch $B \cdot A$ invertierbar und es gilt $(B \cdot A)^{-1} = B^{-1} \cdot A^{-1}$.

Beweis zu 20:
Der erste Teil folgt direkt aus Satz 11. Die anderen Teile gelten in beliebigen Monoiden: Es ist per Definition von A^{-1} klar, dass A auch invers zu A^{-1} ist, und man rechnet nach, dass $B^{-1} \cdot A^{-1}$ invers zu $B \cdot A$ ist.

Erläuterung
Ziel dieses Abschnitts ist es nun, lineare Abbildungen zwischen beliebigen endlich dimensionalen Vektorräumen durch Matrizen zu beschreiben. Da beliebige Vektorräume keine ausgezeicheten Basen haben, wird es – abhängig von gewählten Basen – verschiedene darstellenden Matrizen geben. Eine Hauptfrage wird darin bestehen zu verstehen, wie diese Matrizen miteinander zusammenhängen. Als Spezialfall erhält man dann auch die Darstellung linearer Abbildungen $K^n \to K^m$ bezüglich anderer Basen als den Standardbasen.

Definition: Basiswechsel
Sei V ein n-dimensionaler und W ein m-dimensionaler K-Vektorraum und $\phi : V \to W$ eine K-lineare Abbildung. Sei außerdem $(v_1, \dots, v_n)$ eine angeordnete Basis B von V und $(w_1, \dots, w_m)$ eine angeordnete Basis B' von W. Nach Satz 14 legen B und B' Isomorphismen $i_B : V \to K^n$ und $i_{B'} : W \to K^m$ fest, so dass sich folgendes Diagramm ergibt:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ i_B \downarrow & & \downarrow i_{B'} \\ K^n & & K^m \end{array}$$

Die *Matrix von ϕ bezüglich der Basen B und B'* wird nun definiert als die Matrix der Abbildung $i_{B'} \circ \phi \circ i_B^{-1} : K^n \to K^m$ und wird mit ${}_{B'}\phi_B$ bezeichnet.
Im Spezialfall $V = W$ und $B = B'$ schreibt man kurz ϕ_B für ${}_B\phi_B$.

Bemerkung:
Die Spaltenvektoren der Matrix ${}_{B'}\phi_B$ sind also die Koordinaten von $\phi(v_1), \dots, \phi(v_n)$ bezüglich der angeordneten Basis B'.

Satz 21 Seien V, W, X endlich-dimensionale K-Vektorräume mit angeordneten Basen B, B', B'' und seien $\phi : V \to W$ und $\psi : W \to X$ lineare Abbildungen. Dann gilt

$${}_{B''}(\psi \circ \phi)_B \;=\; ({}_{B''}\psi_{B'}) \cdot ({}_{B'}\phi_B)$$

Beweis zu 21:
${}_{B''}(\psi\circ\phi)_B$ ist nach Definition die Matrix von $i_{B''}\circ(\psi\circ\phi)\circ i_B^{-1} = i_{B''}\circ\psi\circ i_{B'}^{-1}\circ i_{B'}\circ\phi\circ i_B^{-1}$, was gerade das Produkt der Matrix von $i_{B''}\circ\psi\circ i_{B'}^{-1}$ mit der Matrix von $i_{B'}\circ\phi\circ i_B^{-1}$ ist, also $({}_{B''}\psi_{B'}) \cdot ({}_{B'}\phi_B)$.

Satz 22 Sei $\phi : V \to W$ linear, seien B_1, B_2 angeordnete Basen von V und B'_1, B'_2 angeordnete Basen von W. Dann gilt:

$${}_{B'_2}\phi_{B_2} = ({}_{B'_2}\mathrm{id}_W{}_{B'_1}) \cdot ({}_{B'_1}\phi_{B_1}) \cdot ({}_{B_1}\mathrm{id}_V{}_{B_2})$$

Die Matrizen ${}_{B'_2}\mathrm{id}_W{}_{B'_1}$ und ${}_{B_1}\mathrm{id}_V{}_{B_2}$ heißen *Basiswechselmatrizen.*

Im Spezialfall $V = W$ und $B'_i = B_i$ gilt:

$$\phi_{B_2} = ({}_{B_2}\mathrm{id}_{V\,B_1}) \cdot \phi_{B_1} \cdot ({}_{B_1}\mathrm{id}_{V\,B_2}) = ({}_{B_1}\mathrm{id}_{V\,B_2})^{-1} \cdot \phi_{B_1} \cdot ({}_{B_1}\mathrm{id}_{V\,B_2}).$$

Insbesondere sind Basiswechselmatrizen stets invertierbar mit $({}_{B_1}\mathrm{id}_{V\,B_2})^{-1} = {}_{B_2}\mathrm{id}_{V\,B_1}$.

Beweis zu 22:

Der erste Teil folgt direkt aus dem Satz, da $\phi = \mathrm{id}_W \circ \phi \circ \mathrm{id}_V$. Wegen $({}_{B_2}\mathrm{id}_{V\,B_1}) \cdot ({}_{B_1}\mathrm{id}_{V\,B_2}) = {}_{B_2}(\mathrm{id}_V \circ \mathrm{id}_V)_{B_2} = {}_{B_2}\mathrm{id}_{V\,B_2} = I_{\dim V}$ folgt auch die rechte Seite der Gleichung im Spezialfall.

Wie rechnet man die Basiswechselmatrizen aus?

Ist die Basis $B_1 = (v_1, \ldots, v_n)$ von V gegeben und ist v'_j der j-te Vektor in B_2, so muss man also die Koeffizienten a_{ij} mit $v'_j = a_{1j}v_1 + \cdots + a_{nj}v_n$ berechnen; diese stehen als j-te Spalte in der Basiswechselmatrix ${}_{B_1}\mathrm{id}_{V\,B_2}$. Wenn die Basiselemente als Vektoren in K^n gegeben sind (also mit ihren Koordinaten bezüglich der Standardbasis), dann ergibt die Gleichung ein lineares Gleichungssystem, das z. B. nach dem Gauß-Verfahren (siehe folgender Abschnitt) gelöst werden kann. Auch das Invertieren von Matrizen geschieht am besten mit dem Gauß-Verfahren.

Besonders einfach ist es, wenn $V = K^n$ und B_1 die Standardbasis ist: Dann sind die Spaltenvektoren der Basiswechselmatrix ${}_{B_1}\mathrm{id}_{V\,B_2}$ gerade die Vektoren von B_2.

Beispiele

- Sei $V = \mathbb{R}^3$ mit der Basis $B_1 = (e_1, e_2, e_3)$, also der Standardbasis und der Basis $B_2 = (v_1, v_2, v_3)$ mit $v_1 = (0, 0, 1)$, $v_2 = (0, 1, 2)$ und $v_3 = (1, 1, 1)$.
  Sei $W = \mathbb{R}^2$ mit den Basen $B'_1 = (w_1, w_2)$, wobei $w_1 = (1, 1)$ und $w_2 = (1, -1)$, und $B'_2 = (w'_1, w'_2)$, wobei $w'_1 = (1, 0)$ und $w'_2 = (1, 1)$.
  Die eine Basiswechselmatrix von V ergibt sich aus den Vektoren von B_2 als Spalten der Matrix, da B_1 die Standardbasis ist:

  $${}_{B_1}\mathrm{id}_{B_2} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}$$

  Die andere Basiswechselmatrix erhält man als Inverse:

  $${}_{B_2}\mathrm{id}_{B_1} = ({}_{B_1}\mathrm{id}_{B_2})^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

  Man kann zum einen durch Ausmultiplizieren nachprüfen, dass die angegebene Matrix tatsächlich die Inverse ist, also dass${}_{B_1}\mathrm{id}_{B_2} \cdot {}_{B_2}\mathrm{id}_{B_1} = I_3$. Zum andern kann man nachrpüfen, dass ${}_{B_2}id_{B_1}$ tatsächlich die Koeffizienten der Standardbasis bezüglich

B_2 beinhaltet, also dass gilt:

$$\begin{aligned} e_1 &= 1 \cdot v_1 - 1 \cdot v_2 + 1 \cdot v_3 \\ e_2 &= -2 \cdot v_1 + 1 \cdot v_2 + 0 \cdot v_3 \\ e_3 &= 1 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 \end{aligned}$$

Analog sieht man für die Basiswechselmatrizen von W, dass

$$w_1 = 0 \cdot w'_1 + 1 \cdot w'_2 \qquad w'_1 = \frac{1}{2} \cdot w_1 + \frac{1}{2} \cdot w_2$$
$$w_2 = 2 \cdot w'_1 - 1 \cdot w'_2 \qquad w'_2 = 1 \cdot w_1 + 0 \cdot w_2$$

und folglich

$$_{B_{2'}}\mathrm{id}_{B'_1} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \quad \text{und} \quad _{B_{1'}}\mathrm{id}_{B'_2} = (_{B_{2'}}\mathrm{id}_{B'_1})^{-1} = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$$

Sei nun die lineare Abbildung $\psi : V \to W$ bezüglich der Basen B_1, B'_1 beschrieben durch die Matrix

$$_{B'_1}\phi_{B_1} = \begin{pmatrix} 3 & 1 & 2 \\ 0 & 5 & 4 \end{pmatrix}.$$

Dies bedeutet also, dass $\phi(e_1) = 3w_1$, $\phi(e_2) = w_1 + 5w_2$ und $\phi(e_3) = 2w_1 + 4w_2$. Die Matrix von ψ bezüglich der Basen B_2, B'_2 errechnet sich dann als

$$\begin{aligned} _{B_{2'}}\psi_{B'_1} &= ({}'_{B_2}\mathrm{id}_{B'_1}) \cdot (_{B'_1}\psi_{B_1}) \cdot (_{B_1}\mathrm{id}_{B_2}) \\ &= \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 2 \\ 0 & 5 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 10 & 8 \\ 3 & -4 & -2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 26 & 18 \\ -2 & -8 & -3 \end{pmatrix} \end{aligned}$$

Dies bedeutet nun, dass $\phi(v_1) = 8w'_1 - 2w'_2$, $\phi(v_2) = 26w'_1 - 8w'_2$ und $\phi(v_3) = 18w'_1 - 3w'_2$. Exemplarisch kann man dies nachrechnen; so gilt z. B.

$$\begin{aligned} \phi(v_2) &= \phi(e_2 + 2e_3) = \phi(e_2) + 2\phi(e_3) = w_1 + 5w_2 + 2 \cdot (2w_1 + 4w_2) \\ &= 5w_1 + 13w_2 = 5w'_2 + 13(2w'_1 - w'_2) = 26w'_1 - 8w'_2 \end{aligned}$$

- Ein weiteres Beispiel für die Berechnung eines Basiswechsels findet sich bei der Diagonalisierung einer Drehung über den komplexen Zahlen auf Seite 47.
- Ein Spezialfall eines Basiswechsels liegt vor, wenn es sich um die gleichen Basiselemente in anderer Anordnung handelt, wenn der Basiswechsel also in einer Umordnung der Basis besteht:

Wenn B die Basis $(v_1, \dots, v_n)$ ist, wird eine Umordnung beschrieben durch eine *Permutation* der Indizes, also eine Bijektion $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$, wobei die neu angeordnete Basis B^σ dann $(v_{\sigma(1)}, \dots, v_{\sigma(n)})$ ist.[7]

Definition: Permutationsmatrix

Die Basiswechselmatrix $M(\sigma) := {}_{B^\sigma}\mathrm{id}_B$ hat Einträge 1 an den Stellen $(i, \sigma(i))$ und 0 an allen anderen Stellen. Solche Matrizen heißen *Permutationsmatrix*: Sie sind quadratische Matrizen, die in jeder Zeile und in jeder Spalte genau eine 1 haben und sonst überall 0.

Die Inverse zu $M(\sigma)$ ist $M(\sigma^{-1})$, also die Permutationsmatrix mit Einträgen 1 an den Stellen $(i, \sigma^{-1}(i))$. Da jedes i von der Form $\sigma(j)$ ist, ist dann $(i, \sigma^{-1}(i)) = (\sigma(j), j)$, d. h. $M(\sigma^{-1})$ entsteht, indem man $M(\sigma)$ an der Hauptdiagonalen spiegelt. Dies heißt auch die *Transponierte* $M(\sigma)^{\mathrm{T}}$ von $M(\sigma)$.

Beispiel: Sei $n = 3$ und $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1$. Dann ist

$$M(\sigma) = {}_{B^\sigma}\mathrm{id}_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{und } M(\sigma^{-1}) = {}_B\mathrm{id}_{B_\sigma} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

(Kleiner Vorgriff auf Abschnitt **??**: Die Abbildung $\sigma \mapsto M(\sigma)$ ist ein Gruppenhomomorphismus von der Symmetrischen Gruppe $\mathrm{Sym}(n)$ der Permutationen von $\{1, \dots, n\}$ in die multiplikative Gruppe $\mathrm{GL}(n, K)$ der invertierbaren $(n \times n)$-Matrizen.)

Spezialfall: Transpositionen sind spezielle Permutatione, die nur zwei Elemente vertauschen (und damit selbst-invers sind). Die Transposition τ, welche die Elemente i und j vertauscht, schreibt man auch (ij). Der Lesbarkeit halber schreibe ich $M_{(ij)}$ für $M((ij))$. Es gilt dann (alle nicht aufgeführten Einträge sind gleich 0 und o. B. d. A. ist $i < j$):

$$M_{(ij)} = M_{(ij)}^{-1} = \left(\begin{array}{c|c|c|c|c} \begin{smallmatrix}1 & & \\ & \ddots & \\ & & 1\end{smallmatrix} & & & & \\ \hline & & & 1 & \\ \hline & & \begin{smallmatrix}1 & & \\ & \ddots & \\ & & 1\end{smallmatrix} & & \\ \hline & 1 & & & \\ \hline & & & & \begin{smallmatrix}1 & & \\ & \ddots & \\ & & 1\end{smallmatrix} \end{array}\right) \begin{array}{l} \\ i\textit{-te Zeile} \\ \\ j\textit{-te Zeile} \\ \\ \end{array}$$

[7] Bei dieser Version gibt σ also an, welcher Vektor an die jeweilige Stelle gesetzt wird, d. h. $\sigma(2) = 3$ bedeutet, dass v_3 in der neu angeordneten Basis an zweiter Stelle steht. Alternativ könne man als neu angeordnete Basis $(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(n)})$ nehmen. Dann würde σ angeben, an welche Stelle der jeweilige Vektor geschoben wird, d. h. $\sigma(2) = 3$ würde bedeuten, dass v_2 in der neu angeordneten Basis an dritter Stelle käme.

Es ist also etwa (zweite und dritte Zeile und Spalte jeweils vertauschen!)

$$M_{(23)} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 0 & 1 & 2 \\ 3 & 4 & 5 & 6 \end{pmatrix} \cdot M_{(32)} = \begin{pmatrix} 1 & 3 & 2 & 4 \\ 9 & 1 & 0 & 2 \\ 5 & 7 & 6 & 8 \\ 3 & 5 & 4 & 6 \end{pmatrix}.$$

Erläuterung

Ein Ziel der linearen Algebra besteht darin, zu einer gegebenen linearen Abbildung $\phi : V \to V$ eine Basis B zu finden, so dass die Matrix ϕ_B möglichst „schön" ist. Hierzu gibt es eine ganze Reihe von Ergebnissen über sogenannte Normalformen von Matrizen. „Besonders schön" ist eine Matrix in Diagonalgestalt, also von der Form

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

(alles außerhalb der von λ_1 bis λ_n gebikldetetn Diagonalen hat den Eintrag 0).

Für die Basisvektoren $v_1, \ldots, v_n$ gilt dann $\phi(v_i) = \lambda v_i$ und für beliebige Vektoren $\phi(a_1 v_1 + \cdots + a_n v_n) = \lambda a_1 v_1 + \cdots + \lambda a_n v_n$.

Definition: Eigenvektor

Ein Vektor $v \neq 0$ heißt *Eigenvektor* der linearen Abbildung $\phi : V \to V$ zum *Eigenwert* $\lambda \in K$, falls $\phi(v) = \lambda v$.

Der Idealfall besteht also darin, dass man zu einer linearen Abbildung eine Basis aus Eigenvektoren findet. (Wenn man weiß, dass λ ein Eigenwert ist und ϕ durch die Matrix A beschrieben ist, kann man die Eigenvektoren durch Lösen des linearen Gleichungssystemes $A \cdot x = \lambda x$ mit unbekannten Koeffizienten für x finden. Jedes skalare Vielfache eines Eigenvektors ($\neq 0$) ist wieder ein Eigenvektor. Die Eigenwerte wiederum kann man als Nullstellen des sogenannten *charakteristischen Polynoms* bestimmen.)

Im Allgemeinen findet man aber keine Basis aus Eigenvektoren. Es gibt zwei Hinderungsgründe:

(1) *Drehungen im $\mathbb{R}^2$ haben i. a. keine Eigenvektoren.* Dies ist geometrisch sofort ersichtlich. Nur wenn der Drehwinkel ein ganzzahliges Vielfaches von 180° ist, gibt es Eigenvektoren in $\mathbb{R}^2$.

Diese Problem lässt sich dadurch beheben, dass man den Körper erweitert, hier zu den komplexen Zahlen $\mathbb{C}$. So hat z. B. die Drehung um 90° bezüglich der Standardbasis die Matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ – ohne Eigenvektoren in $\mathbb{R}^2$ – aber als Matrix über den komplexen Zahlen sind $\left(\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 \\ -i \end{smallmatrix}\right)$ zwei linear unabhängige Eigenvektoren zu den Eigenwerten $-i$ und i, d. h. bezüglich der aus diesen beiden Vektoren gebildeten Basis ergibt sich die Diagonalform $\left(\begin{smallmatrix} -i & 0 \\ 0 & i \end{smallmatrix}\right)$.

Man kann an diesem Beispiel noch einmal schön den Basiswechsel nachvollziehen: Da eine der Basen die Standardbasis ist, besteht eine der beiden Basiswechselmatrizen aus den Vektoren der anderen Basis als Spalten und die andere Basiswechselmatrix ist deren Inverse:

$$\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \text{ und } \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix};$$

man kann auch nachrechnen, dass diese Matrix tatsächlich die Koeffizienten der Standardbasis bezüglich der neuen Basis enthält, da

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ i \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ und } \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2i}\begin{pmatrix} 1 \\ i \end{pmatrix} - \frac{1}{2i}\begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

Auch den Basiswechsel lässt sich nachrechnen; es gilt:

$$\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\text{und} \quad \begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{1}{2} & -\frac{1}{2i} \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

(2) *Scherungen im $\mathbb{R}^2$ haben i. a. nur einen Eigenvektor (bis auf skalare Vielfache) .* Diese Problem kann nicht durch Vergrößerung des Körpers behoben werden; eine Scherung wie z. B. $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ bildet einen Vektor $v = ae_1 + be_2$ auf $ae_1 + b(e_2 + e_1)$ ab. Man kann leicht nachrechnen, dass nur die skalaren Vielfachen von e_1 Eigenvektoren sind, also wenn $b = 0$.

Man kann nun zeigen, dass dies über $\mathbb{C}$ der einzige Hinderungsgrund ist: Durch geeignete Basiswahl erreicht man die sogenannte Jordan'sche Normalform, bei der die Matrix aus Teilmatrizen der folgenden Form ausgebaut ist, die gewissermaßen höherdimensionale Scherungen beschreiben:

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

2.7. Lineare Gleichungssysteme

Ein lineares Gleichungssystem mit m Gleichungen und n Unbekannten hat die Form

$$\begin{array}{ccccccc} a_{11} \cdot x_1 & + & a_{12} \cdot x_2 & + \dots + & a_{1n} \cdot x_n & = & b_1 \\ & & & & & & \vdots \\ a_{m1} \cdot x_1 & + & a_{m2} \cdot x_2 & + \dots + & a_{mn} \cdot x_n & = & b_m \end{array}$$

wobei a_{ij} aus einem Körper K (in der Regel $\mathbb{R}$) stammen und die x_i Unbekannte sind. Dies entspricht in Matrixschreibweise:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Dabei wird die erste Matrix als A bezeichnet, die zweite als x und die dritte als b. Das Gleichungssystem ist folglich $A \cdot x = b$. Falls $\psi : K^n \to K^m$ die durch A beschriebene lineare Abbildung ist, dann sind die Lösungen des Gleichungssystems genau die $v \in K^n$ mit $\psi(v) = b$.

Definition: Kern und Bild

Sei $\phi : V \to W$ eine lineare Abbildung. Das *Bild von* ϕ ist definiert als $\mathrm{Bild}(\phi) := \{\phi(v) \mid v \in V\}$; der *Kern von* ϕ als $\mathrm{Kern}(\phi) := \{v \in V \mid \phi(v) = 0\}$.

Das Bild wird ebenso für beliebige Abbildungen definiert. Bild und Kern werden auch (nach dem englischen *image* und *kernel*) als $\mathrm{im}(\phi)$ und $\ker(\phi)$ bezeichnet.

Notation: Urbilder

Ist $f : A \to B$ eine beliebige Abbildung und $b \in B$, so bezeichnet $f^{-1}[b]$ die Menge $\{a \in A \mid f(a) = b\}$. Meist wird dafür $f^{-1}(b)$ geschrieben. Falls f bijektiv ist und die Umkehrfunktion $f^{-1} : B \to A$ existiert, so wird die Schreibweise mit runden Klammern aber zweideutig. Mit der exakteren Schreibweise gilt $f^{-1}[b] = \{f^{-1}(b)\}$.

Satz 23 (a) Kern und Bild einer linearen Abbildung $\phi : V \to W$ sind Unterräume von V bzw. W.

(b) Falls $w_0 = \phi(v_0) \in \mathrm{Bild}(\phi)$, so ist $\phi^{-1}[w_0] = v_0 + \mathrm{Kern}(\phi) := \{v_0 + v \mid v \in \mathrm{Kern}(\phi)\}$. Mit anderen Worten, es gilt $\phi(v) = \phi(v') \iff v - v' \in \mathrm{Kern}(\phi)$.

Beweis zu Eigenschaften von Kern und Bild:

(a) Da $\phi(0_V) = 0_W$ ist $0_V \in \mathrm{Kern}(\phi)$ und $0_W \in \mathrm{Bild}(\phi)$.

Seien $w_1 = \phi(v_1)$ und $w_2 = \phi(v_2)$ in $\mathrm{Bild}(\phi)$. Dann sind $w_1 + w_2 = \phi(v_1 + v_2)$ und $k \cdot w_1 = \phi(k \cdot v_1)$ ebenfalls in $\mathrm{Bild}(\phi)$, also ist $\mathrm{Bild}(\phi)$ ein Untervektorraum.

Seien $v_1, v_2 \in \mathrm{Kern}(\phi)$. Dann ist $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2) = 0 + 0 = 0$ und $\phi(k \cdot v_1) = k \cdot \phi(v_1) = k \cdot == 0$. Also ist auch $\mathrm{Kern}(\phi)$ ein Untervektorraum.

(b) Es ist $\phi(v) = \phi(v') \iff \phi(v - v') = \phi(v) - \phi(v') = 0 \iff v - v' \in \mathrm{Kern}(\phi)$.

Erläuterung

Für ein $w \in W$ gibt es also zwei mögliche Fälle: Entweder $w \notin \mathrm{Bild}(\phi)$ und $\phi^{-1}[w] = \emptyset$; oder $w \in \mathrm{Bild}(\phi)$ und $\phi^{-1}[w]$ ist eine sogenannte „Nebenklasse“ $v + \mathrm{Kern}(\phi)$ von $\mathrm{Kern}(\phi)$ mit $\phi(v) = w$.

Natürlich ist ϕ surjektiv, wenn $\mathrm{Bild}(\phi) = W$ (gilt für beliebige Abbildungen $\phi : V \to W$).

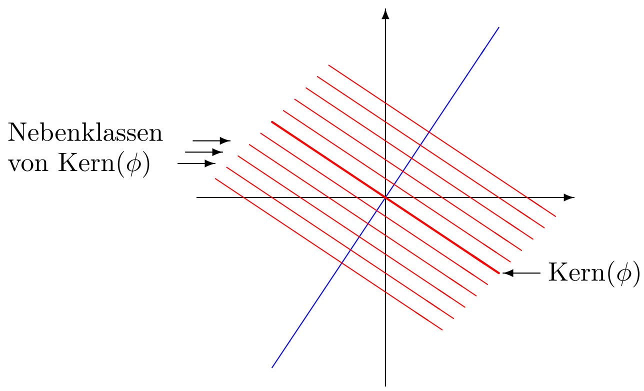
Folgerung 24 ϕ ist injektiv $\iff$ $\mathrm{Kern}(\phi) = \{0\}$ $\iff$ $\dim \mathrm{Kern}(\phi) = 0$.
Wenn W endlich-dimensional ist, so ist ϕ surjektiv $\iff$ $\dim \mathrm{Bild}(\phi) = \dim W$.

Beweis zu 24:

Der erste Teil folgt direkt aus Satz 23. Der zweite Teil folgt, weil ein echter Untervektorraum eines endlich-diemnsionalen Vektorraums kleinere Dimension hat: Eine Basis B des Untervektorraums ist noch keine Basis von W, kann aber zu einer Basis von W ergänzt werden, hat also weniger Elemente.

Beispiele

Betrachte die lineare Abbilung $\phi : \mathbb{R}^2 \to \mathbb{R}^2$, welche die senkrechte Projektion auf die (blaue) Gerade $y = 1,5x$ darstellt. Diese Gerade ist das Bild von ϕ; die im Ursprung dazu senkrecht stehende (fette rote) Gerade ist der Kern von ϕ. Die Parallelen dazu sind die Nebenklassen des Kerns, und zwar ist jede dieser Geraden das volle Urbild ihres Schnittpunktes mit der blauen Geraden.



Satz 25 Sei $\phi : V \to W$ linear. Dann gilt $\dim \mathrm{Kern}(\phi) + \dim \mathrm{Bild}(\phi) = \dim V$.

Beweis zu Dimensionssatz:

[8] Sei $l = \dim \mathrm{Kern}(\phi)$ und $n = \dim V$ und wähle eine Basis $\{v_1, \dots, v_l\}$ von $\mathrm{Kern}(\phi)$. Diese ist eine lineare unabhängige Teilmenge von V, kann also zu einer maximal linear unabhängigen Teilmenge $\{v_1 \dots, v_l, v_{l+1}, \dots, v_n\}$ ergänzt werden, d. h. zu einer Basis von V. Zu zeigen ist also $n - l = \dim \mathrm{Bild}(\phi)$, indem gezeigt wird, dass $\phi(v_{l+1}), \dots, \phi(v_n)$ eine Basis ohne Doppelnennungen von $\mathrm{Bild}(\phi)$ ist.

[8] Für endlich-dimensionales V; der Beweis funktioniert mit den entsprechenden Modifikationen aber auch für unendlich-dimensionale Vektorräume.

Sei $w = \phi(a_1 v_1 + \cdots + a_n v_n) \in \mathrm{Bild}(\phi)$. Dann ist $w = a_1\phi(v_1) + \cdots + a_n\phi(v_n) = a_{l+1}\phi(v_{1+1}) + \cdots + a_n\phi(v_n)$, da $\phi(v_1) = \cdots = \phi(v_l) = 0$. Also ist $\phi(v_{l+1}), \ldots, \phi(v_n)$ ein Erzeugendensystem von $\mathrm{Bild}(\phi)$.

Zu zeigen bleibt die lineare Unabhängigkeit, mit Lemma 4: Sei also $0 = b_{l+1}\phi(v_{l+1}) + \cdots + b_n\phi(v_n) = \phi(b_{l+1}v_{l+1} + \cdots + b_n v_n) \in \mathrm{Kern}(\phi)$. Da $v_1, \ldots, v_l$ eine Basis von $\mathrm{Kern}(\phi)$ ist, gibt es $b_1, \ldots, b_l$ mit $b_{l+1}v_{l+1} + \cdots + b_n v_n = b_1 v_1 + \cdots + b_l v_l$, oder $(-b_1)v_1 + \cdots + (-b_l)v_l + b_{l+1}v_{l+1} + \cdots + b_n v_n = 0$. Aus der linearen Unabhängigkeit der Basis $v_1, \ldots, v_n$ folgt nun aber $b_1 = \cdots = b_n = 0$.

Satz 26 Wenn $\phi : V \to W$ bijektiv ist, dann gilt $\dim V = \dim W$. Wenn $\dim V = \dim W$ endlich ist, dann ist ϕ genau dann injektiv, wenn surjektiv (und damit genau dann, wenn bijektiv).

Beweis zu Dimension und bijektive Abbildungen:

Wenn $\phi : V \to W$ bijektiv ist, so ist $\dim \mathrm{Kern}(\phi) = 0$, da ϕ injektiv, und $\dim \mathrm{Bild}(\phi) = W$, da ϕ surjektiv, also $\mathrm{Bild}(\phi) = W$. Es folgt $\dim V = \dim \mathrm{Kern}(\phi) + \dim \mathrm{Bild}(\phi) = 0 + \dim W$.

Wenn $\dim V = \dim W$ endlich und ϕ injektiv, dann ist $\dim W = \dim V = \dim \mathrm{Kern}(\phi) + \dim \mathrm{Bild}(\phi) = 0 + \dim \mathrm{Bild}(\phi)$, also ϕ surjektiv.

Wenn $\dim V = \dim W$ endlich und ϕ surjektiv, dann ist $\dim \mathrm{Kern}(\phi) = \dim V - \dim \mathrm{Bild}(\phi) = \dim W - \dim \mathrm{Bild}(\phi) = 0$, also ϕ injektiv.

Definition: Lineares Gleichungssystem

Ein *lineares Gleichungssystem* (über einem Körper K, meist $K = \mathbb{R}$) besteht aus linearen Gleichungen

$$\begin{array}{ccccccc} a_{11} \cdot x_1 & + & a_{12} \cdot x_2 & + \cdots + & a_{1n} \cdot x_n & = & b_1 \\ & & & \vdots & & \vdots & \\ a_{m1} \cdot x_1 & + & a_{m2} \cdot x_2 & + \cdots + & a_{mn} \cdot x_n & = & b_m \end{array}$$

mit $a_{ij}, b_i \in K$ und Unbekannten $x_1, \ldots, x_n$. Eine *Lösung* des Gleichungssystems besteht aus Werten $k_1, \ldots, k_n \in K$, welche gleichzeitig alle m Gleichungen erfüllen.

Das zugehörige *homogene (lineare) Gleichungssystem* ist

$$\begin{array}{ccccccc} a_{11} \cdot x_1 & + & a_{12} \cdot x_2 & + \cdots + & a_{1n} \cdot x_n & = & 0 \\ & & & \vdots & & \vdots & \\ a_{m1} \cdot x_1 & + & a_{m2} \cdot x_2 & + \cdots + & a_{mn} \cdot x_n & = & 0 \end{array}.$$

Erläuterung

Offenbar kann man das lineare Gleichungssystem in einer Matrix zusammenfassen als

$$A \cdot x = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = b$$

Eine Lösung des homogenen Gleichungssystems $A \cdot x = 0$ ist dann ein Vektor aus dem Kern von A; die *Lösungsmenge* des homogenen Gleichungssystems (d. h. die Menge aller Lösungen) ist genau $\mathrm{Kern}(A)$.

Erläuterung

Für das allgemeine Gleichungssystem $A \cdot x = b$ gibt es die beiden bereits besprochenen Möglichkeiten: Entweder $b \notin \mathrm{Bild}(A)$ und es gibt keine Lösung, oder $b \in \mathrm{Bild}(A)$ und die Lösungsmenge besteht aus einer Nebenklasse $c+\mathrm{Kern}(A)$, wobei c irgendeine Lösung des Gleichungssystems ist. Um die Lösungsmenge des Gleichungssytems zu bestimmen, muss man also eine sogenannte *spezielle Lösung* c finden – sofern sie existiert! – und den Kern von A bestimmen. Ist $v_1, \dots, v_l$ eine Basis des Kerns, so besteht die Lösungsmenge also aus allen Vektoren der Form $c + k_1 v_1 + \dots + k_l vl$ mit $k_i \in K$ („die *allgemeine Lösung*“).

Definition: Rang einer Matrix

Der *Rang* einer $(m \times n)$-Matrix A, $\mathrm{rg}(A)$, ist die Dimension des Bildes von A als linearer Abbildung $K^n \to K^m$, d. h. die Dimension des von den Spalten $A \cdot e_1, \dots, A \cdot e_n$ von A erzeugten Unterraums.

Satz 27 Nach Definition ist $\mathrm{rg}(A) \leqslant m$ und $= m$ genau dann, wenn A surjektiv ist. Außerdem gilt $n = \dim \mathrm{Kern}(A) + \mathrm{rg}(A)$ nach Satz 25.

Satz 28 Falls A die Matrix eines homogenen linearen Gleichungssystemes mit m Gleichungen und n Unbekannten ist, dann ist die Dimension des Lösungsraums $n - \mathrm{rg}(A)$.

2.7.1. Das Gauß-Verfahren zum Lösen linearer Gleichungssysteme

Erläuterung

Die Idee des Verfahrens besteht darin, das Gleichungssystem bzw. die Matrix durch eine Reihe „elementarer Umformungen“, die die Lösungsmenge nicht oder in einer kontrollierten Weise ändern, in eine „schöne Form“ zu bringen, aus der man die Lösungsmenge leicht errechnen kann.

Definition: elementare Umformungen
Hier betrachten wir drei Arten von elementaren Umformungen. Jede der Umformungen entspricht der Multiplikation von links mit einer invertierbaren Matrix M. Dann gilt

$$A \cdot v = M^{-1}MA \cdot v = b \iff MA \cdot v = M \cdot b$$

und wegen $M \cdot 0 = 0$ ist insbesondere $\mathrm{Kern}(MA) = \mathrm{Kern}(A)$, d. h. die Lösungsmenge ändert sich nicht, wenn A und c gleichermaßen umgeformt werden.

(1) *Vertauschung der i-ten mit der j-ten Gleichung bzw.*
Vertauschung der i-ten mit der j-ten Zeile der Matrix A und des Vektors b.

Dies entspricht der Multiplikation von links mit der Matrix $M_{(ij)} = M_{(ij)}^{-1}$ (siehe Seite 45).

(2) *Addition des k-fachen der j-ten Gleichung zur i-ten Gleichung bzw.*
Addition des k-fachen der j-ten Zeile der Matrix A und des Vektors b zur i-ten Zeile.

Dies entspricht der Multiplikation von links mit der Matrix $E_{ij}(k) = I_m + k \cdot E_{ij}$. ($E_{ij}$ ist die Standardbasenmatrix aus Satz 19). Man sieht leicht ein, dass $E_{ij}(k)^{-1} = E_{ij}(-k)$.

(3) *Multiplikation der i-ten Gleichung mit $k \neq 0$ bzw.*
Multiplikation der i-ten Zeile der Matrix A und des Vektors b mit $k \neq 0$.

Dies entspricht der Multiplikation von links mit der Matrix $E_i(k) = I_m + (k-1) \cdot E_{ii}$. Man sieht wiederum leicht ein, dass $E_i(k)^{-1} = E_i(k^{-1})$.

Erläuterung
Ergänzend können auch Operationen auf den Spalten der Matrix vorgenommen werden (z. B. Vertauscheungen der Spalten, die dann den entsprechenden Vertauschungen der Unbekannten entsprechen). Diese sind aber nur zur Verbesserung von Algorithmen hinsichtlich Stabilität notwendig.

Definition: Zeilenstufenform einer Matrix, Pivot-Elemente
Eine Matrix A ist in *Zeilenstufenform*, falls es $j_1 < \cdots < j_r$ gibt, so dass $a_{1j_1} \neq 0, \ldots, a_{rj_r} \neq 0$ und

$$a_{ij} = 0, \text{ falls } \begin{cases} i > k \text{ und } j \leqslant j_k \\ \text{oder } j < j_1 \\ \text{oder } i > r \end{cases}.$$

Die Elemente a_{ij_i} heißen *Pivot-Elemente*, die Spalten $j_1, \ldots, j_r$ *Pivot-Spalten.*

Erläuterung
Schematisch angedeutet sieht eine Zeilenstufenform (mit $r = 3$) wie folgt aus; $*$ steht für

beliebige Elemente:

$$\begin{pmatrix} 0 & a_{1j_1} & * & * & * & * & * \\ 0 & 0 & a_{2j_2} & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & a_{3j_3} & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Satz 29 (a) Jede Matrix kann durch elementare Umformungen der Art (1) und (2) in Zeilenstufenform gebracht werden.

(b) Jede invertierbare Matrix kann durch elementare Umformungen der Art (1) und (2) in eine Diagonalmatrix und durch elementare Umformungen der Art (1), (2) und (3) in die Identitätmatrix überführt werden.

Beweis zu Mächtigkeit der elementaren Umformungen (Gauß-Verfahren, Gauß-Jordan-Verfahren):

Für (a) gibt es den in Abbildung 2.2 dargestellten Algorithmus, das sogenannte *Gauß-Verfahren*. Die Matrix wird spaltenweise von links nach rechts und zeilenweise von oben nach unten so abgearbeitet, dass die gewünschten Nullen auftreten. Betrachtet wird immer nur der Teil unterhalb der aktuellen Stelle: Ein eventuell vorhandener Eintrag $\neq 0$ in der Spalte wird ggf. durch Zeilenvertauschung an die betrachtete Stelle gebracht; durch die Addition eines passenden Vielfachens der Zeile werden unterhalb der betrachteten Stelle Nullen erzeugt. (Ein formaler Korrektheitsbeweis unterbleibt hier).

Input: $(m \times n)$-Matrix A. Setze $i = 1, j = 1$.

$\boxed{\text{Ist } a_{ij} \neq 0 \text{ ?}}$

Ja: ersetze Zeile $k = i+1, \dots, n$ durch „Zeile $k - \frac{a_{kj}}{a_{ij}} \cdot$ Zeile i“
ersetze i durch $i+1$ und j durch $j+1$

Nein: $\boxed{\text{Gibt es ein } l > i \text{ mit } a_{lj} \neq 0 \text{ ?}}$

Ja: vertausche Zeile i mit Zeile l (für das kleinste l)

Nein: ersetze j durch $j+1$

Falls $i > m$ oder $j > n$: Ende

Abbildung 2.2.: Gauß-Verfahren

(b) Durch das Gauß-Verfahren bringt man zunächst die Matrix in Zeilenstufenform. Sie ist genau dann invertierbar, wenn die Zeilenstufenform Dreiecksform hat, d. h. wenn die

Pivot-Elemente die Diagonalelemente $a_{11}, \ldots, a_{nn}$ sind. Man kann auf diese Matrix nun das Gauß-Verfahren gewissermaßen „punktgespiegelt“, also spaltenweise von rechts nach links und zeilenweise von unten nach oben anwenden, und erhält eine Diagonalmatrix (d. h. $a_{ii} \neq 0$, aber $a_{ij} = 0$ für alle $i \neq j$.) Durch Umformungen der Art (3) kann man schließlich die Diagonaleinträge auf 1 bringen. diese Verfahren heißt manchmal auch *Gauß-Jordan-Verfahren.*

Input: $(n \times n)$-Matrix A. Führe zunächst das Gauß-Verfahren durch.

$\boxed{\text{Sind alle } a_{ii} \neq 0 \text{ ?}}$

→ Nein: die Matrix ist nicht invertierbar

→ Ja: Setze $j = n$

ersetze Zeile $k = 1, \ldots, j-1$ durch „Zeile $k - \frac{a_{kj}}{a_{jj}} \cdot$ eile j“

ersetze Zeile j durch „Zeile $\frac{1}{a_{jj}} \cdot$ Zeile j“

ersetze j durch $j-1$

Falls $j < 1$: Ende

Abbildung 2.3.: Gauß–Jordan-Verfahren

Beispiele

(Fehlender Inhalt: Beispiel)

Satz 30 Was kann mit dem Gauß-Verfahren berechnet werden?

Sei stets E eine invertierbare Matrix, die A in Zeilenstufenform bringt, d. h. E ist eine Matrix $E_k \cdot \ldots \cdot E_1$, wobei $E_1, \ldots, E_k$ Matrizen zu elementaren Umformungen sind, welche nach dem Gauß-Verfahren A in eine Matrix $E_k \cdot \ldots \cdot E_1 \cdot A$ in Zeilenstufenform umformen.

- *Den Rang einer Matrix berechnen:*

  Der Rang der Matrix ist die Anzahl der Pivot-Elemente in der Zeilenstufenform.

- *Testen, ob eine Matrix invertierbar ist:*

  Eine Matrix ist genau dann invertierbar, wenn sie quadratisch ist und der Rang mit der Anzahl der Zeilen/Spalten übereinstimmt.

- *Eine spezielle Lösung eines linearen Gleichungssystems ausrechnen:*

  Man bringt das Gleichungssystem in Zeilenstufenform $EA \cdot x = E \cdot b$ und löst die Gleichungen von unten nach oben auf („Rückwärteinsetzen"). Sind Unbekannte durch eine Gleichung und die vorherigen Festsetzungen nicht eindeutig bestimmt, setzt man einen beliebigen Wert (z. B. 0) ein.

- *Eine Basis des Kerns bestimmen:*

  Man bringt das homogene Gleichungssystem in Zeilenstufenform $EA \cdot x = E \cdot 0 = 0$. Ist der Rang der Matrix gleich n (= Anzahl der Unbekannten), so ist Kern = $\{0\}$, die Basis also die leere Menge. Andernfalls löst man die Gleichungen von unten nach oben durch Rückwärteinsetzen auf. Für jede Unbekannte, die nicht eindeutig festgelegt ist, bekommt man einen Basisvektor des Kerns, indem man diese Unbekannte auf 1 setzt und alle andern dann nicht festgelegten Unbekannten auf 0.

- *Eine Basis des Bilds bestimmen:*

  Spalten $Ae_{i_1}, \ldots, Ae_{i_l}$ von A sind genau dann linear unabhängig, wenn die entsprechenden Spalten $EAe_{i_1}, \ldots, EAe_{i_l}$ der Matrix in Zeilenstufenform linear unabhängig sind. Also bilden die Spalten von A, die Pivot-Spalten von EA sind, eine Basis des Bildes.

- *Eine Menge linear unabhängiger Vektoren zu einer Basis ergänzen:*

  Man fügt die Vektoren als Spalten zu einer $(m \times n)$-Matrix A zusammen und bestimmt eine Basis des Bildes der $(m \times (n+m))$-Matrix $(A \mid I_m)$ wie oben beschrieben.

  alternativ: Man fügt die Vektoren als Zeilen zu der $(n \times m)$-Matrix A^{T} zusammen und bringt sie in Zeilenstufenform. Diejenigen Standardbasisvektoren e_i, für die i keine Pivot-Spalte ist, ergänzen die gegebenen Vektoren zu einer Basis.

- *Das Inverse einer Matrix berechnen:*

  Die elementaren Umformungen, welche A nach dem Gauß-Jordan-Verfahren in die Identitätsmatrix umformen, formen gleichzeitig die Identitätsmatrix in die Inverse von A um: falls $E \cdot A = I_n$, so ist $E \cdot I_n = A^{-1}$.

Mathematische Folgerungen

Definition: Transponierte Matrix
Die *Transponierte* A^{T} einer $(m \times n)$-Matrix $A = (a_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}}$ ist die „an der Diagonalen gespiegelte“ $(n \times m)$-Matrix $(a_{ji})_{\substack{j=1,\ldots,m\\ i=1,\ldots,n}}$.

Beispiele

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad A^{\mathrm{T}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Satz 31 Man sieht auch leicht aus der Multiplikationsformel, dass $(A \cdot B)^{\mathrm{T}} = B^{\mathrm{T}} \cdot A^{\mathrm{T}}$. Insbesondere ist die Transponierte einer invertierbaren Matrix selbst invertierbar mit $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$.

Satz 32 $\mathrm{rg}(A) = \mathrm{rg}(A^{\mathrm{T}})$.

Beweis zu Rang der transponierten Matrix:
Man sieht, dass der Satz für Matrizen in Zeilenstufenform gilt. Da Isomorphismen die Dimension bewahren, ändert die Multiplikation von rechts oder links mit einer invertierbaren Matrix nicht den Rang einer Matrix. Sei also $E \cdot A$ in Zeilenstufenform für ein invertierbares E. Dann gilt: $\mathrm{rg}(A) = \mathrm{rg}(E \cdot A) = \mathrm{rg}((E \cdot A)^{\mathrm{T}}) = \mathrm{rg}(A^{\mathrm{T}} \cdot E^{\mathrm{T}}) = \mathrm{rg}(A^{\mathrm{T}})$.

Beweis zu Beweis Dimensionssatz (alternativ):
Aus dem Gauß-Verfahren gewinnt man auch einen Beweis für Satz **??** im endlich-dimensionalen Fall. Allerdings müsste man sich noch davon überzeugen, dass der Satz für das Gauß-Verfahren nicht gebraucht wurde (und man muss aufpassen, dass man keine Begriffe oder Argumente verwendet, welche bereits auf der Dimension beruhen, wie z. B. den Rang).

Satz 33 Wenn ein Vektorraum V eine Basis mit endlich vielen Elementen besitzt, dann haben alle Basen von V die gleiche Anzahl von Elementen.

Beweis zu Größe der Basen:
Angenommen V hat Basen mit m und mit n Elementen, $m < n$. über die eine Basis ist V isomorph zu K^m; man kann also annehmen, dass $V = K^m$. Nun stellt man die $(m \times n)$-Matrix A auf, deren Spalten die Vektoren der Basis mit n Elementen sind. Diese sind

nach Annahme linear unabhängig, also müssen auch die Spalten der in Zeilenstufenform gebrachten Matrix linear unabhängig sein. In der Zeilenstufenform sieht man aber, dass maximal m Spalten linear unabhängig sein können: Widerspruch.

2.8. Determinanten

Erläuterung

Idee: Wie kann man die Volumenänderung auf einen Quader durch eine lineare Abbildung $\mathbb{R}^3 \to \mathbb{R}^3$ messen?

mathematische Vorgehensweise in diesem Fall: Man möchte das Problem mit einem neuen Konstrukt, den Determinanten lösen. Man stellt die Eigenschaften fest, die die Determinante (= orientierte Volumenänderung) haben soll und stellt fest, dass es nur eine Funktion mit diesen Eigenschaften geben kann. Dann kann man Formeln angeben.

Notation: Determinante

Die Determinante $det(A)$ einer Matrix A wird auch dadruch beschrieben, dass man bei der Angabe der Matrix die äußeren Klammern durch senkrechte Striche ersetzt, also

$$\det(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

Satz 34 Eigenschaften der Determinante (A sei eine $n \times n-$Matrix):

- $det(A) = 0 \Leftrightarrow A$ nicht invertierbar
- $det(id) = 1$
- $det(A \cdot B) = det(A) \cdot det(B)$
- Vertauscht man in einer Matrix die Zeilen i und j oder die Spalten i und j, so ist die zugehörige Determinante das negative der Determinante der ursprünglichen Matrix.
- Addiert man auf eine Spalte oder Zeile der Matrix eine Zeile oder Spalte so kann man das Ergebnis auf als die Addition der beiden Determinanten der Matrizen, die sich nur in der entsprechenden Spalte/Zeile unterscheiden berechnen: z.B.. $det(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}) = det(\begin{pmatrix} 1 & 1 & 3 \\ 4 & 1 & 6 \\ 7 & 1 & 9 \end{pmatrix}) + det(\begin{pmatrix} 1 & 1 & 3 \\ 4 & 4 & 6 \\ 7 & 7 & 9 \end{pmatrix})$
- Multipliziert man eine Zeile/Spalte einer Matrix mit einem Skalar kann die Determinante auch als Multiplikation der ursprünglichen Determinante mit dem Skalar berechnet werden.
- $det(A) = det(A^T)$
- $det(A \cdot B) = det(A) \cdot det(B)$

Satz 35 Berechnung der Determinante

- für kleine n lässt sich eine einfache Formel angeben:
  - $n = 1$: $det(a_{11}) = a_1 1$
  - $n = 2$: $det(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}) = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}$
  - $n = 3$: $det(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix})$
    $= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$
- Formel von Leibniz (A quadratisch) $det(A) = \sum\limits_{\sigma = Sym(\{1,\dots,n\})} sgn(\sigma) \cdot a_{1\sigma(1)} \cdot a_{1\sigma(2)} \cdots a_{n\sigma(n)}$
- Laplace'scher Entwicklungssatz $det(A) = \sum\limits_{j=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij})$, wobei A_{ij} die $(n-1) \times (n-1)-$Matrix ist, die aus A durch Streichen der $i-ten$ Zeile und $j-ten$ Spalte entsteht.
- Mit Gaußverfahren auf Dreiecksmatrix bringen und Diagonale multiplizieren (Bei Vertauschen von zwei Zeilen ändert sich das Vorzeichen), Addition des k-fachen einer Zeile auch eine andere ändert die Determinante nicht.
- **(Fehlender Inhalt: Formel für Inverse)**

2.9. Längen, Winkel, Skalarprodukt

Erläuterung

In diesem Abschnitt soll stets $K = \mathbb{R}$ sein; alle betrachteten Vektorräume seien über $\mathbb{R}$. Es sollen nun die geometrisch anschaulichen Begriffe der Länge eines Vektors, des Abstandes zweier Vektoren und des Winkels zwischen zwei Vektoren eingeführt werden. Dabei ist das Vorgehen – ähnlich wie schon bei der Determinante – wie folgt: Man findet eine Formel für die Berechnung, die in den Fällen der Dimension 1, 2 und 3 das Richtige tut und die Eigenschaften besitzt, die man von den Begriffen erwartet. In den höherdimensionalen Fällen, wo eine direkte geometrische Anschauung fehlt, definiert man die Begriffe dann durch diese Formel.

Länge und Abstand

Definition: Länge, Abstand
Die *Länge* eines Vektors $v = (v_1, \ldots, v_n) \in \mathbb{R}^n)$ ist

$$\|v\| := \sqrt{v_1^2 + \cdots + v_n^2}.$$

Der *Abstand (*oder die *Distanz)* zweier Vektoren ist

$$\mathrm{d}(v, w) := \|v - w\|.$$

Erläuterung

Es gilt also $\|v\| = d(v, 0)$. Im $\mathbb{R}^1$ ist $\|v\| = |v|$; in $\mathbb{R}^2$ und $\mathbb{R}^3$ sieht man mit dem Satz von Pythagoras, dass die Definition den gewöhnlichen Längenbegriff wiedergibt.

Satz 36 Eigenschaften von Länge und Abstand Für alle $u, v, w \in \mathbb{R}^n$ und $k \in \mathbb{R}$ gilt:

Positivität:	$\|v\| \geq 0$	$\mathrm{d}(v,w) \geq 0$
	$\|v\| = 0 \Leftrightarrow v = 0$	$\mathrm{d}(v,w) = 0 \Leftrightarrow v = w$
Symmetrie:	$\|v\| = \|-v\|$	$\mathrm{d}(v,w) = \mathrm{d}(w,v)$
Dreiecksungleichung:	$\|v + w\| \leq \|v\| + \|w\|$	$\mathrm{d}(v,w) \leq \mathrm{d}(v,u) + \mathrm{d}(u,w)$
Skalierung:	$\|r \cdot v\| = \lvert r\rvert \cdot \|v\|$	$\mathrm{d}(r \cdot v, r \cdot w) = \lvert r\rvert \cdot \mathrm{d}(v,w)$

Erläuterung

Neben diesem gewöhnlichen Längenbegriff (der auch „euklidische Norm“ oder „2-Norm“ $\|v\|_2$ genannt wird), gibt es im Mehrdimensionalen auch weitere Längenbegriffe, etwa die „1-Norm“ $\|v\|_1 = |v_1| + \cdots + |v_n|$ oder die „Maximumsnorm“ $\|v\|_\infty := \max\big\{|v_1|, \ldots, |v_n|\big\}$, die ebenfalls alle oben aufgeführten Eigenschaften aufweisen.

Skalarprodukt, Winkel, Orthogonalität

Erläuterung

Der Winkel zwischen zwei Vektoren wird üblicherweise über das Skalarprodukt ausgerechnet. Das Skalarprodukt selbst misst keine ganz elementare geometrische Größe wie Länge oder Winkel, sondern beides in Kombination. Im $\mathbb{R}^2$ wird das Skalarprodukt von $v = (v_1, v_2)$ und $w = (w_1, w_2)$ durch die Formel $\langle v, w\rangle = v_1 w_1 + v_2 w_2$ berechnet; die geometrische Interpretation dieser Größe ist: „$\|v\|$ mal $\|w\|$ mal Cosinus des Winkels zwischen v und w“.

Da in diesem Fall die geometrische Interpretation der Formel viel weniger ersichtlich ist als bei der Länge von Vektoren, soll sie auf zwei Arten erklärt werden (die zwar im wesentlichen übereinstimmen, aber Verschiedenes voraussetzen).

Erläuterung

Erste Methode Da es bei dem Winkel nicht auf die Längen der Vektoren ankommt, kann man o.E. annehmen, dass v und w Länge 1 haben (indem man sie durch $\frac{v}{\|v\|}$ bzw. $\frac{w}{\|w\|}$ ersetzt; den Fall $v = 0$ oder $w = 0$ kann man außer Acht lassen, da $\langle 0, w\rangle = \langle v, 0\rangle = 0$). Falls $v = e_1 = (1, 0)$, so ist w_1 gerade der Cosinus des eingeschlossenen Winkels zwischen e_1 und w (und w_2 ist die (orientierte) Höhe der von e_2 und w aufgespannten Raute, also im wesentlichen deren Flächeninhalt, da die Grundseite e_1 Länge 1 hat). Winkel sollten unter Drehungen invariant sein; man kann daher den allgemeinen Fall auf diesen speziellen Fall durch die Drehung von v auf e_1 zurückführen. Also ist der Cosinus des Winkels zwischen v und w die erste Koordinate von

$$\begin{pmatrix} v_1 & v_2 \\ -v_2 & v_1 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_1 w_1 + v_2 w_2 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Man sieht auch, dass die zweite Komponente die orientierte Höhe der Fläche der von v und w aufgespannten Raute ist, also die Volumenveränderung der Abbildung $\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$ angibt. Also stimmt neben der Formel für das Skalarprodukt auch die Determinantenformel im $\mathbb{R}^2$.

Erläuterung

Zweite Methode Geht man von der gewünschten geometrischen Interpretation des Skalarprodukts aus, so ist klar, dass $\langle v, k \cdot v\rangle = k \cdot \|v\|^2$ sein muss und dass $\langle v, w\rangle = 0$ gelten muss, wenn v und w senkrecht aufeinander stehen. Sicher steht $v = (v_1, v_2)$ senkrecht auf $(-v_2, v_1)$ und allen seinen Vielfachen. Nun schreibt man (nachrechnen durch Ausmultiplizieren!)

$$(w_1, w_2) = \frac{v_1 w_1 + v_2 w_2}{v_1^2 + v_2^2} \cdot (v_1, v_2) + \frac{v_1 w_2 - v_2 w_1}{v_1^2 + v_2^2} \cdot (-v_2, v_1).$$

Der linke Summand gibt dann gerade die orthogonale Projektion von w auf v an; die Länge dieses Vektors mal die Länge von v ist dann gerade $v_1 w_1 + v_2 w_2$.
ähnliche überlegungen kann man für den $\mathbb{R}^3$ anstellen (oder man führt, indem man die beiden Vektoren zunächst in die $\{e_1, e_2\}$-Ebene dreht, den dreidimensionalen auf den zweidimensionalen Fall zurück).

Definition: Standardskalarprodukt
Das *(Standard-)Skalarprodukt* im Vektorraum $\mathbb{R}^n$ ist die folgende Abbildung $\langle \cdot, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$:

$$\langle v, w \rangle := (v_1, \ldots, v_n) \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_2 \end{pmatrix} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i.$$

Satz 37 Eigenschaften des Skalarprodukts Für alle $v, v', w, w' \in \mathbb{R}^n$ und $r \in \mathbb{R}$ gilt:

$$\begin{array}{rll} \text{Positivität:} & \langle v, v \rangle = \|v\|^2 = \geq 0 & \\ & \langle v, v \rangle = 0 \Leftrightarrow v = 0 & \\ \text{Symmetrie:} & \langle v, w \rangle = \langle w, v \rangle & \\ \text{Bilinearität:} & \langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle & \langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle \\ & \langle r \cdot v, w \rangle = r \cdot \langle v, w \rangle & \langle v, r \cdot w \rangle = r \cdot \langle v, w \rangle \end{array}$$

d. h. das Skalarprodukt ist sowohl im ersten als auch im zweiten Argument eine lineare Abbildung.

Satz 38[Cauchy-Schwarz[9]] Seien $v, w \in \mathbb{R}^n$, dann gilt

$$|\langle v, w \rangle| \quad \leq \quad \|v\| \cdot \|w\|$$

oder (quadriert)

$$\left(\sum_{i=1} v_i w_i \right)^2 \leq \sum_{i=1} v_1^2 \cdot \sum_{i=1} w_1^2.$$

Satz 39 Für $v \neq 0$ und $w \neq 0$ gilt

$$-1 \;\leq\; \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \;=\; \left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \right\rangle \;\leq\; 1;$$

somit findet man einen eindeutigen Winkel $\alpha \in [0, \pi]$ mit

$$\langle v, w\rangle = \|v\| \cdot \|w\| \cdot \cos(\alpha).$$

Per Definition nennt man α den *zwischen v und w eingeschlossenen Winkel* $\angle(v, w)$.[10]

Beweis zu eingeschlossener Winkel:
Der Fall $w = 0$ ist klar (beide Seiten ergeben 0); sei also $w \neq 0$. Dann ist

$$\begin{aligned}
0 \;&\leq\; \left\langle v - \frac{\langle v, w\rangle}{\|w\|^2} \cdot w, v - \frac{\langle v, w\rangle}{\|w\|^2} \cdot w \right\rangle \\
&=\; \langle v, v\rangle - 2 \cdot \frac{\langle v, w\rangle}{\|w\|^2} \langle v, w\rangle + \frac{\langle v, w\rangle^2}{\|w\|^4} \langle w, w\rangle \\
&=\; \frac{1}{\|w\|^2} \Big(\langle v, v\rangle \cdot \langle w, w\rangle - 2 \cdot \langle v, w\rangle^2 + \langle v, w\rangle^2 \Big) \;=\; \frac{1}{\|w\|^2} \Big(\langle v, v\rangle \cdot \langle w, w\rangle - \langle v, w\rangle^2 \Big)
\end{aligned}$$

Daraus folgt also $0 \leq \langle v, v\rangle \cdot \langle w, w\rangle - \langle v, w\rangle^2$ bzw. $\langle v, w\rangle^2 \leq \langle v, v\rangle \cdot \langle w, w\rangle = \|v\|^2 \cdot \|w\|^2$, also nach Wurzelziehen das gewünschte Ergebnis.

Satz 40 Insbesondere gilt also:

$$\begin{aligned}
\langle v, w\rangle = 0 \iff \cos\angle(v, w) \text{ ist } \frac{\pi}{2} \text{ oder } \frac{3}{2}\pi \quad (\text{d. h. } 90^\circ \text{ oder } 270^\circ) \\
\iff v \text{ und } w \text{ stehen } \textit{senkrecht} \text{ aufeinander.}
\end{aligned}$$

Erläuterung
In den Dimensionen 1, 2 und 3 stimmt dies also mit dem anschaulichen geometrischen Begriff überein; in den höheren Dimensionen ist es eine sinnvolle Verallgemeinerung. In abstrakten n-dimensionalen Räumen gibt es dagegen kein Standard-Skalarprodukt, also auch keinen natürlichen Winkelbegriff. Durch die Wahl einer Basis kann man aber das Skalarprodukt des $\mathbb{R}^n$ übertragen. Das Standard-Skalarprodukt geht axiomatisch davon aus, dass die Standardbasis eine sogenannte *Orthonormalbasis* ist, also die Basisvektoren Länge 1 haben und paarweise aufeinander senkrecht stehen. Darauf beruhen alle weiteren

[9] Augustin Louis Cauchy (1789-1857), Hermann Amandus Schwarz (1843-1921)
[10] Dieser Begriff ist nicht orientiert, d. h. der Winkel zwischen v und w ist gleich dem Winkel zwischen w und v. Dem entspricht, dass der Cosinus eine gerade Funktion ist, also $\cos(\alpha) = \cos(-\alpha)$ ist.

Längen- und Winkelbestimmungen. Die übertragung des Standard-Skalarprodukts des $\mathbb{R}^n$ auf einen abstrakten n-dimensionalen Vektorraum durch Wahl einer Basis bedeutet, dass man diese Basis zur Orthonormalbasis erklärt.

Satz 41 Sei $v \neq 0$, dann ist die *orthogonale Projektion* von w auf v gleich

$$w_v = \frac{\langle w, v\rangle}{\|v\|} \cdot \frac{v}{\|v\|} = \frac{\langle w, v\rangle}{\langle v, v\rangle} \cdot v.$$

Wenn $v_1, \ldots, v_n$ eine Orthonormalbasis ist, dann gilt

$$w = \sum_{i=1}^{n} \langle w, v_i\rangle \cdot v_i.$$

Beweis zu Orthonormalbasis:
Wenn man $w = \sum_{i=1}^{n} r_i v_i$ ansetzt und $\langle w, v_j\rangle = \langle \sum_{i=1}^{n} r_i v_i, v_j\rangle$ mit Hilfe der Bilinearität des Skalarprodukts ausrechnet, ergibt sich unmittelbar $r_j = \langle w, v_j\rangle$. Der erste Teil folgt aus der geometrischen Interpretation des Skalarprodukts (bzw. durch Skalieren und Ergänzen von v zu einer Orthonormalbasis).

Satz 42[Verallgemeinerter Satz des Pythagoras[11]; Cosinussatz]
Für $v, w \in \mathbb{R}^n$ gilt:

$$\|v + w\|^2 = \|v\|^2 + 2 \cdot \langle v, w\rangle + \|w\|^2.$$

Insbesondere gilt $\|v + w\|^2 = \|v\|^2 + \|w\|^2$ genau dann, wenn $\langle v, w\rangle = 0$, also wenn v und w senkrecht aufeinander stehen.

Beweis zu Cosinussatz:
Einfach ausrechnen:

$$\|v + w\| = \sum_{i=1}^{n} (v_i + w_i)(v_i + w_i) = \sum_{i=1}^{n} v_i^2 + \sum_{i=1}^{n} w_i^2 + 2 \cdot \sum_{i=1}^{n} v_i w_i = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2$$

Orthogonale Abbildungen

Definition: orthogonale lineare Abbildung

[11] Pythagoras (ca. 570 bis ca. 510 v. Chr.)

Eine lineare Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^n$ heißt *orthogonal*, wenn ϕ das Skalarprodukt erhält, wenn also $\langle \phi(v), \phi(w) \rangle = \langle v, w \rangle$ für alle $v, w \in \mathbb{R}^n$ gilt.[12]
Eine $(n \times n)$-Matrix A heißt orthogonal, wenn die zugehörige lineare Abbildung orthogonal ist.

Definition: Orthonormalbasis
Eine Basis $v_1, \ldots, v_n$ des $\mathbb{R}$ ist eine *Orthonormalbasis*, wenn

$$\langle v_i, v_j \rangle := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

Erläuterung
Man rechnet leicht nach, dass

$$\langle Av, w \rangle = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} v_i \right) \cdot w_j = \sum_{i=1}^{n} v_i \cdot \left(\sum_{j=1}^{n} a_{ji} w_j \right) = \langle v, A^{\mathrm{T}} w \rangle.$$

Satz 43 Die folgenden Aussagen sind äquivalent für eine $(n \times n)$-Matrix über $\mathbb{R}^n$:
(a) A ist orthogonal;
(b) A ist invertierbar und $A^{-1} = A^{\mathrm{T}}$;
(c) $Ae_1, \ldots, Ae_n$ ist eine Orthonormalbasis.

Beweis zu Orthoganalität und Matrizen:
Klar ist, dass eine orthogonale Abbildung eine Orthonormalbasis auf eine Orthonormalbasis abbilden muss, also gilt (a)$\Rightarrow$(c). Der (i, j)-Eintrag von $A^{\mathrm{T}} \cdot A$ ist genau $\langle Ae_i, Ae_j \rangle$, also ist $Ae_1, \ldots, Ae_n$ genau dann eine Orthonormalbasis, wenn $A^{\mathrm{T}} \cdot A = \mathrm{Id}$, also wenn (b) gilt. Schließlich folgt aus (b), dass $\langle Av, Aw \rangle = \langle v, A^{\mathrm{T}} Aw \rangle = \langle v, w \rangle$, also dass A orthogonal ist.

Beispiele
Drehungen im $\mathbb{R}^2$ sind orthogonal: $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}^{-1} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$. Ebenso sind Spiegelungen orthogonal. Drehungen und Spiegelungen sind die einzigen orthogonalen

[12] Achtung: Die orthogonale Projektion aus Satz 41 ist keine orthogonale Abbildung.

Abbildungen der Ebene (dabei sind die Drehungen orientierungserhaltend, die Spiegelungen nicht).

Erläuterung

Orthogonale Abbildungen sind *längentreu*, d. h. $\|Av\| = \|v\|$ für alle v, und *winkeltreu*, d. h. $\angle(Av, Aw) = \angle(v, w)$ für alle v, w. Aus $A^{-1} = A^{\mathrm{T}}$ folgt $\det(A)^{-1} = \det(A^{\mathrm{T}}) = \det(A)$ und somit $\det A = \pm 1$. Orthogonale Abbildungen sind also zudem *volumentreu*, allerdings nur im unorientierten Sinn; die Orientierung kann sich ändern (wie man am Beispiel der Spiegelungen sieht).

Scherungen sind Beispiele von volumenerhaltenden Abbildungen, die weder längen- noch winkeltreu sind; Streckungen (aller Vektoren um den gleichen Faktor) sind Beispiele von winkeltreuen Abbildungen, die weder längen- noch volumentreu sind. Man kann aber zeigen, dass längentreue Abbildungen bereits orthogonal sind (dies folgt unmittelbar aus dem verallgemeinerten Satz von Pythagoras). Ebenso sind Abbildungen, die winkel- und volumentreu sind, schon orthogonal.[13]

[13] Winkelerhaltend heißt, dass $Ae_1, \dots, Ae_n$ eine Orthogonalbasis ist, also $A^{\mathrm{T}} \cdot A$ eine Diagonalbasis ist. Betrachtet man den Winkel zwischen e_i und $e_i + e_j$, sieht man schnell, dass alle Diagonaleinträge gleich sein müssen. Da die Abbildung Determinante ± 1 hat, müssen die Diagonaleinträge $= 1$ sein.

3. Lineare Codes

3.1. Codes

Einführung

In der Codierungstheorie geht es um folgende Problematik: Informationen werden als Folgen von Symbolen aufgeschrieben bzw. festgehalten. Man sagt dazu auch, dass die Informationen „codiert“ werden, z. B. durch Morse-Zeichen, durch Zahlenfolgen im ASCII-Code oder, wie in diesem Text hier, durch Symbolfolgen des um Satzzeichen angereicherten lateinischen Alphabets. Bei der Übermittlung von Nachrichten (z. B. Übertragung durch Funk oder Kabel oder Speicherung der Information über längere Zeiträume) können Übertragungsfehler passieren oder Teile der Information verloren gehen. Kann man die Codierung so wählen, dass eine gewisse Anzahl an Übertragungsfehlern erkannt und eventuell auch korrigiert werden können, und die Informationsübermittlung dennoch möglichst effizient geschieht?

Es geht also darum, in die Codierung eine Redundanz einzubauen. Die einfachste Art der Redundanz besteht darin, die Nachricht mehrfach zu wiederholen. Stimmen die empfangenen Informationen nicht überein, so weiß man, dass Übertragungsfehler eingetreten sein. Indem man gegebenenfalls die am häufigsten empfangene Version als die richtige ansieht, kann man u. U. auch Übertragungsfehler ausgleichen. Die Codierung durch Wiederholung ist aber insofern ineffizient, als sich die Länge der übermittelten Nachricht (und damit Zeit und Kosten) vervielfacht. Die Anforderung der Effizienz bezieht sich aber auch auf die Durchführung von Codierung, Decodierung und die eventuelle Fehlerkorrektur: hierfür sollen schnelle Algorithmen vorliegen.

Konkret betrachtet man folgende Situation:

Definition: Hamming-Raum
Man verfügt über ein endliches Alphabet (d. h. eine Symbolmenge) A mit q Elementen und betrachtet Wörter der festen Länge n über A, d. h. Elemente $(a_1, \ldots, a_n)$ von A^n, um Nachrichten zu codieren. Die Menge A^n dieser n-Tupel wird auch der *Hamming-Raum*[1] $H(n, A)$ genannt, bzw. $H(n, q)$, wenn es nur auf die Anzahl der Elemente von A ankommt.

Oft nimmt man als Alphabet eine endliche Gruppe oder einen endlichen Körper, etwa $\mathbb{F}_q$, da die algebraische Struktur beim Ver- und Entschlüsseln helfen kann und geschickte Codierungen ermöglicht. $H(n, \mathbb{F}_q) = \mathbb{F}_q^n$ ist dann ein n-dimensionaler Vektorraum über dem Körper $\mathbb{F}_q$. Besonders häufig ist der Fall $q = 2$ mit $\mathbb{F}_2 = \{0, 1\}$. Der Hamming-Raum

$H(8, \mathbb{F}_2)$ ist zum Beispiel die Menge der Bytes. Den Hamming-Raum $H(4, \mathbb{F}_2)$ kann man mit den hexadezimalen Ziffern identifizieren.

Beispiele

- Im ursprünglichen ASCII-Code wurden Zeichen durch ein Byte $(a_1, \ldots, a_8)$, also ein 8-Tupel über $\mathbb{F}_2$, codiert. Dabei bildeten die ersten sieben Ziffern $a_1, \ldots, a_7$ die eigentliche Information: als Binärzahl gelesen geben sie die Stelle des codierten Zeichens (Buchstabe, Ziffer, Satz- oder Steuerungszeichen) in der Liste der ASCII-Zeichen an. Die letzte Ziffer a_8 war eine Kontrollziffer, welche den sogenannten *parity check* durchführt: a_8 war so gewählt, dass $a_1 + \cdots + a_8 = 0$ in $\mathbb{F}_2$ gilt. Der Code „erkennt", wenn an einer Stelle ein Übertragungsfehler passiert, da dann die Prüfrechnung nicht mehr stimmt. Geht bei der Übertragung eine Stelle verloren, kann man sie errechnen.
- Der alte ISBN-Code bestand aus einer neunstelligen Dezimalzahl, die man als 9-Tupel $(b_1, \ldots, b_9)$ über $\mathbb{F}_{11}$ aufgefasst und um eine Prüfziffer $b_{10} \in \mathbb{F}_{11}$ so ergänzt hat, dass $\sum_{i=1}^{10} i \cdot b_i = 0$ in $\mathbb{F}_{11}$ gilt. (Das Element 10 in $\mathbb{F}_{11}$ wurde übrigens X geschrieben.)

  Dieser Code erkennt eine falsche Ziffer und auch Vertauschungen von zwei Ziffern, d. h. die Prüfrechnung stimmt dann nicht mehr.
- Der aktuelle ISBN-Code ist ein 13-Tupel über $\mathbb{Z}/10\mathbb{Z}$, wobei wieder die letzte Ziffer eine Prüfziffer ist, die so gewählt wird, dass $b_1 + 3b_2 + b_3 + 3b_4 + \cdots + b_{13} = 0$ in $\mathbb{Z}/10\mathbb{Z}$ gilt. Dieser Code erkennt wieder eine falsche Ziffer, aber nur noch gewisse Vertauschungen.
- Bei der neuen internationale Bankkontonummer IBAN folgen nach der anfänglichen Länderkennung (zwei Buchstaben) zwei Prüfziffern, die so gewählt sind, dass für eine gewisse, aus der IBAN gebildete Zahl z die Zahl $z - 1$ durch 97 teilbar ist. z entsteht aus der IBAN, indem man zunächst den Ländercode mit den Prüfziffern ans Ende setzt und dann die Buchstaben des Ländercodes durch zweistellige Zahlen ersetzt (A = 10, B = 11, …).

Fehler und die Hamming-Metrik

Anschaulich gesprochen ist ein Code gut, wenn er besonders viele Fehler erkennt oder sogar deren Korrektur zulässt. Um dies zu präzisieren, muss man festlegen, was Fehler sind und wie man ihre Anzahl misst. Im üblichen Setting legt man dazu fest, dass es nur um die Anzahl der Stellen geht, die nicht übereinstimmen. Eine Vertauschung von zwei (verschiedenen) Ziffern zählt also als zwei Fehler, da anschließend zwei Stellen nicht mehr stimmen. Insbesondere werden alle Stellen als gleichwertig gezählt (während man z. B. bei Dezimalzahlen Fehler in den höheren Stellen als gewichtiger ansehen würde als in den niederen Stellen) und alle Elemente des Alphabets werden ebenfalls untereinander als gleichwertig gezählt (d. h. es ist gleichermaßen ein einziger Fehler, ob z. B. 2 statt 1 empfangen wird oder 9 statt 1).

Mathematisch wird dies durch das Konzept der Hamming-Metrik präzisiert:

Definition: Hamming-Distanz/Hamming-Metrik
Für $v = (v_1, \dots, v_n)$ und $w = (w_1, \dots, w_n)$ in $H(n, A)$ definiert man den *Hamming-Abstand* (oder *Hamming-Metrik*) als

$$\mathrm{d}(v, w) := \big| \{i \mid v_i \neq w_i\} \big|.$$

Satz 44 $\mathrm{d}(\ ,\ )$ ist eine Metrik auf $H(n, A)$, d. h. es gilt für alle $u, v, w \in H(n, A)$:

- Positivität: $\mathrm{d}(v, w) \geqslant 0$ und $\big(\mathrm{d}(v, w) = 0 \iff v = w\big)$
- Symmetrie: $\mathrm{d}(v, w) = \mathrm{d}(w, v)$
- Dreiecksungleichung: $\mathrm{d}(u, v) \leqslant \mathrm{d}(u, v) + \mathrm{d}(v, w)$.

Falls $(A, +)$ eine Gruppe ist, gilt zusätzlich:

- Translationsinvarianz: $\mathrm{d}(v, w) = \mathrm{d}(v + u, w + u)$,
  insbesondere $\mathrm{d}(v, w) = \mathrm{d}(v - w, 0) = \mathrm{d}(-w, -v)$

Falls $A = K$ ein Körper und $H(n, A)$ ein K-Vektorraum ist, gilt außerdem:

- Invarianz unter Skalarmultiplikation: $\mathrm{d}(v, w) = \mathrm{d}(kv, kw)$ für alle $k \in K \setminus \{0\}$

Beweis zu 44:
Die ersten beiden Eigenschaften folgen unmittelbar aus der Definition. Die Dreiecksungleichung sieht man aus der Transitivität der Gleichheit: Wenn $u_i \neq w_i$, dann gilt $u_i \neq v_i$ oder $v_i \neq w_i$. Offensichtlich gilt $\mathrm{d}(v, w) \geqslant \mathrm{d}(f(v), f(w))$ für eine beliebige komponentenweise definierte Abbildung $f : H(n, A) \to H(n, A)$, also

$$\mathrm{d}(v, w) \geqslant \mathrm{d}(f(v), f(w)) \geqslant \mathrm{d}(f^{-1}(f(v)), f^{-1}(f(w))) = \mathrm{d}(v, w)$$

für bijektive solche f. Damit folgt die Invarianz unter Translationen und unter Skalarmultiplikation, da die Abbildungen $v \mapsto v + u$ und $v \mapsto k \cdot v$ für $k \neq 0$ bijektiv sind (Umkehrabbildungen sind $v \mapsto v - u$ und $v \mapsto k^{-1} \cdot v$).

Bemerkung:
Während die übliche euklidische Metrik $\|v - w\|$ im $\mathbb{R}^n$ ebenfalls translationsinvariant ist, gilt dort $\|rv - rw\| = |r| \cdot \|v - w\|$. Die Invarianz der Hamming-Metrik unter Skalarmultiplikation ist also eine „ungeometrische“ Eigenschaft.

Satz 45 Falls $(A, +)$ eine kommutative Gruppe ist und p eine Primzahl, dann ist A genau dann ein $\mathbb{F}_p$-Vektorraum, wenn $\underbrace{a + \cdots + a}_{p \text{ mal}} = 0$ für alle $a \in A$ gilt. Die Skalarmultiplikation ist dann durch $m \cdot a = \underbrace{a + \cdots + a}_{m \text{ mal}}$ gegeben und es gilt $-a = (p - 1) \cdot a$.
Insbesondere folgt daraus: Wenn $C \subseteq \mathbb{F}_p^n$ unter Addition abgeschlossen ist, dann ist C bereits ein Untervektorraum!

Beweis zu 45:
Nachrechnen. Der Beweis folgt auch aus Satz **??** im Kapitel II.

Definition: Codes und ihre Eigenschaften
(a) Ein *Code* ist eine Teilmenge von $H(n,A)$ bzw. $H(n,q)$. Man spricht von einem „*Code der Länge n über A*“ bzw. einem „*q-ären Code der Länge n*“. Der *Minimalabstand* des Codes ist $\min\big\{\mathrm{d}(v,w) \;\big|\; v,w \in C, v \neq w\big\}$.
(b) Ein *linearer Code* ist ein Untervektorraum von $\mathbb{F}_q^n$. Das *Gewicht* von $v \in C$ ist $\mathrm{d}(v,0)$ und das *Minimalgewicht* des Codes ist $\min\big\{\mathrm{d}(v,0) \;\big|\; v \in C, v \neq 0\big\}$.
(c) Ein Code C *erkennt* (mindestens) e *Fehler*, falls der Minimalabstand größer als e ist.
(d) Ein Code C *korrigiert* (mindestens) e *Fehler*, falls es zu jedem $v \in H(n,A)$ höchstens ein $c \in C$ gibt mit $\mathrm{d}(v,c) \leqslant e$.

Notation: Beschreibung linearer Codes
Lineare Codes werden meist durch zwei oder drei Parameter beschrieben: als „(q-äre) $[n,k]$-*Codes*“ oder „(q-äre) $[n,k,d]$-*Codes*“. Dabei ist n die Länge der Wörter, $k = \dim C$ und d das Minimalgewicht. Es gilt dann $|C| = q^k$ bzw. $k = \log_q |C|$. [2]

Bemerkung:
Wegen $\mathrm{d}(v,w) = \mathrm{d}(v-w,0)$ ist das Minimalgewicht eines linearen Codes gleich seinem Minimalabstand. Unmittelbar aus der Definition sieht man auch:

Satz 46
(a) Ein Code mit Minimalabstand d erkennt $d-1$ Fehler und korrigiert $\left\lfloor \frac{d-1}{2} \right\rfloor$ Fehler.
(b) Ein Code, der e Fehler korrigiert, erkennt $2e$ Fehler und hat Minimalabstand mindestens $2e+1$.

Beispiele

- Der alte ISBN-Code ist ein 11-ärer Code der Länge 10, der einen Fehler erkennt und keinen korrigiert.
- Der ursprüngliche ASCII-Code ist ein binärer linearer $[8,7,2]$-Code, der also einen Fehler erkennt und keinen korrigiert.
- Der Wiederholungscode $\{(x,x,x) \mid x \in \mathbb{F}_q\} \subseteq H(3,q)$ ist ein q-ärer linearer $[3,1,3]$-Code, der zwei Fehler erkennt und einen korrigiert.

[2] Manche Autoren bevorzugen, statt der Dimension eines Codes C an zweiter Stelle die Anzahl der Elemente von C anzugeben.

Definition: Ball
Der *Ball vom Radius* e um v ist [3]

$$B_e(v) := \{w \in H(n,A) \mid \mathrm{d}(v,w) \leqslant e\}.$$

Satz 47 Die Anzahl der Elemente eines Balls kann man ausrechnen durch

$$|B_e(v)| = \sum_{i=0}^{e} \binom{n}{i}(q-1)^i.$$

Für $q = 2$ gilt insbesondere

$$|B_e(c)| = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{e}.$$

Beweis zu 47:
i durchläuft die möglichen Abstände zu v; der Binomialkoeffizient gibt die Anzahl der Möglichkeiten für die i Stellen, an denen die Abweichungen auftreten; $q-1$ ist für jede Stelle die Anzahl der alternativen Elemente des Alphabets.

Satz 48 Es gilt nun offensichtlich:

- C erkennt genau dann e Fehler, wenn $c' \notin B_e(c)$ für $c, c' \in C$, $c \neq c'$.
- C korrigiert genau dann e Fehler, wenn die Bälle $B_e(c)$ für $c \in C$ paarweise disjunkt sind.

3.2. Gütekriterien und Schranken für Codes

Zur Motivation: Zwei Beispiele für einen 1-fehlerkorrigierenden Code
Ausgangslage: Man hat als eigentliche Information Wörter der Länge 4 über $\mathbb{F}_2$ (also etwa die Binärdarstellung von hexadezimalen Zeichen). Man möchte den Code nun z. B. durch Anhängen von Prüfziffern so verändern, dass er einen Fehler korrigiert.

Erster Code Die „naive“ Methode besteht darin, das Ausgangswort dreifach zu senden. Wörter aus $H(4,\mathbb{F}_2)$ werden also codiert als Wörter in $H(12,\mathbb{F}_2)$, nämlich $v = (v_1, v_2, v_3, v_4)$ als $v^\frown v^\frown v := (v_1, v_2, v_3, v_4, v_1, v_2, v_3, v_4, v_1, v_2, v_3, v_4)$.

[3] In der Analysis sind Bälle üblicherweise als offene Bälle definiert, d. h. man fordert „$< e$“ statt „$\leqslant e$“. Dies ist in der diskreten Situation hier nicht besonders sinnvoll.

$C_1 = \left\{ v^\frown v^\frown v \mid v \in H(4, \mathbb{F}_2) \right\}$ ist dann ein binärer $[12, 4, 3]$-Code: Die Wortlänge ist 12, die Dimension 4 und das Minimalgewicht 3, d. h. der Code erkennt zwei Fehler und korrigiert einen.

Dieser Code ist aber nicht besonders effizient: die Raumgröße ist $|H(12, \mathbb{F}_2)| = 2^{12} = 4.096$. Es gibt 16 Codewörter, die mit Ihren „Korrekturbereichen“ einen Platz von $16 \cdot |B_1(c)| = 16 \cdot 13 = 208$ einnehmen. Es gibt also einen „verschwendeten Platz“ von $4.096 - 208 = 3.888$ Wörtern.

Zweiter Code C_2 besteht aus folgenden Wörtern in $H(7, \mathbb{F}_2)$:

$$\begin{array}{cccc}
(0,0,0,0,0,0,0) & (0,1,0,0,1,0,1) & (1,0,0,0,0,1,1) & (1,1,0,0,1,1,0) \\
(0,0,0,1,1,1,1) & (0,1,0,1,0,1,0) & (1,0,0,1,1,0,0) & (1,1,0,1,0,0,1) \\
(0,0,1,0,1,1,0) & (0,1,1,0,0,1,1) & (1,0,1,0,1,0,1) & (1,1,1,0,0,0,0) \\
(0,0,1,1,0,0,1) & (0,1,1,1,1,0,0) & (1,0,1,1,0,1,0) & (1,1,1,1,1,1,1)
\end{array}$$

Ein Wort v aus $H(4, \mathbb{F}_2)$ wird codiert durch dasjenige Wort aus $H(7, \mathbb{F}_2)$ in der Liste, dessen Anfangsstück gerade v ist. Man kann nun überprüfen, dass C_2 ein binärer $[7, 4, 3]$-Code ist. Der Code erkennt also ebenfalls zwei Fehler und korrigiert einen, bei gleicher Anzahl von Codewörtern (d. h. bei gleicher Dimension 4).

Die Raumgröße ist hier aber $|H(7, \mathbb{F}_2)| = 2^7 = 128$. Die 16 Codewörter nehmen mit Ihren „Korrekturbereichen“ einen Platz von $16 \cdot |B_1(c)| = 16 \cdot 8 = 128$ ein, d. h. es gibt keinen verschwendeten Platz. Solche Codes heißen *perfekte Codes*.

C_2 ist übrigens ein Beispiel für einen *Hamming-Code*. Im folgenden wird erklärt werden, wie man C_2 systematisch konstruieren kann und wie Codierung und Decodierung funktionieren. Denn C_1 hat gegenüber C_2 zunächst den Vorteil, dass die Codierungs- und Decodierungsschritte offensichtlich sind, während man bei C_2 in der Tabelle nachschauen muss.

Definition: Anforderungen an einen guten Code

Ein guter Code sollte

- möglichst viele Fehler erkennen und korrigieren, d. h. großen Minimalabstand haben;
- möglichst viele Codewörter im Verhältnis zur Wortlänge n haben;
- und dabei eine effiziente Codierung (Verschlüsselung), Decodierung (Entschlüsselung) und ggf. Fehlerkorrektur gestatten.

Erläuterung

Für Codierung und Decodierung gibt es immer die Möglichkeit, eine Codierungstafel aufzustellen. Für die Entschlüsselung eines fehlerhaft übertragenen Worts muss dann in der Tafel nach dem Wort im Code gesucht werden, das den kleinsten Hamming-Abstand zum übertragenen Wort hat (wenn man davon ausgeht, dass höchstens so viele Fehler aufgetreten sind, wie der Code korrigieren kann). Bei einem großen Hamming-Raum ist dies aber ein eher langwieriger Algorithmen. Schnelle Algorithmen setzen voraus, dass der Code eine interne Struktur besitzt. Daher sind lineare Codes interessant.

Die ersten beiden Anforderungen laufen einander zuwider: Redundanzen (Prüfziffern) erhöhen die Wortlänge. Es gibt daher Schranken für das Verhältnis von Codegröße und Minimalabstand bei gegebenen Hamming-Raum.

Satz 49 [Die Hamming-Schranke] Die Anzahl der Codewörter eines q-ären Codes der Länge n mit Mindestabstand $\geqslant d$ ist höchstens

$$\frac{q^n}{\sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} \cdot (q-1)^i}$$

Beweis zu 49:
Die Schranke folgt sofort aus der Formel für die Anzahl der Elemente von $|B_e(c)|$ und der Größe des Hamming-Raumes $|H(n,q)| = q^n$.

Definition: perfekter Code
Ein Code heißt *perfekt*, wenn er die Hamming-Schranke erreicht.

Beispiele

- $q = 2, n = 7, d = 3$: Hier ergibt die Hamming-Schranke $2^7/(1+7) = 16$. Der Hamming-Code C_2 im Beispiel oben erreicht als perfekter Code diese Schranke.
- $q = 2, n = 6$: Die Folge der Binomialkoeffizienten $\binom{6}{i}$ ist 1, 6, 15, 20, 15, 6, 1. Keine der Summen $\binom{6}{0} + \ldots \binom{6}{e}$ ist ein Teiler von $2^6 = 64$ außer für $e = 0$ und $e = 6$. Diese entsprechen den sogenannten *trivialen Codes*: es sind alle Wörter Codewörter (bei Minimalabstand 1) oder es gibt überhaupt nur ein Codewort (bei Minimalabstand ∞). Beide Codes sind perfekt, aber aus Sicht der Codierungstheorie vollkommen uninteressant. Für die Länge 6 gibt es also keine nicht-trivialen perfekten binären Codes.

Satz 50 [Die Gilbert-Schranke[4]] Gegeben q, n, d, so gibt es einen q-ären Code der Länge n und vom Minimalabstand mindestens d mit mindestens

$$\frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} \cdot (q-1)^i}$$

Codewörtern. Ist q eine Primzahlpotenz, so kann man den Code linear über $\mathbb{F}_q$ wählen.

Beweis zu 50:

Sei C ein Code vom Minimalabstand $\geqslant d$, so dass $|C|$ kleiner als die Gilbert-Schranke ist. Dann gibt es ein $x \in H(n,q)$, welches zu allen $c \in C$ mindestens Abstand d hat, denn nach Annahme gilt $|C| \cdot |B_{d-1}(c)| < q^n = |H(n,q)|$. (Die Größe der Bälle $B_{d-1}(c)$ hängt nicht von c ab!) Dann ist $C \cup \{x\}$ ein größerer Code vom Minimalabstand $\geqslant d$. Durch sukzessives Vergrößern erhält man also einen Code, der die Gilbert-Schranke erfüllt.

Für den linearen Fall nimmt man an, dass $C \subseteq H(n, \mathbb{F}_q)$ bereits ein linearer Code ist (z. B. der triviale Code $\{0\}$) und wählt x wie oben. Statt $C \cup \{x\}$ betrachtet man nun den erzeugten linearen Code $\langle x, C\rangle$, muss aber noch zeigen, dass dieser weiterhin Mindestgewicht $\geqslant d$ hat.

Ein typisches Element darin hat die Form $kx + c$ mit $k \in \mathbb{F}_q$ und $c \in C$.
– Falls $k = 0$, so ist $\mathrm{d}(kx + c, 0) = \mathrm{d}(c, 0) \geqslant d$ nach Annahme an C.
– Falls $k \neq 0$, so ist $\mathrm{d}(kx + c, 0) = \mathrm{d}(x, -\frac{1}{k}c) \geqslant d$ nach Wahl von x, da $\frac{1}{k}c \in C$.

Beispiele

Für $q = 2, n = 7, d = 3$ ergibt die Gilbert-Schranke $2^7/(1 + 7 + 21) \approx 4,41$. Die Gilbert-Schranke stellt also die Existenz eines Codes C vom Minimalabstand 3 mit mindestens 5 Codewörtern sicher. Im linearen Fall weiß man, dass die Anzahl der Element von C als Untervektorraum von $\mathbb{F}_2^7$ eine Zweierpotenz sein muss, also erhält man $|C| \geqslant 8$. Aus dem obigen Beispiel wissen wir aber, dass es sogar den Hamming-Code mit 16 Wörtern gibt.

3.3. Erzeuger- und Prüfmatrizen

Sei C nun ein q-ärer $[n, k]$-Code, also ein k-dimensionaler Unterraum von $H(n,q) = \mathbb{F}_q^n$.

Definition: Erzeugermatrix

Eine *Erzeugermatrix* G für einen linearen $[n, k]$-Code C ist eine $(k \times n)$-Matrix, deren Zeilen eine Basis von C bilden.

Erläuterung

Eine Erzeugermatrix eines Codes ist nicht eindeutig bestimmt. Man kann sie aber durch elementare Umformungen auf die Form

$$\left(\begin{array}{c|c} \mathrm{Id}_k & A \end{array}\right)$$

bringen, wobei A eine $(k \times (n-k))$-Matrix ist. Im allgemeinen wird der Code durch solche Umformungen verändert und durch einen *äquivalenten* Code ersetzt (d. h. man betrachtet das Bild des Codes unter einem Automorphismus des Vektorraums $H(n,q)$). Äquivalente Codes haben zwar u. U. andere Codewörter, aber dieselben Parameter: Anzahl der Codewörter, Dimension, Minimalabstand.

[4]Edgar Gilbert (1923–2013)

Wir werden auch sehen, dass diese spezielle Form der Erzeugermatrix einer Codierung durch Anhängen von Prüfziffern entspricht, also einer sehr üblichen Art von Codes.

Beispiele

Der im vorherigen Abschnitt angegebene $[7,4,3]$-Hamming-Codes hat mit

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

eine Erzeugermatrix in der Form $(\mathrm{Id}_k \mid A)$.

Bemerkung:

Man sieht hier leicht, dass die Basisvektoren ein Gewicht und paarweise einen Abstand von mindestens 3 haben. Dies ist natürlich eine notwendige Bedingung für einen Minimalabstand von mindestens 3, aber keine hinreichende: Es reicht nicht um zu folgern, dass der erzeugte Code Minimalabstand mindestens 3 hat. Zum Beispiel haben die Vektoren $(1,1,1,0,0,0)$, $(0,0,0,1,1,1)$ und $(1,1,0,1,1,0)$ Gewicht und paarweisen Abstand $\geqslant 3$, der von ihnen erzeugte Code hat aber nur Minimalgewicht 2 (betrachte die Summe der drei Vektoren!). Man kann nur, wie im Beweis der Gilbert-Schranke, von einem Vektor v, der zu einem Untervektorraum U einen Minimalabstand hat, auf den Minimalabstand des von v und U erzeugten Untervektorraums schließen.

Notation:

Elemente eines Hamming-Raums fasse ich als Zeilenvektoren auf; v^{T} ist dann der zum Zeilenvektor v gehörende Spaltenvektor.

Definition: Codierung mit Hilfe der Erzeugermatrix:

Die Codierung eines (Zeilen-)Vektors $v \in H(k,q)$ erfolgt nun durch

$$v \cdot G = (G^{\mathrm{T}} \cdot v^{\mathrm{T}})^{\mathrm{T}} \in H(n,q).$$

Hat G die besondere Form $(\mathrm{Id}_k \mid A)$, so entsteht der Codevektor also durch das Anhängen der $n-k$ Prüfziffern $v \cdot A$, da $v \cdot G$ die Form $(v \cdot \mathrm{Id}_k)^\frown(v \cdot A) = v^\frown w$ für einen Vektor w der Länge $n-k$ hat. Die Prüfziffern erhält man als Linearkombination der Prüfziffern der Basiselemente.

Beispiele

Im angegebenen Beispiel des $[7,4,3]$-Hamming-Codes wird etwa der Vektor $v = (1,0,1,1)$, den man als Darstellung der Binärzahl 1011 bzw. der Hexadezimalzahl B auffassen kann, durch $(1,0,1,1) \cdot G = (1,0,1,1,0,1,0)$ codiert. Der Vektor schreibt sich als $e_1 + e_3 + e_4$, demgemäß ergeben sich die Prüfziffern für v als die analoge Linearkombination $(0,1,1) + (1,1,0) + (1,1,1)$ der Prüfziffern der Standardbasisvektoren.

Satz 51 Die Codierungsabbildung ist injektiv.

Beweis zu 51:
Im Falle des Anhängens von Prüfziffern ist dies trivialerweise gegeben; da jeder Code zu einem solchen äquivalent ist, also durch Isomorphie dazu übergeht, gilt es auch allgemein.
Alternativ: Die Zeilen von G sind linear unabhängig, also gilt

$$\operatorname{rg}(G) = \operatorname{rg}(G^{\mathrm{T}}) = \dim \operatorname{Bild}(G^{\mathrm{T}}) = k$$

und somit $\dim \operatorname{Kern}(G^{\mathrm{T}}) = \dim H(k, q) - \dim \operatorname{Bild}(G^{\mathrm{T}}) = k - k = 0$.

Definition: Prüfmatrix
Eine *Prüfmatrix* oder (*Kontrollmatrix*) H für einen $[n, k]$-Code C ist eine $((n-k) \times n)$-Matrix, für die $C = \operatorname{Kern}(H)$ gilt.

Erläuterung
Mit der Prüfmatrix kann man also die Kontrollrechnung ausführen: Gilt $H \cdot v = 0$, so liegt v im Code, andernfalls nicht.

Satz 52 Die folgenden Aussagen über einen linearen $[n, k]$-Code C und eine $((n-k) \times n)$-Matrix H sind äquivalent:
(a) H ist eine Prüfmatrix von C.
(b) Es gilt $H \cdot c^{\mathrm{T}} = 0$ für alle $c \in C$ und die Zeilen von H sind linear unabhängig.
(c) Es gilt $G \cdot H^{\mathrm{T}} = 0$ und die Zeilen von H sind linear unabhängig.

Beweis zu äquivalente Definitionen der Prüfmatrix:
$G \cdot H^{\mathrm{T}} = 0$ ist äquivalent mit $H \cdot G^{\mathrm{T}} = 0$ und impliziert $H \cdot c^{\mathrm{T}} = 0$ für alle $c \in C$, da jedes $c \in C$ Linearkombination von Zeilen von G ist. Beides bedeutet also, dass C im Kern von H liegt.
Die lineare Unabhängigkeit der Zeilen von H ist gleichbedeutend mit $n - k = \operatorname{rg}(H) = \dim \operatorname{Bild}(H)$ und damit, wegen $\dim \operatorname{Kern}(H) = n - \dim \operatorname{Bild}(H)$, gleichbedeutend mit $\dim \operatorname{Kern}(H) = n - (n - k) = k = \dim C$.
Zusammen sind beide Bedingungen äquivalent mit $C = \operatorname{Kern}(H)$.

Satz 53 Genau dann sind G und H Erzeuger- und Prüfmatrix eines $[n, k]$-Codes, wenn G eine $(k \times n)$-Matrix vom Rang k und H eine $((n-k) \times n)$-Matrix vom Rang $(n - k)$ ist, für die $G \cdot H^{\mathrm{T}} = 0$ ist.

Beweis zu 53:
Erzeuger- und Prüfmatrix haben nach Definition und Satz 52 diese Eigenschaften. Umgekehrt kann es eine $(k \times n)$-Matrix G vom Rang k nur für $k \leqslant n$ geben; solch eine Matrix ist dann per Definition Erzeugermatrix des Codes $\mathrm{Bild}(G^{\mathrm{T}})$. H ist dann wieder nach Satz 52 eine zugehörige Prüfmatrix.

Folgerung 54 Wenn eine Erzeugermatrix G die Form $(\mathrm{Id}_k \mid A)$ hat, so ist

$$H = \left(\begin{array}{c|c} -A^{\mathrm{T}} & \mathrm{Id}_{n-k} \end{array} \right)$$

eine zugehörige Prüfmatrix.

Beweis zu 54:
Wenn also

$$G = \begin{pmatrix} 1 & 0 & \dots\dots & 0 & a_{11} & \dots & a_{1\,n-k} \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & & \vdots \\ \vdots & & \ddots & \ddots\ 0 & \vdots & & \vdots \\ 0 & \dots\dots & 0 & 1 & a_{k1} & \dots & a_{k\,n-k} \end{pmatrix},$$

dann ist

$$H = \begin{pmatrix} -a_{11} & \dots & -a_{k1} & 1 & 0 & \dots & \dots & 0 \\ \vdots & & \vdots & 0 & \ddots & \ddots & & \vdots \\ \vdots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & \ddots & 0 \\ -a_{1\,n-k} & \dots & -a_{k\,n-k} & 0 & \dots & \dots & 0 & 1 \end{pmatrix},$$

denn man sieht leicht, dass dann $G \cdot H^{\mathrm{T}} = 0$, und durch die jeweiligen Identitätsmatrizen in G und H sieht man auch, dass die Bedingungen an den Rang erfüllt sind.

Bemerkung:
Der Code C ist jeweils durch G und durch H festgelegt; umgekehrt sind G und H aber nicht eindeutig durch C bestimmt. In der speziellen Form sind sie zwar durch C festgelegt, für äquivalente Codes sind sie aber im allgemeinen verschieden.

Beispiele
Im Falle des $[7, 4, 3]$-Hamming-Codes und der Erzeugermatrix G von oben ist

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

die passende Prüfmatrix.

Decodierung mit Hilfe der Prüfmatrix

Angenommen $w \in H(n,q)$ wird empfangen. Als Decodierung wird dasjenige $c \in C$ gesucht, welches minimalen Hamming-Abstand zu w hat, und dann das Urbild von c unter der Codierung (die ja injektiv ist) bestimmt.

Um die Existenz von c sicherzustellen, nehmen wir an, dass entweder ein perfekter, e-fehlerkorrigierender Code vorliegt, oder dass höchstens e Übertragungsfehler vorgekommen sind, wobei $2e+1$ höchstens so groß wie das Minimalgewicht d des Codes sein darf. Es gilt aufgrund dieser Annahmen, dass sich w als $w = c + f$ schreiben lässt mit $c \in C$ und einem Fehler f mit $\mathrm{d}(f, 0) \leqslant e$.

Man berechnet nun zunächst das sogenannte *Syndrom* von w, das ist $H \cdot w^{\mathrm{T}}$. Es gilt dafür

$$H \cdot w^{\mathrm{T}} = H \cdot (c+f)^{\mathrm{T}} = H \cdot c^{\mathrm{T}} + H \cdot f^{\mathrm{T}} = 0 + H \cdot f^{\mathrm{T}} = H \cdot f^{\mathrm{T}},$$

d. h. das Syndrom von w ist gleich dem Syndrom des Fehlers f.

Für zwei mögliche Fehler f, f' gilt zudem

$$\mathrm{d}(f - f', 0) = \mathrm{d}(f, f') \leqslant \mathrm{d}(f, 0) + \mathrm{d}(0, f') \leqslant e + e < d,$$

also ist $f - f' \notin C$ und somit $H \cdot f^{\mathrm{T}} - H \cdot f'^{\mathrm{T}} = H \cdot (f - f')^{\mathrm{T}} \neq 0$. Verschiedene Fehler haben also verschiedene Syndrome.

Man kann nun eine Liste der Syndrome der möglichen Fehler aufstellen. Dies sind $|B_e(0)|$ viele; eine im Vergleich mit $|H(n,q)|$ deutlich kleinere Zahl. Die Decodierung geht dann folgendermaßen: Man berechnet das Syndrom $H \cdot w^{\mathrm{T}}$, schaut in der Tabelle nach, welchem Fehler f es entspricht, korrigiert den Fehler und erhält so das zugehörige Codewort $c \in C$. Zu diesem Codewort kann man nun das Urbild unter der Codierungsabbildung bestimmen; hat G die besondere Form $(\mathrm{Id}_k \mid A)$, dann erhält man das Urbild einfach durch das Weglassen der Prüfziffern.

Im Falle der Hamming-Codes ist das Decodierungsverfahren sogar noch einfacher: siehe unten.

Auf dem Weg zur Definition der Hamming-Codes ist der folgende Satz essentiell, der aussagt, dass man das Minimalgewicht des Codes unmittelbar der Prüfmatrix ablesen kann:

Satz 55 Ein linearer Code C hat genau dann Minimalgewicht mindestens d, wenn je $d-1$ Spalten der Prüfmatrix linear unabhängig sind.

Beweis zu Prüfmatrix und Minimalgewicht des Codes:

Eine Linearkombination $0 = \sum_{i=0}^{n} a_i S_i$ der Spalten S_i von H entspricht gerade einem Vektor $a = (a_1, \ldots, a_n)$, für welchen $H \cdot a = 0$ gilt, also einem $a \in \mathrm{Kern}(H) = C$. Die Anzahl der Komponenten $a_i \neq 0$ in a, also der tatsächlich vorkommenden Spalten in der Linearkombination, ist gerade das Gewicht von a.

Folgerung 56 Insbesondere hat also ein Code Minimalgewicht mindestens 3, wenn je zwei Spalten der Prüfmatrix linear unabhängig sind, d. h. wenn keine null ist und keine das skalare Vielfache einer anderen ist.

Definition: Hamming-Code
Ein *Hamming-Code* ist ein linearer Code C vom Minimalgewicht 3, dessen Prüfmatrix zu gegebener Zeilenanzahl die maximale Anzahl von Spalten hat.

Erläuterung
Die zu einem gegebenen Vektor v linear abhängigen Vektoren sind gerade die Elemente des von v erzeugten Untervektorraums. Ist m die Anzahl der Zeilen der Prüfmatrix H, so bilden die Spalten von H ein den Nullvektor nicht enthaltendes Repräsentantensystem der eindimensionalen Untervektorräume von $\mathbb{F}_q^m$, d. h. keine Spalte von H ist die Nullspalte und für jeden Vektor $v \in \mathbb{F}_q^m, v \neq 0$ gibt es genau einen Spaltenvektor von H, der ein skalares Vielfaches von v ist.

Die Anzahl der Vektoren $\neq 0$ in $\mathbb{F}_q^m$ ist $q^m - 1$. Da es $q-1$ Skalarfaktoren $\neq 0$ gibt, gibt es also $\frac{q^m-1}{q-1}$ eindimensionale Untervektorräume von $\mathbb{F}_q^m$, d. h. der zugehörige Hamming-Code besteht aus Wörtern der Länge $n = \frac{q^m-1}{q-1}$. Die Dimension des Hamming-Codes ist dann $k = \frac{q^m-1}{q-1} - m$.

Satz 57 Hamming-Codes sind perfekt.

Beweis zu 57:
Es ist also $n = \frac{q^m-1}{q-1}$ und $k = \frac{q^m-1}{q-1} - m$ und man rechnet nach, dass

$$|C| \cdot |B_1(c)| = q^{\frac{q^m-1}{q-1}-m} \cdot \big(1 + \tfrac{q^m-1}{q-1} \cdot (q-1)\big) = q^{\frac{q^m-1}{q-1}-m} \cdot q^m = q^{\frac{q^m-1}{q-1}} = |H(\tfrac{q^m-1}{q-1}, \mathbb{F}_q)|.$$

Spezialfall $q = 2$:
Hier ist die Situation besonders einfach: Da die eindimensionalen Untervektorräume jeweils aus zwei Vektoren bestehen – dem Nullvektor und einem anderen Vektor – treten sämtliche Vektoren in $\mathbb{F}_q^m \setminus \{0\}$ als Spaltenvektoren von H auf. Ihre Anzahl n ist $2^m - 1$, die Dimension des Codes ist $2^m - (m+1)$. Alle binären Hamming-Codes fester Länge sind außerdem äquivalent.

Auch die Decodierung ist im Falle $q = 2$ besonders einfach: Die möglichen Fehler sind gerade die Standardbasisvektoren $e_1, \ldots, e_n$ in $\mathbb{F}_2^n$. Das Syndrom $H \cdot e_i^{\mathrm{T}}$ von e_i ist dann gerade die i-te Spalte von H. Zur Decodierung von w berechnet man also das Syndrom $H \cdot w^{\mathrm{T}}$. Ist das Syndrom 0, so ist kein Fehler aufgetreten. Andernfalls schaut man nach, in welcher Spalte i von H das Syndrom auftritt, ändert das i-te Bit von w und lässt die Prüfziffern, d. h. die letzten $n-k$ Stellen von w, weg.

3.4. Liste der perfekten Codes

Ist q eine Primzahlpotenz, so gibt es die folgenden perfekten q-ären Codes (wobei e die Anzahl der korrigierbaren Fehler bezeichnet):

- Triviale Codes, die nur aus einem Wort bestehen.
  (nimmt man dafür den Nullvektor, so ist es ein $[n, 0, \infty]$-Code mit $e = n$)
- Der trivialen $[n, n, 1]$-Code mit $e = 0$ (alle Wörter sind im Code).
- Die q-ären Hamming-Codes: $[\frac{q^m-1}{q-1}, \frac{q^m-1}{q-1} - m, 3]$-Codes mit $e = 1$.
  Außerdem einige nicht-lineare Codes mit gleichen Parametern wie Hamming-Codes.
- Die binären Wiederholungscodes ungerader Länge:
  zu jedem $e \in \mathbb{N}$ der $[2e + 1, 1, 2e + 1]$-Code, der nur die beiden Wörter $(0, 0, \ldots, 0)$ und $(1, 1, \ldots, 1)$ enthält.
- Der binäre Golay-Code[5]: ein $[23, 12, 7]$-Code mit $e = 3$.
- Der ternäre Golay-Code: ein $[11, 6, 5]$-Code mit $e = 2$.

Der binäre Wiederholungscode stimmt für $e = 0$ mit dem trivialen $[1, 1, 1]$-Code und für $e = 1$ mit dem $[3, 1, 3]$-Hamming-Code überein.

Falls q keine Primzahlpotenz ist, so weiß man nicht, ob es perfekte nicht-triviale q-äre Codes gibt. Nur für einige wenige Werte weiß man, dass keine perfekten q-ären Codes existieren außer den trivialen. Im allgemeinen weiß man auch wenig darüber, welches die (hinsichtlich der „Packungsdichte“) „besten“ Codes sind.

[5] Marcel Golay (1902-1989)